NICE National Institute for Health and Care Excellence

Single Technology Appraisal

Pembrolizumab for treating relapsed or refractory classical Hodgkin lymphoma [Review of TA540] [ID5084]

Committee Papers

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

SINGLE TECHNOLOGY APPRAISAL

Pembrolizumab for treating relapsed or refractory classical Hodgkin lymphoma [Review of TA540] [ID5084]

Contents:

The following documents are made available to stakeholders:

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

1. **Company submission** from MSD
   a. **Company summary of information for patients (SIP)**
2. **Clarification questions and company responses**
3. **Patient group, professional group and NHS organisation submissions** from:
   a. Lymphoma Action
   b. Royal College of Pathologists
   c. SACT report
4. **Expert personal perspectives** from:
   a. Dr Patrick Medd, Consultant Haematologist – clinical expert, nominated by RCP-ACP-RCR
   b. Dr Beth Phillips, Clinical Senior Lecturer/Hon. Consultant Haematologist – clinical expert, nominated by RCP-ACP-RCR
5. **External Assessment Report** prepared by Kleijnen Systematic Reviews (KSR)
6. **External Assessment Report – factual accuracy check**

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

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single technology appraisal

Pembrolizumab for treating relapsed or refractory classical Hodgkin lymphoma [review of TA540]

Document B

Company evidence submission

September 2023

File name	Version	Contains confidential information	Date
ID5084_Pembrolizumab_NICE_STA_Document B_CIC	V1.0	Yes (CiC)	20th September 2023

Contents

List of tables

List of figures

Abbreviations

Abbreviation	Definition
ABVD	Doxorubicin, bleomycin, vinblastine, dacarbazine
AE	Adverse event
AESI	Adverse events of special interest
AIC	Akaike Information Criterion
AlloSCT	Allogenic stem cell transplant
ASaT	All subjects as treated
AutoSCT	Autologous stem cell transplant
BEACOPP	Bleomycin, etoposide, doxorubicin, cyclophosphamide, vincristine, procarbazine and prednisone
BIC	Bayesian information criterion
BICR	Blinded independent central review
BNF	British National Formulary
BSC	Best supportive care
BV	Brentuximab vedotin
CDF	Cancer Drugs Fund
cHL	Classical Hodgkin Lymphoma
CHOP	Cyclophosphamide, doxorubicin, prednisolone, vincristine
CI	Confidence interval
CR	Complete remission
CSR	Clinical Study Report
CT	Computed tomography
CTLA-4	Cytotoxic T-lymphocyte antigen 4
DoR	Duration of response
ECOG	Eastern Cooperative Oncology Group
EMA	European Medicines Agency
EORTC-QLQC30	European Organisation for Research and Treatment Cancer Quality of Life Questionnaire
EQ-5D	EuroQoL 5 Dimensions
FAD	Final appraisal document

GVHD	Graft versus host disease
HIV	Human immunodeficiency virus
HL	Hodgkin Lymphoma
HR	Hazard ratio
HRQoL	Health-related quality of life
ICER	Incremental cost-effectiveness ratio
ITT	Intention-to-treat
IV	Intravenous
KM	Kaplan–Meier
LY	Life year
mAB	Monoclonal antibody
MAIC	Matching adjusted indirect treatment comparison
MSD	Merck Sharp & Dohme Ltd
NHS	National Health Service
NICE	National Institute for Health and Care Excellence
NMA	Network Meta-Analysis
ORR	Overall response rate
OS	Overall survival
PD	Progressed disease
PD-L1	Programmed death-ligand 1
PET	Positron emission tomography
PFS	Progression-free survival
PR	Partial remission
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PSS	Personal Social Services
Q3W	Every 3 weeks
Q6W	Every 6 weeks
QALY	Quality adjusted life year
QoL	Quality of life
RECIST	Response Evaluation Criteria in Solid Tumours

R/R cHL	Relapsed or refractory classical Hodgkin lymphoma
SACT	Systemic anti-cancer therapy
SAE	Serious adverse event
SD	Stable disease
SEE	Structured expert elicitation
SLR	Systemic literature review
SmPC	Summary of Product Characteristics
SoC	Standard of care
STA	Single Technology Appraisal
ToT	Time on treatment
UK	United Kingdom of Great Britain and Northern Ireland

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

Summary of the decision problem, technology, and clinical care pathway

Decision problem

- Pembrolizumab is proposed as an option for the management of adults with relapsed or refractory classical Hodgkin Lymphoma (R/R cHL) who have relapsed after or not responded to treatment with brentuximab vedotin (BV) and are ineligible for autologous stem cell transplant (autoSCT).
- The indication specified in the decision problem is narrower than the marketing authorisation for pembrolizumab in cHL.
- MSD's submission aligns with the final scope issued by NICE in most areas, deviating from the scope only in the composition of standard of care (SoC).

Technology

- Pembrolizumab (KEYTRUDA®) is a humanised monoclonal antibody (mAb) against the programmed death-1 (PD-1) receptor.
- Expression of PD-1 protein, and its ligands PD-L1 and PD-L2, triggers a signalling cascade that culminates in the suppression of T cell proliferation, cytokine release and cytotoxicity, a process that modulates the immune response (to prevent destruction of healthy cells).

Health condition

- cHL is a neoplastic disease that affects the lymphatic system, which is an important part of the body's immune system.
- cHL originates from uncontrolled proliferation of B lymphocytes (white blood cells) in predominantly lymph node tissues and is characterised by the presence of Reed–Sternberg cells.
- The incidence of HL follows a bimodal age distribution, with the first peak in young adults (20–24 years) and the second occurring in the elderly (75–79 years).
- In 2020, 1,525 new cases of cHL were reported in England.
- Data from pembrolizumab's time in the CDF indicate that there were approximately 50 unique applications annually for use of pembrolizumab for

patients with R/R cHL who have been treated with BV and remain ineligible for autoSCT.

Clinical care pathway

- Cure rates for cHL are considered high, with some people achieving cure or long-term remission after standard first-line chemotherapy and radiotherapy.
- However, those who experience relapse or are refractory to treatment typically have a poor prognosis, and there is an unmet need for the management of those with R/R cHL.
- Patients with primary refractory disease have been reported to have a median overall survival (OS) of 19 months compared with 27 months for patients who achieved a complete response to their primary treatment.
- The current standard of care (SoC) for patients with R/R cHL who remain ineligible for autoSCT after BV comprises various generic chemotherapeutic options, with choice dependent on individual patient factors.
- No equity or equality considerations are anticipated.

B.1.1. Decision problem

The submission focuses on part of the technology's marketing authorisation. The proposed use of pembrolizumab in the management of adults with relapsed or refractory classical Hodgkin Lymphoma (R/R cHL) who have relapsed after or not responded to treatment with brentuximab vedotin (BV) and are ineligible for autologous stem cell transplant (autoSCT) is narrower than the marketing authorisation for pembrolizumab in cHL (Table 2). Pembrolizumab has previously been appraised, and recommended as a treatment option, by the National Institute for Health and Care Excellence (NICE) for those with R/R cHL and no prior exposure to BV and who have either failed autoSCT or cannot undergo autoSCT but have had at least two previous therapies (TA772; described in greater detail in Section B.1.3.4).[(1)] The STA presented here is a review of TA540,[(2)] which recommended pembrolizumab enter the Cancer Drugs Fund (CDF) as an option for the management of those with R/R cHL who are autoSCT-naïve and have failed treatment with BV and remain ineligible for autoSCT.

The decision problem addressed in this submission is presented in Table 1.

Table 1. The decision problem[3]

	Final scope issued by NICE[3]	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 had BV and cannot have autologous stem cell transplant.	As per final scope Note of clarification: Pembrolizumab was not recommended for treating relapsed or refractory classical Hodgkin lymphoma in adults who have had autologous stem cell transplant and brentuximab vedotin (TA540[2]) and so the population of interest to this STA are those who transplant naïve.	N/A
Intervention	Pembrolizumab	As per final scope	N/A
Comparator(s)	• Single or combination chemotherapy including drugs such as gemcitabine, vinblastine and cisplatin; • Best supportive care.	MSD recognise that patients with R/R cHL who are not suitable for SCT after BV have few treatment options, and most people are likely to receive single-agent chemotherapy in this setting. However, no new evidence on treatment post BV was identified. For clinical effectiveness, MSD have considered the standard of care to be as set out in Cheah (2016),[4] including: • Investigational agent;	The new SLR for studies on clinical effectiveness, and updates to the economic SLRs that formed the basis of MSD's original submission to TA540,[2] identified no new study evaluating the comparators specified by NICE. Thus, MSD consider that Cheah (2016) remains the most relevant study to generate estimates of comparative clinical effectiveness versus pembrolizumab for those not responding to BV. As noted in the

		• Gemcitabine; • Bendamustine; • Other alkylatory; • BV retreatment; • Platinum based; • AutoSCT; • Other. To inform the economic model, MSD have created a blended comparator (Table 50) based on Cheah (2016), Eyre (2017) and expert opinion.[5]	original submission, Cheah (2016) reported combined outcome data for all included chemotherapy regimens, and the lack of IPD for any of the treatments precludes MSD from providing estimates of comparative effectiveness for pembrolizumab versus specific regimens. MSD understand the term Best Supportive Care to mean "no active treatment". Based on feedback from clinicians this is not a comparator of interest and is therefore not included in the submission.
Outcomes	The outcome measures to be considered include: • Overall survival; • Progression-free survival; • Response rates; • Adverse effects of treatment; • Health-related quality of life; • Time to allogeneic stem cell transplantation.	As per final scope, with the exception that time to stem cell transplant is not broken down by type of transplant (autologous versus allogeneic) as separate data were not available.	N/A
Abbreviations: autoSCT, autologous stem cell transplant; BV, brentuximab vedotin; IPD, individual patient data; N/A, not applicable; NICE, National Institute for Health and Care Excellence; STA, Single Technology Appraisal.			

B.1.2. Description of the technology being evaluated

Pembrolizumab (KEYTRUDA®) is a humanised monoclonal antibody (mAb) against the programmed death-1 (PD-1) receptor. Expression of PD-1 protein, and its ligands PD-L1 and PD-L2, triggers a signalling cascade that culminates in the suppression of T cell proliferation, cytokine release and cytotoxicity, a process that modulates the immune response (to prevent destruction of healthy cells).(6) Expression of PD-L1 and PD-L2 is frequently upregulated on the surface of tumour cells, as well as other cells in the tumour microenvironment. By binding to the PD-1 receptor, and thus blocking its interaction with PD-L1 and PD-L2, pembrolizumab reverses PD-1-mediated T-cell suppression, thereby reactivating tumour-specific cytotoxic T lymphocytes and restoring antitumour immunity.

A description of pembrolizumab and its proposed use for the treatment of R/R cHL is available in Table 2. The draft Summary of Product Characteristics (SmPC) is provided in Appendix C.

Table 2. Technology being evaluated

UK approved name and brand name	Pembrolizumab (KEYTRUDA®)
Mechanism of action	Pembrolizumab is a monoclonal antibody that binds to the PD-1 receptor, thereby potentiating an immune response to tumour cells.
Marketing authorisation/CE mark status	On 28 January 2021, the EMA granted an amendment to the marketing authorisation of KEYTRUDA®,(7) approving its use as a monotherapy in the treatment of adult and paediatric patients aged 3 years and older with relapsed or refractory classical Hodgkin lymphoma who have failed autoSCT or following at least two prior therapies when autoSCT is not a treatment option.
Indications and any restriction(s) as described in the summary of product characteristics (SmPC)	The proposed indication under appraisal is: People with relapsed or refractory classical Hodgkin lymphoma who have had brentuximab vedotin and cannot have autologous stem cell transplant. Pembrolizumab, as monotherapy or in combination with other agents, is also licensed for the management of:(8) • Melanoma; • Non-small-cell lung cancer;

	• Cervical cancer; • Urothelial carcinoma; • Head and neck squamous cell carcinoma; • Renal cell carcinoma; • Colorectal cancer; • Oesophageal cancer; • Triple-negative breast cancer; • Endometrial cancer.
Method of administration and dosage	The recommended dose of KEYTRUDA in adults is either 200 mg every 3 weeks or 400 mg every 6 weeks administered as an intravenous infusion over 30 minutes. Therapy must be initiated and supervised by specialist physicians experienced in the treatment of cancer. In KEYNOTE-087, patients received pembrolizumab 200 mg once every 3 weeks until disease progression, unacceptable toxicity, or patient withdrawal. Pembrolizumab could be administered for a maximum of 35 cycles (~24 months).
Additional tests or investigations	None
List price and average cost of a course of treatment	The list price of pembrolizumab is £2,630 per 100 mg vial and the total cost per administration is £5,260.
Patient access scheme (if applicable)	The price of pembrolizumab is subject to a CAA, with a simple discount of [redacted]; therefore, administration of 200 mg pembrolizumab will cost £[redacted]. Due to the highly confidential nature of the CAA discount, MSD requests that documentation submitted by the Evidence Assessment Group does not include the CAA price but instead refers the reader back to this table.
Abbreviations: autoSCT, autologous stem cell transplant; CAA, Commercial Access Arrangement; EMA, European Medicines Agency; PD-1, programmed cell death-1; PD-L1, programmed cell death ligand-1.	

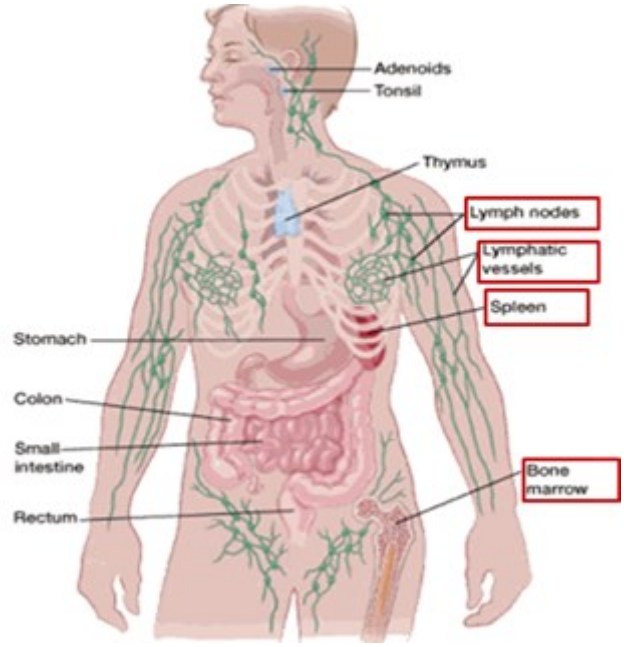
B.1.3. Health condition and position of the technology in the treatment pathway

B.1.3.1 Health condition

Lymphomas are neoplastic diseases of the lymphatic system, which is an important part of the body's immune system.[9] Neoplastic disorders cause abnormal or excessive growth of cells in a confined area of the body, which leads to the development of a neoplasm or tumour. Lymphomas are classified into two main subtypes – Hodgkin's lymphoma (HL) and non-Hodgkin's lymphoma.[9] HL is a relatively rare cancer, accounting for less than 1% of all new cancer cases in the UK.[10] HL is further subdivided into nodular lymphocyte predominant Hodgkin lymphoma (NLPHL) and classical Hodgkin lymphoma (cHL), with cHL making up 95% of cases of HL.[11]

cHL originates from uncontrolled proliferation of B lymphocytes (white blood cells) in predominantly lymph node tissues and is characterised by the presence of Reed–Sternberg cells, which are large bi- or multi-nucleated cells with a unique morphology. There are four subtypes of cHL: nodular sclerosis; mixed cellularity (mostly observed in people infected with HIV); lymphocyte-rich; and lymphocyte-depleted. The subtypes of cHL have similar prognoses and their individual management follows the same treatment pathway. In Europe and North America, the nodular sclerosis subtype accounts for 70% of all cHL. Although typically localised to a group of connected lymph nodes, cHL can spread to involve multiple lymph nodes throughout the lymphatic system (Figure 1), and, in advanced stages, may spread to extranodal sites, such as the bone marrow and spleen, and to organs outside the lymphatic system, such as the liver or lungs. Other types of extranodal involvement (e.g., digestive tract, skin, and brain) are rare in cHL.[11]

Figure 1. Lymphatic system of the human body



Source: Adapted from American Cancer Society.[12] Red boxes indicate primary sites of cHL.

The aetiology of cHL is not fully understood but is believed to involve a complex interplay of altered immune responses, environmental factors and genetic mutations.[13] Factors thought to be associated with an increased risk of developing cHL include:[9, 13]

- Lowered immunity;

  Having a medical condition that weakens the immune system; or

  Taking medicines that suppress the immune system;
- Previous exposure to Epstein–Barr virus, which causes glandular fever;
- Age;

  cHL is most often diagnosed in people in their 20s and 30s and those over age 55 (i.e. incidence is bi-modal by age);
- Family history of cHL, particularly a first-degree relative (i.e., parent, sibling or child).

Clinically, cHL most frequently presents as painless, persistent swelling of lymph nodes in the neck, armpit, or groin. Other symptoms can include persistent fatigue,

fevers and chills, night sweats, unintentional weight loss, loss of appetite, and pruritus (itching).[9] Patients are typically divided into those who are experiencing B symptoms and those who are not, where B symptoms are specified as:[11]

- Unintentional loss of more than 10% of body weight over the previous 6 months;
- Unexplained fever of at least 100.4°F (38°C);
- Drenching night sweats.

The presence of B symptoms is associated with the development of advanced forms of cHL and, consequently, worse outcomes. In some patients whose disease affects the lymph nodes in the chest, swelling of those nodes may press against the trachea and manifest as coughing or other breathing difficulties.

Diagnosis of cHL is certified by lymph node biopsy and confirmation of the presence of the hallmark Reed–Sternberg cells.[9] After diagnosis of cHL, the stage of disease is ascertained through blood tests, assessment of kidney and liver function, and imaging using computed tomography (CT) and positron emission tomography (PET), or a combination of the two. Staging of disease is established based on the Lugano classification, which is derived from the older Ann Arbor system,[14] with both tools having four stages (Table 3). Information on the staging of HL is not consistently available across the UK due to disparity in the collecting and recording of staging data.[10]

Table 3. Classification of stages of classical Hodgkin Lymphoma

Stage	Lugano classification[a]	Ann Arbor classification[a,b]
I	• Found in only one lymph node area or lymphoid organ such as the thymus (I). • Found only in one part of one organ outside the lymph system (IE).	• A single lymph-node area is involved.
II	• Found in two or more lymph node areas on the same side (above or below) of the diaphragm (II). • Extends locally from one lymph node area into a nearby organ (IIE).	• More than one lymph-node area is involved, confined to one side of the diaphragm.
III	• Found in lymph node areas above and below the diaphragm. • Found in lymph nodes above the diaphragm and in the spleen.	• Lymph nodes on both sides of the diaphragm are involved, including eventually the spleen.

IV	• Has spread widely into at least one organ outside of the lymph system, such as the liver, bone marrow, or lungs.	• One or more areas of extranodal involvement (e.g., lung, liver, bone, bone marrow).
[a] In both systems, the A and B notation is used to denote the absence or presence of B symptoms, respectively. [b] The letter E can be added to indicate one extranodal involvement contiguous to a lymph node.		

After the stage of cHL has been established, patients are further categorised by prognostic score, which is based on presence of favourable or unfavourable characteristics and provides an estimate of the likelihood of relapse (Table 4).[11, 15]

Applying the prognostic criteria generates three strata of outcome:[11]

- Favourable: 20–30% of patients with cHL, overall survival (OS) rate of up to 98% and a relapse rate of <5%;
- Intermediate: 40–50% of patients with cHL, OS rate of up to 94% and a relapse rate of 10–15%;
- Advanced: ~30% of patients with cHL, OS rate of 80–87% and a relapse rate of 15–30%.

Patients are also classified as relapsed or refractory, where refractory denotes cHL that does not respond to treatment or when the response to treatment is short lived. Stage of disease (early vs late), clinical judgement of prognosis, and relapsed or refractory status guide the management of cHL (see Section B.1.3.4), with the overall goal of treatment being achievement of sufficient clinical response to enable stem cell transplantation.

Table 4. Overview of factors considered when determining likely prognosis[11]

German Hodgkin Study Group	**Lymphoma Study Association and European Organisation for Research and Therapy in Cancer**
Favourable	**Favourable**
Stage I/II with <3 lymph nodes involved; an erythrocyte sedimentation rate of ≤50; no	Stage I/II with none of the factors listed in the unfavourable classification

extranodal disease; and non-bulky mediastinal involvement	
Intermediate	**Unfavourable**
• Stage I/IIA with >2 lymph nodes involved; an erythrocyte sedimentation rate of >50; or bulky mediastinal involvement • Stage I/IIB with an erythrocyte sedimentation rate of >50 or >2 lymph nodes involved	Stage I/II with one of these factors: • Aged older than 49 years; • >3 lymph nodes involved; • An erythrocyte sedimentation rate of >50 and A symptoms or an erythrocyte sedimentation rate of >30 and B symptoms; • Extranodal disease; • Bulky mediastinal involvement or diameter >10 cm.
Advanced	**Advanced**
• Stage III/IV • Stage IIB with extranodal involvement or bulky mediastinal involvement	Stage III/IV

B.1.3.2 Epidemiology

Surveillance data from England, Scotland, and Wales, as reported by Cancer Research UK, show that the incidence of HL follows a bimodal age distribution, with the first peak in young adults (20–24 years) and the second occurring in the elderly (75–79 years).[10] Of HL cases reported in the UK, the division between females and males is 42% and 58%, respectively, and incidence rates are significantly lower in females than males in several (mainly older) age groups.

The most up-to-date estimates of incidence of cHL are available for 2020, when 1,525 new cases of cHL were reported in England, with a breakdown of cases by age and stage of:[16]

- 0–19 years: 170 (11%);

  stage not reported

- 20–59 years: 910 (60.0%);

  stage 1, 71 (8%); stage 2, 298 (33%); stage 3, 126 (14%), stage 4, 280 (31%), stage unknown, 136 (15%);

- over 60 years: 445 (19%);

stage 1, 47 (11%); stage 2, 81 (18%); stage 3, 99 (22%), stage 4, 139 (31%), stage unknown, 79 (18%)

The incidence of cHL reported for 2020 may underestimate the number of new cases diagnosed annually in subsequent years, when taking into consideration the strain the COVID-19 pandemic, which began in 2020, placed on healthcare settings across the UK, a situation from which many services have not yet fully recovered. Cancer services continue to face significant challenges, with persistent delays to and decreases in tests, diagnosis, and treatment. Research on cancer statistics commissioned by Cancer Research UK indicates that, in 2023, the number of people awaiting a key diagnostic test in England is amongst the worst on record (since 2006).[17] Additionally, the target of urgent suspected cancer patients being seen by a specialist within 2 weeks of referral has not been met since May 2020.[17]

Incidences of HL pre-2020 may offer an estimate that better reflects the number of cHL diagnoses in future years, when taking together population growth and the anticipated return of healthcare services to levels recorded pre-COVID-19 pandemic. Annual incidence of HL in England for the period 2015–2019 suggests fluctuation across the years in the number of people diagnosed with HL, but the number of cases seems to be consistently higher than that recorded in England in 2020 (Table 5).[18] Additionally, Cancer Research UK estimates that 2,400 cases of HL will be recorded annually in the UK for the period 2023–2025, rising to around 2,900 cases annually in 2038–2040.[10]

Table 5. Annual incidence of Hodgkin lymphoma in England for 2015–2019[18]

Year	Hodgkin lymphoma	Age standardised rate [a] (LCI, UCI)	Classical Hodgkin lymphoma [b]
2015	1,803	3.4 (3.2, 3.5)	1,713
2016	1,735	3.2 (3.1, 3.4)	1,648
2017	1,819	3.3 (3.1, 3.4)	1,728
2018	1,786	3.2 (3.1, 3.4)	1,697
2019	1,841	3.3 (3.2, 3.5)	1,749

[a] Rate per 100,000 persons.

[b] Calculated by MSD based on estimate that 95% of HL cases are attributed to cHL.

Abbreviations: cHL, classical Hodgkin lymphoma; HL, Hodgkin lymphoma; LCI, lower confidence interval; UCI, upper confidence interval.

Advances in treatment have led to a substantial improvement in survival for HL, and therefore cHL, with a decline of 74% in mortality from HL from the 1970s to 2017–2019.[10] Outcomes of HL are usually favourable, with 75% of people alive at 10 years and longer after initial diagnosis.[10] Around 90% of people are alive at 1 year after diagnosis, falling to 80% at 5 years. However, analysis of OS by stage of disease shows a considerable difference in 5-year OS between early and advanced cHL, with the proportion of people alive at 5 years ranging from 96.0% to 99.4% for those diagnosed at an early stage compared with 56% to 89% for people with advanced stage disease. Additionally, those diagnosed with early stage cHL and favourable characteristics have a better prognosis than those presenting with unfavourable characteristics, with 5-year OS of 98.8%–99.4% and 96.0–96.2%, respectively. The reported survival values should be interpreted with caution because they are likely to be substantially lower in the context of the later lines of therapy being considered for the proposed position of pembrolizumab in the treatment pathway.

There is evidence that patients who are described as R/R have poor prognosis compared with their counterparts who respond to therapy.[19, 20] In patients with R/R cHL, time to initial relapse after high-dose chemotherapy and autoSCT was identified as a key prognostic factor for survival. Patients who relapsed within 12 months of treatment showed significantly lower survival compared with patients who relapse at 12 months or longer after finishing treatment (n=115: 5-year OS: 44% with relapse at <12 months vs 63% with relapse ≥12 months; P=0.03). A retrospective trial of patients with R/R cHL who had progressed after high-dose chemotherapy and autoSCT (n=71) also identified time to relapse as a key prognostic factor. Patients who relapsed within 6 months had a worse prognosis than those who relapsed after 6 months, with a median OS of 15 months and 36 months, respectively. Most people (96%) had relapsed within 2 years, and 5-year OS was about 20%, which is markedly lower than the estimate of 80% for HL reported by Cancer Research UK.[10] Patients with a history of primary refractory disease also had worse outcomes, with a

median OS of 19 months compared with 27 months for patients who achieved a complete response to their primary treatment.

B.1.3.3 Burden of disease

A retrospective UK observational study evaluated direct costs of the management of recurrent HL after autoSCT, encompassing treatment, hospital stay, outpatient visits, scans, and day care visits.[21] The treatment pathways captured were: chemotherapy followed by allogeneic SCT (alloSCT); chemotherapy followed by second autoSCT; best supportive care; and palliative chemotherapy. Chemotherapy followed by alloSCT was the most expensive treatment option, with best supportive care the least costly (Table 6). The direct costs associated with the management of R/R cHL are, therefore, also substantial, and are likely to increase with the management of disease progression and multiple lines of therapy.

Table 6. Summary of direct costs for treatments used in the management of recurrent HL[21]

Regimen	Mean cost per patient (range), £
Chemotherapy followed by alloSCT	110,374 (69,289–191,670)
Palliative chemotherapy	32,264 (2,686–119,820)
Chemotherapy followed by second autoSCT	21,612 (21,612–21,612)
Best supportive care	13,288 (8,485–23,295)
Abbreviations: alloSCT, allogeneic stem cell transplant; autoSCT, autologous stem cell transplant; HL, Hodgkin lymphoma.	

In addition to direct costs, given that cHL is one of the more common cancers observed in young adults, indirect costs affecting both patients and caregivers are also likely to be substantial, driven by the age (working age) of patients and life lost. One study estimating paid and unpaid productivity lost due to cancer-related premature mortality in 31 European countries, including the UK, reported that the average loss per premature death was highest for HL (€506,345).[22] Caregiver burden, in terms of time and resource, is also likely to be substantial.

B.1.3.4 Treatment pathway and proposed positioning of the technology

Cure rates for cHL are considered high, with some people achieving cure or long-term remission after standard first-line chemotherapy and radiotherapy. Despite the high survival rate, particularly for those with early stage disease (detailed in Section B.1.3.2), there remains an unmet need for those who are refractory to treatment or experience relapse. In the absence of NICE guidelines for the treatment of R/R cHL, the company outlines below the treatment pathway, and pembrolizumab's proposed position, based on NICE recommendations and those of the British Committee for Standards in Haematology (BCSH).[15, 23, 24]

B.1.3.4.1 Treatment pathway for relapsed/refractory cHL

Treatment strategies for cHL are typically determined by the stage and characteristics of disease, lymph node size, involvement of extranodal sites and, importantly the patient's age and general health. Updated guidance from the BCSH details first-line treatment strategies by stage of disease, that is early versus advanced, and for early stage cHL by presence of favourable or unfavourable characteristics.[24] First-line therapies are based on combination of ABVD (doxorubicin, bleomycin, vinblastine and dacarbazine) and escalated BEACOPP (bleomycin, etoposide, doxorubicin, cyclophosphamide, vincristine, procarbazine and prednisone), with or without radiotherapy (Figure 2). It is noted that older patients (>60 years) have poorer outcomes, especially with advanced-stage cHL, and they are more likely to die from non-lymphoma causes, including bleomycin lung toxicity, and treatment with ABVD and escalated BEACOPP is challenging.[24]

About 10–15% of patients with early stage, and 15–30% with advanced-stage cHL, will have refractory disease or will relapse after achieving a complete remission.[23, 25] For those patients who do not achieve long-term remission, salvage treatment may be an option, comprising chemotherapy and/or radiotherapy with the intent to enable autoSCT, which is regarded as potentially curative: autoSCT involves replacing the patient's diseased or damaged bone marrow with healthy stem cells harvested from their body. Following salvage therapy, the goal of which is to produce a major clinical response, a subset of patients will remain ineligible for autoSCT, typically due to a lack of sufficient remission to proceed (i.e., do not achieve complete [CR] or partial

remission [PR]) or the presence of characteristics, such as older age or a comorbidity, that would preclude a transplant.

Based on NICE recommendations, treatment options available to those who have relapsed or refractory cHL after autoSCT are BV (TA524)[26] or pembrolizumab (TA772),[1] if the patient has not previously received BV (Figure 2). Nivolumab is an option for those who have undergone autoSCT and treatment with BV, irrespective of whether BV is given before or after autoSCT (TA462).[27] From 2020 until mid-2023, interim COVID guidance was in place allowing use of BV in 2L instead of 3L therapy.

Those who are ineligible for autoSCT after initial treatment with chemotherapy, with or without radiotherapy, are eligible for treatment with BV (TA524)[26] or pembrolizumab (TA772),[1] if the patient has not previously received BV. The KEYNOTE-204 (N=304) open label, randomised Phase III trial evaluated the clinical efficacy of pembrolizumab against that of BV in people with R/R cHL, and comprised a mixture of those had received prior autoSCT (37%) and those who were ineligible for autoSCT (63%). In the subgroup of autoSCT-naïve patients, pembrolizumab was found to statistically significantly improve progression free survival (PFS) compared with BV, reducing the risk of progression by 39% compared with BV (hazard ratio [HR] 0.61; 95% confidence interval [CI] 0.42 to 0.89).

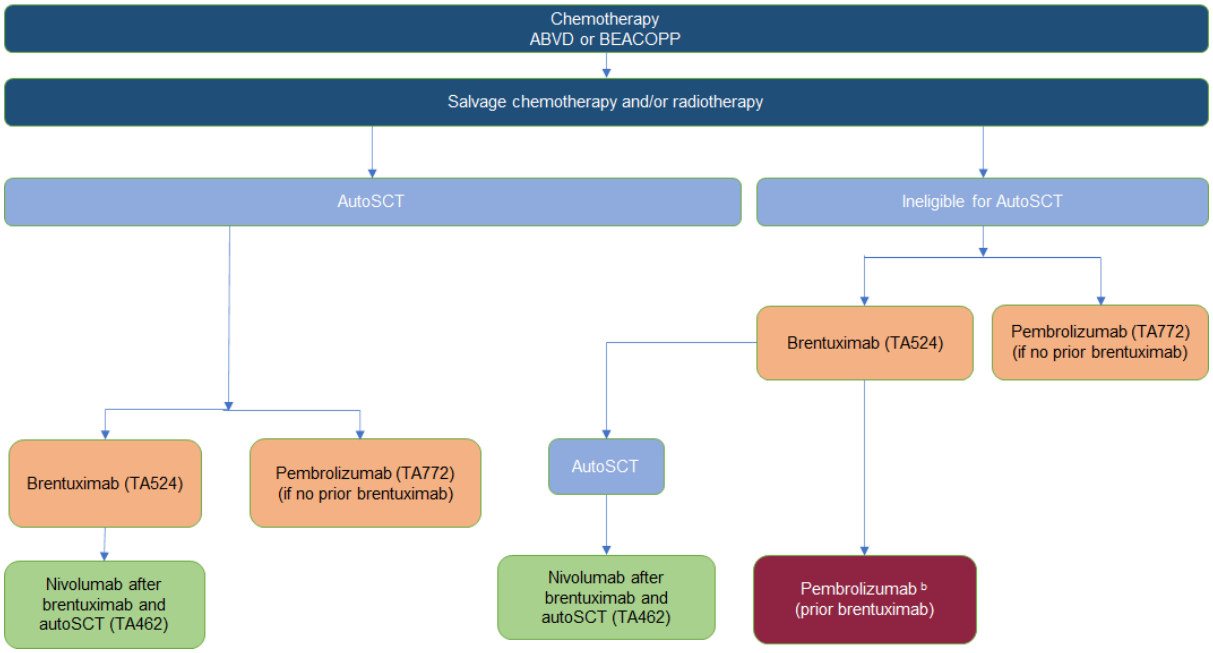
Outcomes after high-dose chemotherapy and autoSCT or alloSCT are influenced by disease status at the time of SCT.[28] More favourable outcomes are noted for patients with chemosensitive (i.e., responsive to chemotherapy) disease, predominantly those who achieve a complete metabolic response (CMR) as determined by pre-SCT PET and younger, fitter patients. By contrast, patients who have chemorefractory cHL, chiefly those who fail to respond to ≥2 lines of salvage therapy, are typically considered poor candidates for SCT. SCT is typically only considered an option for chemosensitive patients. There is some evidence that treatment with PD-1 blockade (e.g., pembrolizumab and nivolumab) may result in higher-than-expected response rates to subsequent cytotoxic therapy, a proposal echoed by clinical experts consulted by MSD, thus potentially improving the likelihood of subsequent autoSCT or alloSCT.[5] A retrospective analysis in patients

with R/R cHL who received autoSCT after PD-1 blockade found that response to anti-PD-1 therapy was a statistically significant predictor of improved PFS (18-month PFS, 51% for non-responders vs 88% for responders; P<0.001; discussed in more detail in Section B.2.6).[28]

Given that pembrolizumab has been shown to be associated with improved PFS compared with BV,[29, 30] and considering that pembrolizumab has a favourable safety profile among a heavily pre-treated patient population, it is anticipated that the use of BV at the third-line setting in autoSCT-naïve patients will decline, with pembrolizumab the preferred treatment option. However, there will be some patients for whom clinicians will consider BV to have a more favourable benefit/risk profile compared with pembrolizumab. In this late-stage setting, those who do not respond to BV have a poor prognosis, with few effective treatment options available post BV. High-dose chemotherapy is unlikely to be a viable treatment option at this stage in the treatment pathway, with many patients unlikely to be fit enough to tolerate the associated toxicity. Clinical experts consulted by MSD fed back that single-agent chemotherapy, often bendamustine due to its low cost, would the treatment of choice on failure to respond to BV.[5]

For the reasons outlined above, and the data presented on clinical effectiveness (Section B.2.6), the company propose that pembrolizumab, in line with its licensed indication (Table 2), would offer both patients and clinicians a much needed treatment option for those autoSCT-naïve patients who have failed to respond to BV and remain ineligible for autoSCT (Figure 2).

Figure 2. R/R cHL treatment pathway in the UK[a]



[a] MSD's interpretation of the treatment pathway for R/R cHL based on NICE recommendations and available guidelines.

[b] Pembrolizumab's proposed position in the treatment pathway that is under consideration in this STA.

Abbreviations: ABVD, doxorubicin, bleomycin, vinblastine and dacarbazine; autoSCT, autologous stem cell transplant; BEACOPP, bleomycin, etoposide, doxorubicin, cyclophosphamide, vincristine, procarbazine and prednisone.

B.1.4. Equality considerations

No equity or equality considerations are anticipated.

B.2. Clinical effectiveness

Summary of key clinical effectiveness information

Pivotal study:

- A systematic literature review (SLR) identified one study (KEYNOTE-087) that provided direct evidence on the efficacy and safety of pembrolizumab in the management of R/R cHL after BV.
- KEYNOTE-087 is an ongoing phase II, multicentre, multi-cohort, single arm, study of pembrolizumab in patients with R/R cHL. KEYNOTE-087 comprised three cohorts of patients, of which cohort 2 (N=81) is the population relevant to the STA, that is, those who are SCT-naïve and have relapsed after treatment with, or failed to respond to, BV.
- Clinical outcomes are reported from a database cutoff of 15 March 2021, which represents a median follow up of 62.2 months.
- 52 (64.2%) and 55 (67.9%) of 81 patients in cohort 2 achieved either CR or PR as determined by IWG and Lugano criteria, respectively.
- 29.6% of people from cohort 2 went on to receive SCT.
- Median PFS for cohort 2 was 11.1 months (95% CI: 7.5 to 13.7).
- Estimated mean OS was 53.7 months (SE 1.8 months). Median OS was not reached.

Supporting data from the CDF

- Data captured prospectively on pembrolizumab between 25 July 2018 and 30 September 2022.
- 215 unique patients with CDF applications were reported in the SACT dataset and were included in the final cohort.
- SCT was received by 30.2% of all patients and 49% of those identified as SCT candidate, with an approximate ratio of autoSCT to alloSCT of 35% to 65%.
- After a median follow-up time of 19.2 months, 73 (34.0%) of people had died. Median OS was not reached: mean OS was not available.

Indirect treatment comparisons:

- No relevant randomised controlled trial or comparative observational study was identified by the SLR.

- Estimates of comparative clinical effectiveness for OS for pembrolizumab versus SoC were derived from unanchored indirect treatment comparisons versus the excluded study deemed to be most representative of SoC.
- For cohort 2 from KEYNOTE-087, the HRs for OS versus the study most representative of SoC were 0.25 and 0.24 for a naïve-unadjusted comparison and a MAIC, respectively. Both analyses achieved statistical significance.
- HR for OS for pembrolizumab versus SoC using the same comparator study and using the SACT dataset was 0.59 (95% CI: 0.40 to 0.86).
- Bucher analyses using a common comparator indicated HRs of [redacted] to 0.41 for pembrolizumab versus SoC, with no 95% confidence interval for the analyses crossing the line of no effect

Clinical effectiveness conclusions

- Analyses carried out by MSD demonstrate that pembrolizumab improves clinical outcomes in those with R/R cHL who are ineligible for autoSCT after BV, a population that typically has a poor prognosis and has exhausted available treatment options.
- MSD consider that KEYNOTE-087 and the SACT dataset represent the most robust evidence on the clinical effectiveness of pembrolizumab relevant to the decision problem, but acknowledge the limitation that there is a paucity of data on comparators relevant to the decision problem.

B.2.1. Identification and selection of relevant studies

A systematic literature review (SLR) was carried out to identify and select relevant evidence on the efficacy and safety of treatments for patients with R/R cHL. The SLR focused on those who had failed on treatment with BV and could not have autoSCT, as per the indication recommended for entry into the CDF in TA540.[2] Given the narrow target population, to maximize the available evidence base, both clinical trials (randomised and non-randomised) and observational studies were eligible for inclusion. As the manufacturer of pembrolizumab, MSD are aware of all relevant clinical trials for pembrolizumab in the relevant indication. Because BV was approved by the EMA in 2012,[31] a time restriction of 2010 until the date of the search was applied. Full details on the SLR methodology and results are provided in Appendix D.1.

B.2.2. List of relevant clinical effectiveness evidence

The SLR retrieved 1,959 unique records, from which 2 unique trials (10 publications) were identified and considered relevant to the decision problem. Both trials are single arm studies evaluating the clinical efficacy and safety of pembrolizumab in R/R cHL – KEYNOTE-013 and KEYNOTE-087. Given the small number of people in the relevant population in KEYNOTE-013 (N=9), the study is not discussed further in Document B, but details on trial design and conduct are provided in Appendix D.1. Thus, evidence on the clinical efficacy and safety of pembrolizumab is derived from the relevant subgroup enrolled in KEYNOTE-087 (N=81; Table 7). Data collection for KEYNOTE-087 is ongoing. Here, 5-year data captured on clinical outcomes and adverse effects are presented in support of MSD's application (data cutoff 15 March 2021; Section B.2.6).[32, 33]

Additionally, to address areas of uncertainty identified by the Appraisal Committee during TA540, data were collected prospectively for pembrolizumab on specific outcomes during its time in the CDF (Table 8).[34] Data from the systemic anti-cancer therapy (SACT) cohort for available outcomes are presented alongside outcomes from KEYNOTE-087. The final cohort for the SACT dataset comprised 215/220 (98%) unique patients with CDF applications.

Estimate of OS for pembrolizumab derived from the SACT dataset is implemented in the base-case analysis of the economic model, with OS from KEYNOTE-087 used in a scenario analysis (Section B.3). The SACT data were chosen to inform the base-case economic analysis as, compared with KEYNOTE-087, the cohort forming the dataset is more generalisable to the UK population likely to be treated with pembrolizumab in clinical practice in England. However, as a well-conducted clinical study, KEYNOTE-087 remains relevant to the decision problem and provides details on other outcomes of interest encompassing the clinical efficacy and safety profile of pembrolizumab.

Table 7. Clinical effectiveness evidence: KEYNOTE-087

Study	KEYNOTE-087[32, 33]
Study design	Multi-centre, multi-cohort, single-arm, non-randomised study
Population	KEYNOTE-087 enrolled three cohorts, of which cohort 2 is relevant to the decision problem that is the focus of this STA. **Cohort 1** Participants with R/R cHL who failed to achieve a response or progressed after autoSCT and relapsed after treatment with, or failed to respond to treatment with, BV post-autoSCT. **Cohort 2** Participants with R/R cHL who were unable to achieve a CR or PR to salvage chemotherapy and who did not receive autoSCT, but relapsed after treatment with, or failed to respond to treatment with BV. **Cohort 3** Participants with R/R cHL who failed to achieve a response to, or progressed after, autoSCT, and had not received BV after autoSCT and did or did not receive BV as part of primary treatment or salvage treatment.
Intervention(s)	Pembrolizumab 200 mg administered via intravenous infusion (infused over 30 minutes) on day 1 of each 3-week cycle for up to 35 cycles.
Comparator(s)	N/A
Indicate if study supports application for marketing authorisation	Yes
Indicate if study used in the economic model	Yes; scenario analysis
Rationale if study not used in model	N/A
Reported outcomes specified in the decision problem	• Overall survival • Progression-free survival • Response rates • Health-related quality of life • Time to alloSCT • Adverse effects of treatment
All other reported outcomes	• Time to autoSCT

Abbreviations: autoSCT, autologous SCT; alloSCT, allogeneic SCT; BV, brentuximab vedotin; CR, complete remission; N/A, not applicable; PR, partial remission; R/R cHL, relapsed or refractory classical Hodgkin lymphoma; STA, Single Technology Appraisal.

Table 8. Clinical effectiveness evidence: SACT dataset

Study	Systemic Anti-Cancer Therapy dataset[34]
Study design	Real-world evidence Data captured prospectively between 25 Jul 2018 and 30 September 2022
Population	Criteria for access to pembrolizumab through the CDF included:[35] • Adult with histologically documented classical Hodgkin lymphoma; • Failed two lines of chemotherapy and also BV; • Has not received SCT of any kind and is ineligible for SCT; • Patient is **either** a candidate for future SCT if there is sufficient benefit from pembrolizumab, **or** not a candidate for SCT however good the response to pembrolizumab may be; • ECOG performance status of 0 or 1; • No previous treatment with PDL-1, PDL-2, CD137 or CTLA-4 inhibitors.
Intervention(s)	Pembrolizumab 200 mg administered via intravenous infusion (infused over 30 minutes) on day 1 of each 3-week cycle for up to 35 cycles.
Comparator(s)	N/A
Indicate if study supports application for marketing authorisation	No
Indicate if study used in the economic model	Yes
Rationale if study not used in model	N/A
Reported outcomes specified in the decision problem	• Time to SCT • Proportion of patients who receive a SCT; • Overall survival
All other reported outcomes	• Treatment duration
Abbreviations: autoSCT, autologous SCT; alloSCT, allogeneic SCT; BV, brentuximab vedotin; CDF, Cancer Drugs Fund; CTLA-4, cytotoxic T-lymphocyte antigen 4; ECOG,	

Eastern Cooperative Oncology Group; N/A, not applicable; PDL, programmed death-ligand; SCT, stem cell transplant.

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

B.2.3.1 Summary of the methodology of the KEYNOTE-087 study

B.2.3.1.1 Trial design

KEYNOTE-087 (NCT02453594[36]) is a phase II, multicentre, multi-cohort, single arm, trial of pembrolizumab in patients with R/R cHL. The rationale for selecting a single-arm, non-comparative trial was largely based on the absence of established clinical practice at this later-line setting and the limited number of eligible patients for treatment.

As detailed earlier (Table 7), KEYNOTE-087 comprised three cohorts of patients, of which cohort 2 is the population relevant to the STA:

- 1: 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 CR or PR to salvage chemotherapy and did not receive autoSCT, but have relapsed after treatment with, or failed to respond to, BV;
- 3: 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 or salvage treatment.

Patients who experienced a CR or PR or had stable disease (SD) could continue treatment with pembrolizumab 200 mg every 3 weeks (Q3W) for up to 2 years (approximately 37 administrations via intravenous infusion), or until unacceptable toxicity or documented disease progression. After documented disease progression, or the start of new antineoplastic therapy, each patient 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 investigators' discretion, patients who attained a CR could consider stopping pembrolizumab after a minimum of 24 weeks of treatment. Additionally, at least two

doses of pembrolizumab had to be received after documentation and confirmation of CR before cessation of treatment was allowed. Patients who later experienced disease progression were eligible for retreatment with pembrolizumab at the discretion of the investigator if:

- no cancer treatment had been administered since the last dose of pembrolizumab;
- the subject met the safety parameters listed in the inclusion/exclusion criteria of the protocol; and
- the trial was open.

Patients would resume therapy at the same dose and schedule as at the time of initial discontinuation.

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

Tumour imaging could be performed using CT and/or PET. After screening, CT scans were to be repeated every 12 weeks, with PET repeated at week 12 and week 24 to confirm CR or progressed disease (PD), and as clinically indicated. Disease assessments and imaging were to continue until documented disease progression, the start of new anti-cancer treatment, withdrawal of consent, death, or the end of the study, whichever occurred first.

As of the date of data cutoff (15 March 2021), enrolment was closed, and all enrolled participants had either completed or discontinued original protocol treatment. The data presented here correspond to approximately 5 years of follow-up and include data from participants who were retreated with pembrolizumab after experiencing CR and relapsing (Section B.2.6).[32]

B.2.3.1.2 Eligibility criteria

Key inclusion and exclusion criteria for eligibility for KEYNOTE-087 are listed below. Full criteria are available in the Clinical Study Report (CSR) for KEYNOTE-087.[33]

To be eligible for entry into KEYNOTE-087, patients must:

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

  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 not have received BV post autoSCT. These patients may or may not have received BV as part of primary treatment, or salvage treatment.

  Relapsed was defined as disease progression after most recent therapy, and refractory to treatment as failure to achieve CR or PR to most recent therapy.

- Have measurable disease, which was defined as at least one lesion that can be accurately measured in at least two dimensions with spiral 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 ECOG Performance scale.

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

- Was participating and receiving study therapy or has participated in a study of an investigational agent and received study therapy or used an investigation device within 4weeks of the first dose of treatment;
- Had a diagnosis of immunosuppression or was 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;

- Had received a prior monoclonal antibody within 4 weeks prior to study Day 1 or had not recovered (i.e., ≤ Grade 1 or at baseline) from adverse events due to agents administered more than 4 weeks earlier;
- Had received prior chemotherapy, targeted small molecule therapy, or radiation therapy within 2 weeks prior to study Day 1 or had not recovered (i.e., ≤ Grade 1 or at baseline) from adverse events due to a previously administered agent;
- Had undergone 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 versus host disease;
- Had 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;
- Had evidence of active, non-infectious pneumonitis;
- Had an active infection requiring intravenous systemic therapy;
- Was 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;
- Had received prior therapy with an anti-PD-1, anti-PD-L1, anti-PD-L2, anti-CD137, or anti-CTLA-4 antibody (including ipilimumab or any other antibody or drug specifically targeting T-cell co-stimulation or checkpoint pathways).

B.2.3.1.3 Settings and locations where the data were collected

KEYNOTE-087 is a global study that enrolled 210 patients between 26 June 2015 and March 21 2016 across 51 study sites.[33] Of the 51 study sites, three were located in the UK, with the remaining sites located as follows: 23 across Europe (France, Russia, Italy, Spain, Germany, Greece, Hungary, Sweden, and Norway); 11 in the USA; seven in Japan; four in Israel; two in Australia, and one in Canada.

Of the 210 patients enrolled into KEYNOTE-087, 81 were recruited to cohort 2, with 69 and 60 being eligible for cohorts 1 and 3, respectively. Fourteen patients were enrolled from the three UK study sites (cohort 1, n=4; cohort 2, n=10).

B.2.3.1.4 Trial drugs and concomitant medications

Administration of pembrolizumab

All patients received pembrolizumab 200 mg via intravenous (IV) infusion over 30 minutes every 3 weeks (Q3W), with treatment administered in the outpatient setting by qualified site personnel.[33] Initiation of pembrolizumab was to take place as close as possible to the date on which the participant entered into KEYNOTE-087. For administrative reasons, pembrolizumab could be administered up to 3 days before or after the scheduled Day 1 of each cycle. Interruptions to the treatment plan for longer 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 the management of the patient.

Modification of pembrolizumab dose was permitted on occurrence of adverse events (AEs), both non-serious and serious, that were thought to be treatment-related. As an immune-checkpoint inhibitor, pembrolizumab, like other members of this class of drug, is associated with low grade and manageable immune-mediated AEs, such as dermatologic, gastrointestinal, and endocrine toxicities. Immune-mediated AEs can occur shortly after the first dose or several months after the last dose of treatment.

Concomitant medication

Concomitant treatments were allowed at the discretion of the investigator and in keeping with the community standards of medical care. Use of concomitant medications, including prescription, over-the-counter, herbal supplements, and IV medications and fluids, was recorded on the case report form. Changes to medication during the trial period could also be documented on the case report form, including dosage, frequency, route, and date of change. Patients taking anticoagulation therapy were allowed to continue treatment as long as the prothrombin time or activated partial thromboplastin time was within the therapeutic range of the intended use of anticoagulants.

All concomitant medications received up to 28 days before the first dose of trial treatment and 30 days after the last dose of trial treatment were 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 resulted in discontinuation of study therapy.
- Live vaccines within 30 days prior to the first dose of trial treatment and while participating in the trial;

  Live vaccines include, but are not limited to, measles, mumps, rubella, and chicken pox. Seasonal influenza vaccines for injection are generally killed virus vaccines and are permitted, but, intranasal influenza vaccines are live attenuated vaccines, and are prohibited.

- Glucocorticoids for any purpose other than to modulate symptoms from an event of clinical interest of suspected immunologic aetiology.

B.2.3.1.5 Outcomes assessed

Primary outcomes

The primary efficacy endpoint of KEYNOTE-087 is overall response rate (ORR), which is defined as the proportion of patients who achieved CR or PR at any time during the study. Classification of CR or PR followed criteria set out by the International Working Group (IWG) that was established to standardise response criteria for malignant lymphomas.[37] The primary analysis of ORR was based on classification of response by blinded, independent central review (BICR). ORR was reported for the "all subjects as treated" (ASaT) population and for each cohort. A co-primary objective of KEYNOTE-087 was to determine the safety and tolerability of pembrolizumab. Efficacy outcomes, in addition to ORR, are reported, as are data on adverse effects in Section B.2.6.

Other outcomes

Secondary outcomes captured were:

- ORR by investigator assessment according to the IWG response criteria;[37]

- ORR by BICR using the 5-point scale according to the Lugano Classification;[38]
- Complete remission rate (CRR) by BICR and investigator assessment according to the IWG response criteria;[37]
- CRR by BICR using the 5-point scale according to the Lugano Classification;[38]
- Progression Free Survival (PFS) by BICR and investigator assessment according to the IWG response criteria;[37]
- Duration of response (DOR) by BICR and investigator assessment according to the IWG response criteria;[37]
- OS.

For those patients who achieved CR or PR, DOR was defined 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 occurred first.

B.2.3.1.6 Baseline characteristics of participants in KEYNOTE-087 and the cohort forming the SACT dataset

KEYNOTE-087

Most people in cohort 2 of KEYNOTE-087 were aged less than 65 years at baseline (81.5%, Table 9), with a mean age of 42.3 years (standard deviation [SD] 17.4) for the group. Of the 81 patients enrolled, 44 (54.3%) had an ECOG score of 0 (54.3%), with the remaining patients having an ECOG score of 1 (45.7%). At study entry, 26 people (32.1%) reported experiencing B symptoms. As expected, based on inclusion criteria, the patients forming cohort 2 were heavily pre-treated, with a median of 4 lines (range 1 to 11) of prior treatment, and most people had received at least 3 lines of previous therapy (96.3%).

Table 9. Baseline characteristics of the population of Cohort 2 enrolled in KEYNOTE-087[33]

Characteristic	Cohort 2 (N=81), n, %
Gender	
Male	43 (53.1)
Female	38 (46.9)

Age, years	
Mean	42.3
SD	17.4
Median	40
Range	20 to 76
<65	66 (81.5)
≥65	15 (18.5)
20–24	16 (19.8)
25–29	12 (14.8)
30–34	7 (8.6)
35–39	4 (4.9)
40–44	7 (8.6)
45–49	5 (6.2)
50–54	8 (9.9)
55–59	5 (6.2)
60–64	2 (2.5)
65–69	8 (9.9)
70–74	6 (7.4)
75–79	1 (1.2)
Race	
American Indian or Alaska native	1 (1.2)
Asian	4 (4.9)
Black or African American	2 (2.5)
Missing	1 (1.2)
Multi-racial	0
White	73 (90.1)
Ethnicity	
Hispanic or Latino	5 (6.2)
Not Hispanic or Latino	65 (80.2)
Not reported	7 (8.6)
Unknown	4 (4.9)

Race group	
White	73 (90.1)
Non-White	7 (8.6)
Missing	1 (1.2)
US region	
US	20 (24.7)
Ex-US	61 (75.3)
Disease subtype	
Classical Hodgkin Lymphoma – Nodular sclerosis	65 (80.2)
Classical Hodgkin Lymphoma – Mixed cellularity	10 (12.3)
Classical Hodgkin Lymphoma – Lymphocyte rich	1 (1.2)
Classical Hodgkin Lymphoma – Lymphocyte depleted	4 (4.9)
Missing	1 (1.2)
ECOG performance status	
0	44 (54.3)
1	37 (45.7)
Prior lines of therapy group	
>=3	78 (96.3)
<3	3 (3.7)
Prior lines of therapy	
Subjects with data	81
Mean	4.0
SD	1.7
Median	4.0
Range	1 to 11.0
Prior radiation	
Yes	21 (25.9)
No	60 (74.1)

Bulky lymphadenopathy	
Yes	12 (14.8)
No	69 (85.2)
Baseline B symptoms	
Yes	26 (32.1)
No	55 (67.9)
Baseline bone marrow involvement	
Yes	5 (6.2)
No	75 (92.6)
Missing	1 (1.2)
Abbreviations: ECOG, Eastern Cooperative Oncology Group; SD, standard deviation.	

SACT dataset

In comparison with cohort 2 from KEYNOTE-087, the group of patients forming the SACT dataset are older and less fit (Table 10).[34] Where only 15 (18.5%) patients from cohort 2 were 65 years or older, 49 (23%) from the SACT dataset were aged 70 years and above at baseline. Considering performance status, where a lower ECOG score indicates better functionality, 59 (27%) of patients from the SACT dataset had a baseline ECOG score of 0 compared with 44 (54.3%) for cohort 2. It should be noted that information on baseline ECOG score was missing for 52 (24%) people from the SACT dataset.

Table 10. Baseline characteristics of the SACT cohort[34]

Characteristic	SACT cohort (N=215), n,%
Gender	
Male	130 (60)
Female	85 (40)
Age	
<40	75 (35)
40 to 49	22 (10)
50 to 59	37 (17)
60 to 69	32 (15)

70 to 79	43 (20)
80+	6 (3)
ECOG status at the start of regimen	
0	59 (27)
1	86 (40)
2	16 (7)
3	2 (1)
4	0 (0)
Missing	52 (24)
Abbreviations: ECOG, Eastern Cooperative Oncology Group; SACT, Systemic Anti-Cancer Therapy**.**	

The differences noted in age and performance score between the cohorts from KEYNOTE-087 and the SACT dataset are to be expected to some degree as it is recognized that, because of stringent inclusion criteria, participants in clinical studies are frequently younger, fitter and with fewer comorbidities than patients seen in clinical practice. However, clinical experts consulted by MSD commented that the cohort forming the SACT dataset might not be wholly generalisable to patients with R/R cHL who would now be considered a candidate for pembrolizumab in the fourth-line setting, being older and less fit than a typical patient at this time.[5] Clinical experts highlighted that much of the data informing the SACT dataset were collected during the COVID-19 pandemic, at which time there were greater concerns about immune suppression, resulting in adaptations to the typical treatment pathway. Additionally, there may be challenges to accurately recording ECOG performance status when completing the form required for access to pembrolizumab through the CDF.[5] MSD consider the cohort forming the SACT dataset to be the most appropriate population to inform the evaluation of the cost effectiveness of pembrolizumab, with the caveat that there is potentially bias against pembrolizumab in the analyses, the extent of which cannot be accurately quantified.

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

B.2.4.1 Statistical analysis and definition of study groups in KEYNOTE-087

B.2.4.1.1 Objectives, hypotheses, and endpoints

Study endpoints were evaluated within each cohort. The statistical methods used to evaluate primary and secondary efficacy endpoints for KEYNOTE-087 are summarised in Table 11.

The primary hypothesis tested was whether, after treatment with pembrolizumab, the ORR for each cohort was greater than a fixed control rate, with ORR assessed by independent central review and with classification for CR, PR, SD and progressive disease (PD) as per IWG criteria. As per the earlier definition of ORR, final analysis was conducted for each cohort when the last patient in the group reached the week 12 response assessment or has discontinued study therapy. The analysis of ORR consisted of the point estimate and 95% 2-sided exact CI using the Clopper–Pearson method, which had least 95% coverage of the true rate (Table 11). As per the protocol, investigators could continue to treat patients who were classified as having PD by central review or by site assessment, and so exploratory analyses (point estimate and 95% 2-sided exact confidence interval) were conducted for ORR to re-classify those patients who achieved PR or CR after progression as responders.

Secondary clinical endpoints (i.e., CRR, PFS, DOR, and OS) were evaluated within each cohort but did not involve hypothesis testing. Considering PFS, disease progression was assessed periodically and progression could occur any time in the time interval between the last assessment documenting absence of PD and the subsequent assessment at which PD was confirmed. For the primary analysis, the date of progression was approximated by the date of the first assessment at which PD was objectively documented as per IWG criteria, regardless of discontinuation of pembrolizumab. Death was always considered as a confirmed PD event. A secondary analysis was performed for PFS based on investigator's assessment.

DOR data were censored on the date of the last disease assessment documenting absence of PD for patients who were:

- still receiving pembrolizumab at the time of an analysis;
- given antitumour treatment (including SCT) other than pembrolizumab; or
- removed from the study prior to documentation of tumour progression.

Additionally, SCT post-initiation of pembrolizumab was considered an indicator of positive efficacy rather than failure of the current treatment.

Table 11. Analysis strategy for primary and key secondary efficacy endpoints for KEYNOTE-087

Endpoint/Variable	Statistical method	Analysis population	Missing data approach
Primary outcome			
Overall response rate • IWG criteria (2007)[37] • Independent central review	Exact test of binomial parameter; 2-sided 95% exact CI	ASaT/FAS	People with missing data were considered non-responders
Secondary outcomes			
Overall response rate • IWG criteria (2007)[37] ○ Study site • Lugano criteria (2014)[38] ○ Independent central review	Point estimate; 2-sided 95% exact CI	ASaT/FAS	People with missing data were considered non-responders
Complete remission rate • IWG criteria (2007)[37] ○ Independent central review ○ Study site • Lugano criteria (2014)[38] ○ Independent central review	Point estimate; 2-sided 95% exact CI	ASaT/FAS	People with missing data were considered non-responders
Progression-free survival • IWG criteria (2007)[37]	Summary statistics using Kaplan–Meier method	ASaT/FAS	Censored at last assessment

○ Independent central review ○ Study site			
Duration of response • IWG criteria (2007)[37] ○ Independent central review ○ Study site	Summary statistics using Kaplan–Meier method	All responders	Non-responders were excluded from the analysis
Overall survival	Summary statistics using Kaplan–Meier method	ASaT/FAS	Censored at last assessment
Abbreviations: ASaT, all subjects as treated; CI, confidence interval; FAS, full analysis set; IWG, International Working Group.			

B.2.4.1.2 Analysis populations

Efficacy analysis population

The analysis of primary efficacy endpoint for all cohorts was based on the all subjects as treated (ASaT) population, that is, all enrolled patients who received at least one dose of pembrolizumab. Supportive analyses were conducted using the full analysis set (FAS) population, which comprised all patients who: received at least one dose of pembrolizumab; had a baseline disease assessment; and had a post baseline disease assessment OR discontinued the trial due to progressive disease/drug related AE.

Safety analysis population

The ASaT population informed the analysis of safety data. At least one laboratory or vital sign measurement obtained subsequent to at least one dose of pembrolizumab was required for inclusion in the analysis of each specific parameter. To assess change from baseline, a baseline measurement was also required.

B.2.4.1.3 Statistical methods

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). To ensure inclusion of 180 patients, the protocol outlined that 190 patients

would need to be enrolled, assuming that approximately 5% of those enrolled would not receive one dose of pembrolizumab. With 60 patients in the ASaT, for cohort 2 specifically, there would be at least 93% power (one-sided 2.5% alpha level) to demonstrate that pembrolizumab is superior to a fixed control rate of 5%, assuming the underlying ORR for pembrolizumab is at least 20%. If the true ORR is 25% or greater for pembrolizumab, the analysis has >99% power. The choice of 5% for the fixed control rate for cohort 2 was based on a reported response rate of 30% for patients with R/R HL who were treated with BV, and who were either SCT-ineligible or had refused SCT (N=20).[39]

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.

B.2.5. Critical appraisal of the relevant clinical effectiveness evidence

The critical appraisal of KEYNOTE-087 was based on the Newcastle–Ottawa Scale, which was designed to assess the quality of nonrandomised studies that include a comparative cohort.[40] In brief, the tool evaluates three key domains: selection of study groups; comparability of groups; and ascertainment of exposure or outcome of interest. Studies are awarded a star or stars, depending on the domain, if they are deemed to have little to no risk of bias for that parameter, with a maximum possible score of 9. KEYNOTE-087 was deemed to have a score of 6, which is the maximum score possible for a single-arm study (Table 12).

Full details of the domains considered and supportive evidence required when using the Newcastle-Ottawa scale are available in Appendix D1.3.

Table 12. Quality assessment of KEYNOTE-087

Newcastle–Ottawa Scale domain	**KEYNOTE-087 score**
Selection	
Representativeness of the exposed cohort	*

Selection of the non-exposed cohort	Not applicable
Ascertainment of exposure	*
Outcome of interest not present at start of study	*
Comparability	Not applicable
Outcome	
Outcome assessment	*
Duration of follow-up	*
Loss to follow-up	*

B.2.6. Clinical effectiveness results of the relevant studies

Cohort 2 of KEYNOTE-087 comprised 81 patients. The first patient in Cohort 2 enrolled on 24 June 2015, and the last entered the study on 16 December 2015. Information relating to participant enrolment and baseline characteristics, as reported in Table 9, are reported from the CSR drafted in June 2016. Data on clinical efficacy and safety for pembrolizumab, as relevant to this submission, are based on a database cutoff of 15 March 2021.[33] Median duration of follow-up for cohort 2, as of 15 March 2021, was 62.2 months (range: 2.1 months to 67.5 months).

At the time of the latest data cutoff, 15 March 2021, 13 (16.0%) of the 81 patients from cohort 2 had completed treatment (Table 13). Of the 68 (84.0%) patients who discontinued pembrolizumab, 37 (45.7%) stopped treatment due to progressive disease, 9 (11.1%) due to complete response, and 5 (6.2%) due to adverse event.

Table 13. Disposition of patients forming cohort 2 of KEYNOTE-087

	Cohort 2 (N=81) **n (%)**
Status for study medication in trial segment treatment	
Started	81
Completed	13 (16.0)
Discontinued	68 (84.0)
• Adverse event	5 (6.2)
• Bone marrow transplant	2 (2.5)
• Clinical progression	1 (1.2)

• Complete response	9 (11.1)
• Lost to follow up	2 (2.5)
• Physicians Decision	6 (7.4)
• Pregnancy	1 (1.2)
• Progressive disease	37 (45.7)
• Withdrawal by subject	5 (6.2)

B.2.6.1 KEYNOTE-087: clinical outcomes

B.2.6.1.1 Overall response

The primary clinical outcome evaluated in KEYNOTE-087 was ORR as determined by BICR and based on IWG response criteria, with a secondary endpoint of ORR assessed following the Lugano classification system.

In the ASaT population, 52 (64.2%) and 55 (67.9%) of 81 patients achieved either CR or PR as per IWG and Lugano criteria, respectively (Table 14).

Median time to response by BICR and IWG criteria for those achieving CR or PR from cohort 2 (n=52) was 2.8 months (range: 2.2 to 11.0 months; Table 15), and the median DOR was 11.1 months (range: 0.0+ to 59.0+ months). Response durations of ≥12 and ≥24 months were observed in 15 (45.6% by KM estimation; Figure 3) and 10 (32.6% by KM estimation) participants, respectively (Table 15).

Table 14. Summary of best overall response based on central review as per IWG and Lugano classifications[33]

Level of response	**Cohort 2 (N=81)**			
	n (%)	**95% CI[a]**	**n (%)**	**95% CI[a]**
	IWG criteria		**Lugano criteria**	
OR (CR + PR)	52 (64.2)	52.8 to 74.6	55 (67.9)	56.6 to 77.8
CR	21 (25.9)	16.8 to 36.9	23 (28.4)	18.9 to 39.5
PR	31 (38.3)	27.7 to 49.7	32 (39.5)	28.8 to 51.0
SD	8 (9.9)	4.4 to 18.5	6 (7.4)	2.8 to 15.4
PD	19 (23.5)	14.8 to 34.2	18 (22.2)	13.7 to 32.8
NA	2 (2.5)	0.3 to 8.5	2 (2.5)	0.3 to 8.5
[a] Based on binomial exact confidence interval method.				

Abbreviations: CI, confidence interval; CR, complete remission; IWG, International Working Group; NA, no assessment; OR, objective response; PD, progressive disease; PR, partial remission; SD, stable disease.

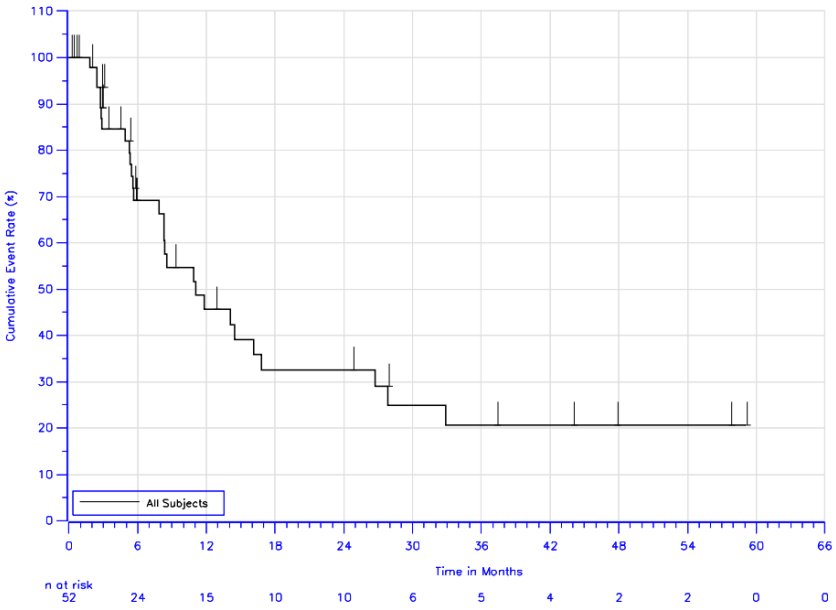
Table 15. Summary of time to response and response duration for those achieving CR or PR in cohort 2 and based on BICR and IWG criteria[33]

Outcome	Cohort 2 (n=52)
Time to response (months)	
• Mean (SD)	3.2 (1.4)
• Median (range)	2.8 (2.2 to 11.0)
Response duration (months)	
• Median (range)	11.1 (0.0+ to 59.0+)
• 95% CI	7.9 to 16.8
• Number with response lasting ≥3 months (%)[a]	36 (84.5)
• Number with response lasting ≥6 months (%)[a]	24 (69.1)
• Number with response lasting ≥9 months (%)[a]	19 (54.7)
• Number with response lasting ≥12 months (%)[a]	15 (45.6)
• Number with response lasting ≥24 months (%)[a]	10 (32.6)
• Number with response lasting ≥36 months (%)[a]	5 (20.7)
• Number with response lasting ≥48 months (%)[a]	2 (20.7)

[a] From product-limit (Kaplan–Meier) method for censored data.

Abbreviations: BICR, blinded independent central review; CI, confidence interval; CR, complete remission; IWG, international working group; PR, partial remission; SD, standard deviation.

Figure 3. Kaplan–Meier estimates of duration of objective response for cohort 2 based on central review and IWG criteria[33]



B.2.6.1.2 HRQoL

In the ASaT population of KEYNOTE-087 (N=210), pembrolizumab was associated with a clinically meaningful improvement in EQ-5D from baseline to week 12, with a mean change in score of 8.4 points, where >8 point increase represents the threshold for a minimal clinically important difference (Table 16). The greatest improvements in HRQoL were recorded for those patients classed as responders, that is, those achieving CR or PR. Responders demonstrated a clinically meaningful improvement in QoL, as assessed by both tools, at week 12 compared with patients with SD or PD (Table 16). Considering the disease-specific questionnaire EORTC QLQ-C30, responders had a mean change in score of 10.4 points from baseline to week 12, where a ≥10 point increase represents the threshold for a minimal clinically important difference.

Table 16. Summary of HRQoL endpoints in KEYNOTE-087

Population	**Change from baseline at week 12 EORTC QLQ-C30**		**Change from baseline at week 12 EQ-5D-VAS**	
	N[a]	Mean (Sdev)	N[a]	Mean (Sdev)
ASaT	184	8.6 (1.6)	191	8.4 (1.4)

CR/PR	110	10.4 (2.1)	116	10.9 (1.8)
SD	48	7.3 (3.2)	49	5.4 (3.0)
PD	26	3.5 (3.6)	26	2.6 (2.7)

[a] Number of patients in the ASaT population with observations at baseline and week 12. Data cutoff of 25 September 2016.

Abbreviations: ASaT, all subjects as treated; CR, complete remission; EORTC QLQ-C30, European Organization for Research and Treatment of Cancer quality of life questionnaire-core 30; EQ-5D VAS, EuroQOl 5 dimensions visual analogue scale; PD, progressive disease; PR, partial remission; SD: stable disease; Sdev, standard deviation.

B.2.6.1.3 SCT

Data on the proportion of people receiving a SCT and time to SCT are available from both KEYNOTE-087 and the SACT dataset.

Proportion of people receiving SCT

The proportion of people undergoing a SCT was similar for cohort 2 and the SACT dataset, with 29.6% and 30.2% of people receiving SCT, respectively.

Considering cohort 2 of KEYNOTE-087, of the 24 (29.6%) patients who went on to receive SCT, 14 (58.3%) and 9 (37.5%) underwent autoSCT and alloSCT, respectively.(41) One person received both auto- and alloSCT. Baseline characteristics for the subgroup of patients from cohort 2 receiving SCT are presented in Appendix E.

Of the 215 people forming the SACT dataset, 132 (61%) patients were identified in Blueteq as being suitable candidates for SCT (Table 17).(34) In contrast to cohort 2 of KEYNOTE-087, of the 65 (30.2%) patients from the SACT dataset having undergone SCT, the majority received alloSCT (23 [35.4%] receiving autoSCT vs 42 [64.6%] receiving alloSCT; Table 17). The proportion of autoSCT versus alloSCT recorded for the SACT dataset has been incorporated into MSD's economic model.

Table 17. Applications to CDF for pembrolizumab for treating cHL – SCT suitability in Blueteq and SCT procedures in HES(34)

SCT suitability	Blueteq SCT suitability[a] (N)	HES AlloSCT (N)	HES AutoSCT (N)	HES SCT (N)	HES SCT (%)

Candidate for future SCT	132	42	23	65	49
Not a candidate for SCT	83	3	1	–	–
Total	215	45	24	65[b]	–

[a] Applications made between 25 July 2018 and 30 September 2022.

[b] Total number of SCTs including those who were not initially deemed a candidate for SCT seems to be 69 rather than 65. MSD have taken the 65/215 at face value and discuss an alternate assumption of 69/215 in the economic analysis.

Abbreviations: alloSCT, allogeneic SCT; autoSCT, autologous SCT; cHL, classical Hodgkin lymphoma; HES, Hospital Episodes Statistics.

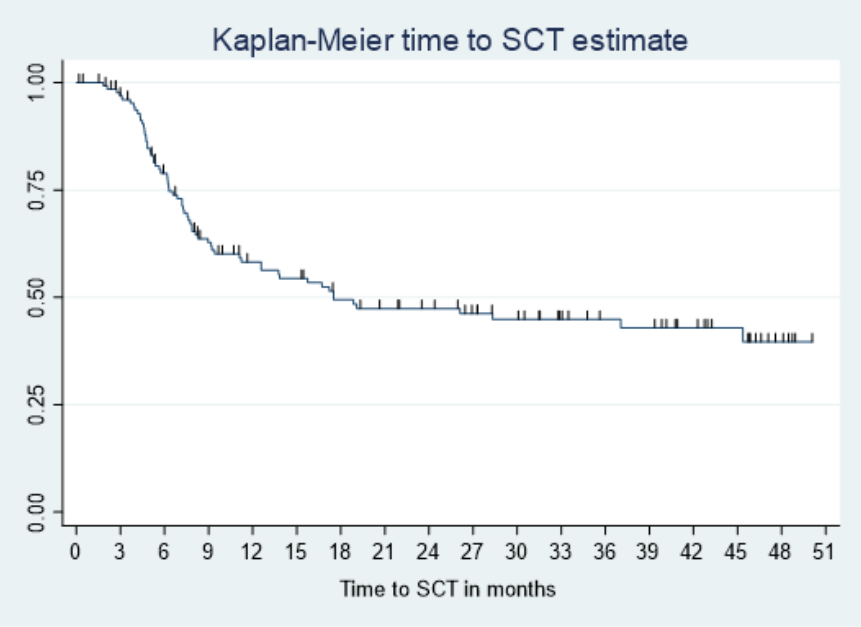
Clinical experts consulted by MSD indicated that they considered the ratio of autoSCT to alloSCT in the SACT, which is approximately 35% autoSCT to 65% alloSCT to be higher for alloSCT than would be seen in clinical practice in England.(5) One clinical expert suggested that alloSCT is typically now only performed on relapse post-autoSCT. Patients with cHL who are SCT-naïve, including those who have relapsed after achieving remission from initial treatment, would be considered for autoSCT if they are fit enough to undergo the procedure and have achieved a sufficient clinical response to chemotherapy.(11) A complication that can occur after SCT, predominantly alloSCT, is graft-versus-host disease (GVHD), which is an immune complication arising from the recognition and destruction of the recipient's tissues and organs by the donor's T cells.(42) GVHD, which can be acute or chronic, is associated with high morbidity and mortality. The potential for the complications associated with alloSCT lead to its use typically being limited to young patients who relapse quickly after high-dose chemotherapy and autoSCT, or those who are chemorefractory and achieve a response to checkpoint inhibitors.(11) Additionally, people may choose not to undergo SCT. Given that the patients forming the SACT dataset are older and less fit than those enrolled into cohort 2 of KEYNOTE-087, and that all patients are SCT-naïve, it could be expected that a larger proportion of people in the SACT dataset would have undergone autoSCT compared with alloSCT.

Time to SCT

An outcome of interest specified in the decision problem was time to alloSCT (Table 1). However, data are only available for combined time to autoSCT or alloSCT for both cohort 2 of KEYNOTE-087 and the SACT dataset. There was a considerable difference in time to SCT between the two cohorts.

For cohort 2 of KEYNOTE-087, the estimated mean time to SCT was 30.3 months (131.4 weeks).[41] For the SACT dataset, of patients eligible for SCT, median time to SCT was 17.5 months (532 days; Figure 4).[34] MSD note that time to SCT limited to those receiving an SCT is not reported separately.

Figure 4. Kaplan–Meier estimates for time to SCT in HES (N=132)[34]



A potential explanation for the marked difference in time to SCT could be variation in clinical practice between England and the other countries participating in KEYNOTE-087 in transitioning people from chemotherapy to SCT, particularly if a patient is deemed a potential candidate for SCT. As noted in Section B.2.3.1.3, only 10 people from the UK were enrolled into cohort 2, of whom five went on to receive SCT. However, it is unclear whether the five patients receiving SCT would be classed as a future candidate for SCT. MSD consider the time to SCT derived from the SACT dataset to be more generalisable to clinical practice in England.

B.2.6.1.4 PFS

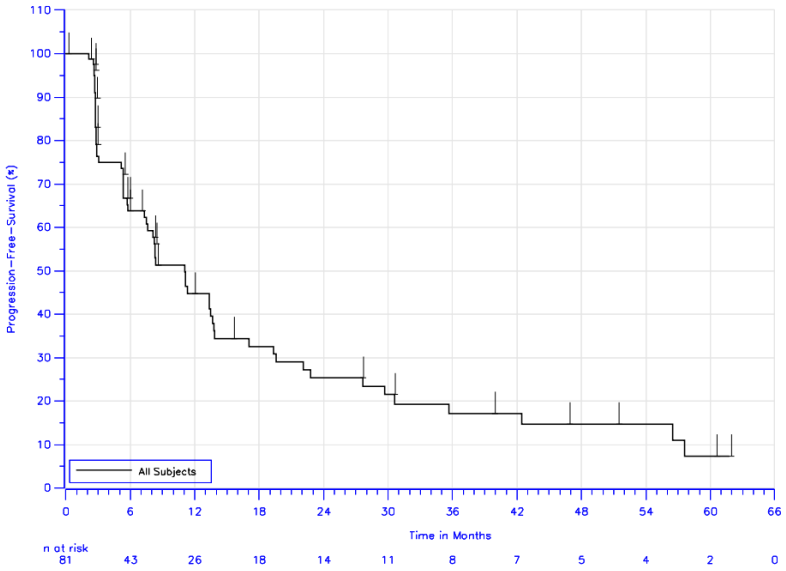
PFS was defined as the time from first dose to the first documented disease progression by BICR and by site assessment according to IWG criteria or death due to any cause, whichever occurred first.

Of the 81 people forming cohort 2, 57 (70.4%) experienced an event (Table 18). Median PFS in the ASaT population, as assessed by BICR and IWG criteria, was 11.1 months (95% CI: 7.5 to 13.7). PFS rates at various time points up to 60 months are presented in Table 18 and Kaplan–Meier estimates of PFS are depicted in Figure 5.

Table 18. Summary of PFS for cohort 2 based on central review as per IWG criteria[33]

Outcome	**Cohort 2 (N=81)**
Number of events, n (%)	57 (70.4)
Person-months	1080
Event rate/100 person-months (%)	5.3
Median PFS (95% CI), months	11.1 (7.5 to 13.7)
PFS rate at various time points	
• 6 months (%)[a]	63.8
• 12 months (%)[a]	44.7
• 24 months (%)[a]	25.4
• 36 months (%)[a]	17.2
• 48 months (%)[a]	14.7
• 60 months (%)[a]	7.4
[a] From product-limit (Kaplan–Meier) method for censored data. **Abbreviations:** BICR, blinded independent central review; CI, confidence interval; IWG, international working group; PFS, progression-free survival; SD, standard deviation.	

Figure 5. Kaplan–Meier estimates of PFS for cohort 2 based on BICR and IWG criteria[33]



B.2.6.1.5 OS

Of the 81 people forming cohort 2, after a median follow-up of 62.2 months, 24 (29.6%) had died (Table 19).[33] Mean OS for the cohort was 53.7 months (SE 1.8 months). Median OS was not reached.

Considering the SACT dataset, after a median follow-up time of 19.2 months, 73 (34.0%) of people had died (Table 19).[34] Median OS was not reached: mean OS was not available.

Table 19. Summary of OS for cohort 2 from KEYNOTE-087[33] and for the SACT dataset[34]

Outcome	Cohort 2 (N=81)	SACT dataset (N=215)
Number of events, n (%)	24 (29.6)	73 (34.0%)
Median OS (95% CI), months	Not reached	Not reached
OS rate at various time points		
• 6 months (%)	100.0[a]	88% (95% CI: 83% to 92%)
• 12 months (%)	96.3[a]	82% (95% CI: 76% to 87%)
• 18 months (%)	93.7[a]	75% (95% CI: 68% to 80%)
• 24 months (%)	91.1[a]	68% (95% CI: 61% to 75%)
• 36 months (%)	85.9[a]	56% (95% CI: 47% to 64%)

• 48 months (%)	76.5[a]	55% (95% CI: 46% to 63%)
• 60 months (%)	69.2[a]	N/A

* From product-limit (Kaplan–Meier) method for censored data.

Abbreviations: BICR, blinded independent central review; CI, confidence interval; IWG, international working group; N/A, not available; OS, overall survival; SACT, Systemic Anti-Cancer Therapy; SD, standard deviation.

Figure 6. Kaplan–Meier estimates of OS for cohort 2[33]

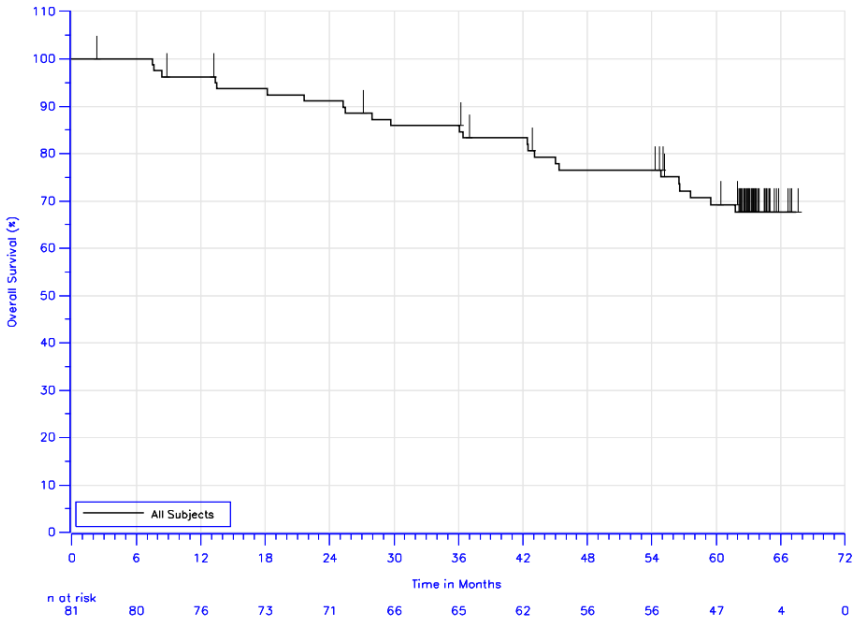
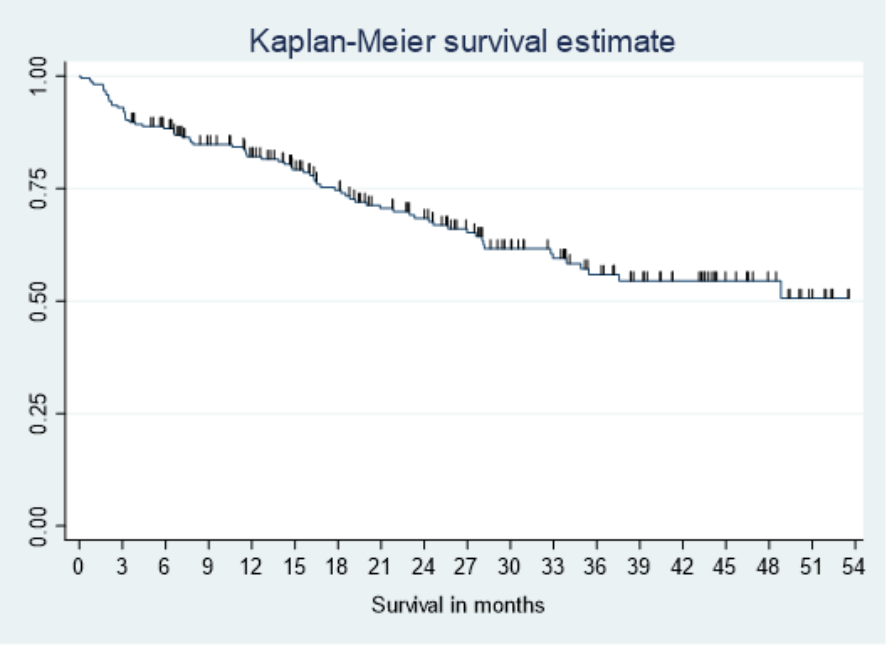


Figure 7. Kaplan–Meier estimates of OS for the SACT dataset[34]



As would be expected based on previously discussed differences in patient characteristics between the cohorts, OS is less favourable for the SACT dataset compared with cohort 2, in terms of proportion of people alive at set time points (Table 19). However, MSD consider that prognosis of those in the SACT dataset is better than those not receiving pembrolizumab in this setting (discussed in Section B.2.9).

OS rates at various time points up to 60 months for both cohorts are presented in Table 19 and Kaplan–Meier estimates of OS are depicted in Figure 6 for cohort 2 and Figure 7 for the SACT dataset.

OS for patients who did and did not receive a SCT

SCT can potentially be curative, even for those with R/R cHL. To evaluate the impact of SCT on OS, MSD carried out *post hoc* subgroup analyses for those in cohort 2 of KEYNOTE-087 based on who did and did not receive SCT. Additionally, NHS England helpfully provided data on OS for the subgroup of people from the SACT dataset who did not receive SCT. MSD note that, as *post hoc* subgroup analyses, the results are hypothesis generating and should be interpreted with a level of caution.

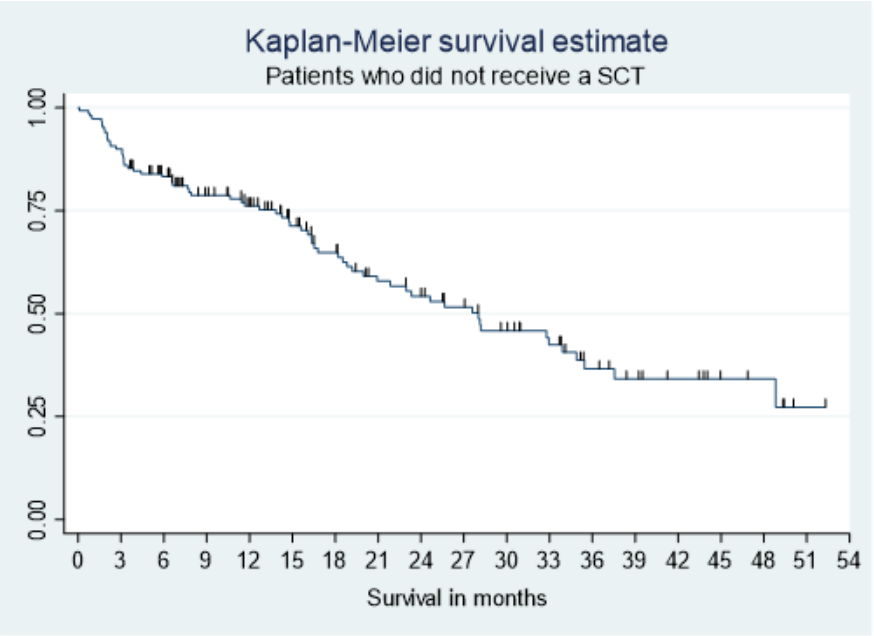
Of the 24 people who received SCT from cohort 2, 19 (79.2%) remained alive at the time of last follow-up.[41] Median OS was not reached. Mean OS for those receiving SCT was 53.7 months (SE 2.7 months), and mean OS after SCT was 42.5 months (SE 2.5 months). For the SACT cohort, of the 65 people identified as undergoing SCT, 59 (91%) remained alive at the time of data cutoff. Median and mean OS are not available.[34]

For those not receiving SCT, at the data cutoff, 38 (66.7%) from cohort 2 remained alive, and mean OS was 52.1 months (SE 2.3 months).[41] Again, median OS was not reached. Among the 150 (70%) patients from the SACT dataset identified as not receiving a SCT, median OS was 28 months (95% CI: 20.0 to 34.9 months; Figure 8).[34]

For cohort 2 of KEYNOTE-087, the mean OS is similar for those receiving and not receiving SCT. However, MSD note that, despite a median follow-up of 62.2 months,

median OS has not been reached in either subgroup, and data may be too immature to discern the impact of SCT on OS.

Figure 8. Kaplan–Meier survival plot for patients from the SACT dataset who did not receive SCT



B.2.7. Subgroup analysis

Prespecified subgroup analysis focused on evaluating whether ORR was consistent in the following groups:

- Age category (≤65 vs >65 years);
- Sex (female vs male);
- Race (white vs non-white);
- Region (US vs ex-US);
- Number of prior therapies (<4 vs ≥4).

If the observed numbers for a particular subgroup were too small to make a clinically meaningful interpretation, the subgroup analysis was not conducted. Data on ORR by subgroup are not available in the CSR.[33]

Subgroup analyses for the outcomes of PFS and OS were not prespecified, and are not available in the CSR for the groups detailed above.[33]

B.2.8. Meta-analysis

No head-to-head study evaluating pembrolizumab in R/R cHL after failure to respond to BV was identified and, thus, meta-analysis was not possible. Estimates of comparative clinical effectiveness for pembrolizumab versus standard of care (SoC) for OS were generated via unanchored matching adjusted indirect treatment comparison (MAIC), the results of which are presented in section B.2.9.

B.2.9. Indirect and mixed treatment comparisons

As noted in Section B.2.1, a SLR was carried out to identify relevant published evidence on the clinical effectiveness of pharmacological treatments for autoSCT-naïve R/R cHL patients who have received BV. The SLR identified only two relevant studies — KEYNOTE-087 and KEYNOTE-013 — both of which are single arm studies, and, thus, do not provide estimates of comparative clinical effectiveness for pembrolizumab. As no study was identified that evaluated a comparator of interest, as set out in the decision problem (Table 1), to inform the Technology Appraisal MSD chose to replicate the matching adjusted indirect comparison (MAIC) presented in TA540[(2)] using Cheah (2016) to represent SoC. MSD acknowledge the limitations associated with using Cheah (2016) (discussed in detail in Section B.2.9.3), but consider that the study remains the most appropriate dataset to generate estimates of comparative clinical effectiveness versus pembrolizumab for those with R/R cHL who are SCT-naïve and did not respond to BV. Because only OS is available from the SACT dataset, ITCs were carried out for only OS using both KEYNOTE-087 and the SACT dataset.

Full details of the SLR search strategy, study selection process and results are presented in Appendix D. Full details of the methodology followed for the MAIC are available in Appendix D.

B.2.9.1 Summary of the studies included in the MAIC

B.2.9.1.1 Cheah (2016)[(4)]

The study is a retrospective observational study designed to evaluate outcomes after treatment with BV in patients with cHL who were either refractory to BV or

experienced disease relapse. The study evaluated the records for patients treated at the MD Anderson Centre, USA, between June 2007 and January 2015.

Patients included were those with histologically confirmed cHL and who had received treatment with BV for relapsed cHL and subsequently experienced disease progression at any time after treatment with BV. In total, 97 patients met the study inclusion criteria. Baseline characteristics at the time of documented disease progression after treatment with BV were available for 89 patients. The full cohort (N=97) was predominantly male (53%) and had a median age of 28 years (range 16 to 83). Most patients had an ECOG score of 0 (84%) and had nodular sclerosing histology (88%). Many patients had stage III/IV cHL (56%) and B symptoms (60%) at the time of initial diagnosis.

Median number of prior lines of therapy was three, with a range of 0 to 9. Importantly, at the time of second remission, before treatment with BV, 70 (72%) people had undergone SCT, predominantly autoSCT (n=66), and an additional 10 (10.3%) patients received alloSCT after autoSCT but before BV. As noted by the EAG in TA540, the large proportion of people having undergone SCT does not align with the autoSCT naïve status of people enrolled into cohort 2 of KEYNOTE-087 or receiving pembrolizumab through the CDF.

The most common reason for discontinuation of BV was disease progression (n=76, 78%). Ten (10.3%) patients electively discontinued BV while in remission in order to receive a SCT. Treatments subsequent to BV comprised investigational agents, gemcitabine, bendamustine, another alkylator, BV retreatment, platinum-based treatment, and autoSCT.

The authors reported that 46 people from the cohort died, with a median OS of 25.2 months: OS was measured from the time of progression post-BV to death. The main cause of death was reported to be lymphoma (39%), but with an equal number of deaths attributed to unknown cause (39%).

Baseline characteristics, quality assessment and study results for Cheah (2016) are available in Appendix D. The single centre, retrospective design, US setting, youth and SCT history of the patients are major concerns. A priori, the patient level factors

just mentioned suggest these patients would have a favourable prognosis compared to the population treated with SoC in this Technology Appraisal.

B.2.9.2 OS for pembrolizumab versus SoC

Considering results for cohort 2 from KEYNOTE-087, the naïve and MAIC analyses generated similar estimates of OS for pembrolizumab versus SoC, with HRs of 0.25 and 0.24, respectively, favouring pembrolizumab (Table 20). Each analysis reached statistical significance, and associated p values were <0.001.

The reduction in risk of death associated with pembrolizumab compared with SoC is lower for the SACT dataset than for cohort 2, with a HR of 0.59 (Table 20). However, the difference between pembrolizumab and SoC remains statistically significant, with a 95% CI of 0.40 to 0.86 (Table 20).

Supporting evidence on the clinical effectiveness of pembrolizumab versus SoC in improving OS is derived from a Bucher ITC (described in more detail in Section 3.3.1.3), with BV as the common comparator. An HR for OS for pembrolizumab was derived from the subgroup of patients from KEYNOTE-204 who were SCT-naïve at baseline, and an HR for OS for BV versus SoC was estimated from data presented in TA524. The Bucher comparison generated an HR for OS of [redacted] in favour of pembrolizumab, and the 95% CI does not cross the line of no effect. MSD appreciate that the cohort of patients informing the Bucher ITC are not receiving 4L treatment after BV, but consider that this analysis lends support to the clinical effectiveness of pembrolizumab in R/R cHL.

MSD also investigated comparative clinical effectiveness of pembrolizumab versus BV through MAIC using data on clinical outcomes from KEYNOTE-204 and Eyre (2017), which evaluated BV in SCT-naïve patients, a population MSD considers is perhaps most analogous to those forming cohort 2 of KEYNOTE-087. Although BV is not a comparator of interest as specified in the decision problem (Table 1), MSD consider that the MAICs versus BV are informative for decision-making, and include the estimates of effect in economic analysis (see Section B.3.2). Full results of the MAICs are presented in Table 33 and Appendix D2.3.6.2.

Kaplan–Meier curves generated for the indirect comparison of pembrolizumab versus Soc are presented in Figure 9 for cohort 2 and Figure 10 for the SACT dataset.

Table 20. Summary of OS estimates derived from indirect comparisons of pembrolizumab versus SoC

Dataset	Comparison	Sample size/ effective sample size, n	Pembro Events, n (%)	Post BV[a] (N=89) Events, n (%)	Hazard ratio (95% CI),[b] p value[b,c]
Cohort 2	Unadjusted	81	24 (29.6)	46 (58.2)	0.25 (0.15 to 0.41) <0.001
Cohort 2	MAIC	75.4[d]	21 (27.9)	46 (58.2)	0.24 (0.14 to 0.40) <0.001
SACT[e]	Unadjusted	215	68 (31.6)	46 (58.2)	0.59 (0.40 to 0.86) 0.006

[a] Based on mAPaT population from Cheah (2016).

[b] Based on Cox regression model with treatment as a single covariate.

[c] Two-sided p-value using Wald test (Score test in case of zero event in one treatment group).

[d] Effective sample size computed as sum of weights.

[e] Based on Real World Evidence source (SACT database) in participants relapsing after treatment with or failing to respond to BV and ineligible for SCT, mAPaT.

Abbreviations: BV, brentuximab vedotin; CI, confidence interval; HR, hazard ratio; MAIC, matching adjusted indirect comparison; mAPaT, defined as analysis populations used to report comparator study results; SACT, systemic anti-cancer therapy.

Figure 9. Kaplan–Meier curve for OS for pembrolizumab versus SoC with pembrolizumab data derived from Cohort 2 from KEYNOTE-087

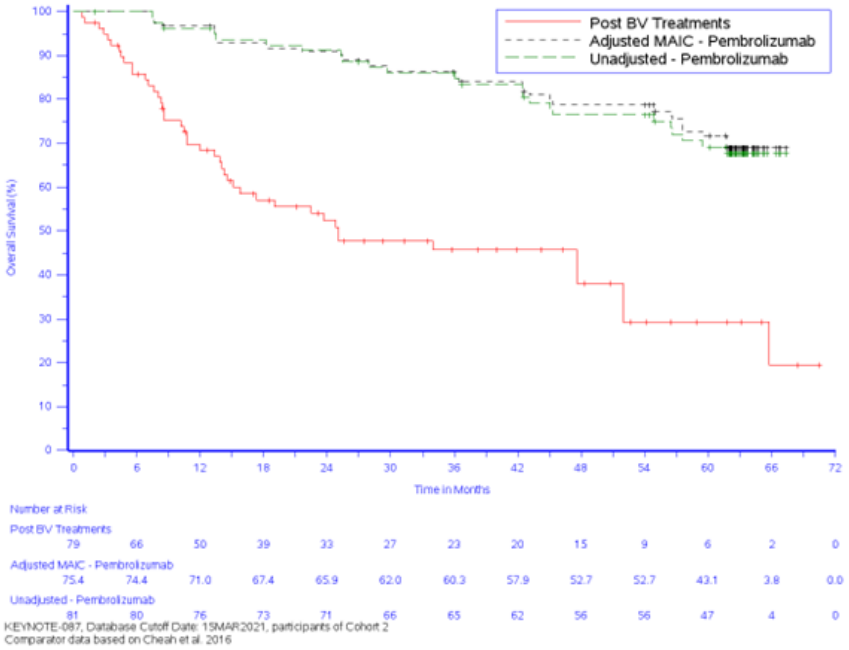
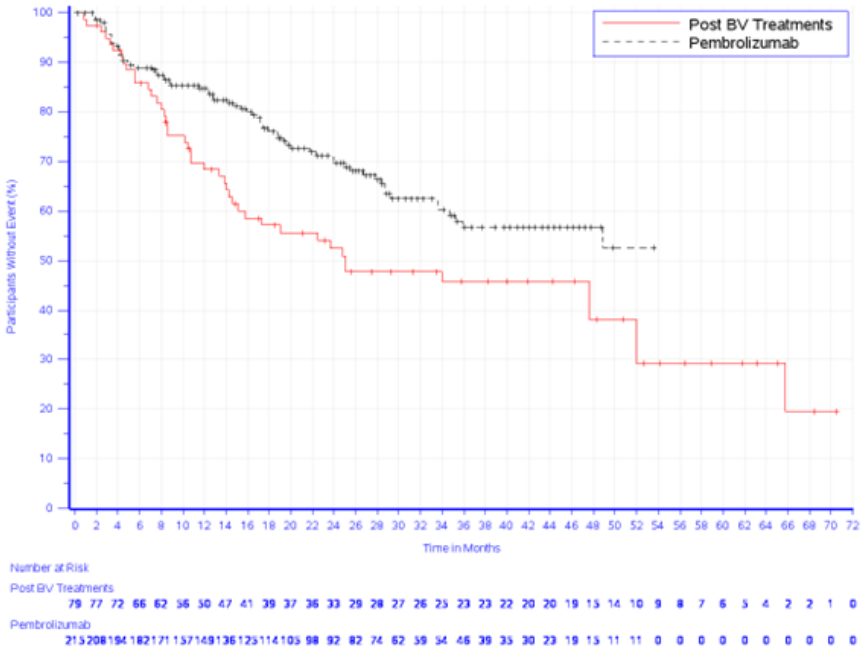


Figure 10. Kaplan–Meier curve for OS for pembrolizumab versus SoC with pembrolizumab data derived from the SACT dataset



B.2.9.3 Limitations associated with the indirect comparison

The identification of only single-arm and observational studies relevant to the decision problem meant that comparative analyses were limited to unanchored indirect comparisons.

Naïve-unadjusted and population-adjusted analyses were possible for the ITC of pembrolizumab data derived from cohort 2 of KEYNOTE-087 versus SoC. However, due to a lack of individual patient data for either study, only a naïve-unadjusted comparison was feasible for pembrolizumab data generated from the SACT cohort versus SoC.

Of the two types of analysis, given the differences across the studies in terms of design and baseline characteristics, MSD acknowledges that the naïve-unadjusted ITCs should be interpreted with caution. Considering study design, KEYNOTE-087 and the SACT dataset are prospective in nature, whereas Cheah (2016) is a retrospective review of patient records, and, as such, is open to, for example, selection and recall bias, as well as inaccurate or incomplete recording of information. Additionally, naïve-unadjusted ITCs do not match patient populations, and the cohorts evaluated in the three studies differ in patient characteristics that could be prognostic factors or treatment effect modifiers. Of note, considering age and ECOG score, patients in cohort 2 were older but fitter than those in Cheah (2016), in contrast to those forming the SACT dataset, who were considerably older and less fit than patients in Cheah (2016):

- Proportion of people aged >45 years: 43.2% in cohort 2 versus ≥55% in the SACT dataset versus 14% in Cheah (2016)
- ECOG performance status:

  0: 54.3% in cohort 2 versus 27% in SACT dataset versus 41% in Cheah (2016);

  1: 45.7% in cohort 2 versus 40% in SACT dataset versus 54% in Cheah (2016).

A key characteristic across the studies, that cannot be adjusted for is the baseline SCT-naïve status of patients in cohort 2 and the SACT dataset, whereas 72% of patients in Cheah (2016) had undergone SCT.

Considering the MAIC for pembrolizumab versus SoC, as detailed in the NICE DSU report on carrying out population-adjusted analyses in submissions to NICE,[43] in unanchored comparisons it is necessary to assume that all effect modifiers and prognostic factors have been accounted for, which is considered impossible to do.

MAICs adjust the characteristics of the population for which IPD data are available to those of the comparator population. Thus, the variables matched in the MAIC for pembrolizumab versus SoC were determined by the characteristics reported in Cheah (2016), which are unlikely to be all relevant prognostic variables and treatment effect modifiers for OS. Therefore, the results are likely to contain systematic error, as per DSU guidance,[43] but it is not possible to quantify the extent of any potential errors.

Other limitations with the use of Cheah 2016, as highlighted by the authors of a study comparing pembrolizumab versus SoC in R/R cHL,[44] include the acknowledged selection bias within the study, which has implications for the generalisability of results, and the diversity noted in post-progression treatment regimens, which do not reflect clinical practice in England. However, as mentioned in Section B.2.1, no new relevant study was identified in MSD's SLR, and, so, Cheah (2016) remains the most relevant study to inform estimates of comparative clinical effectiveness for this Technology Appraisal.

MSD acknowledge that there are limitations associated with the ITCs of pembrolizumab versus SoC that lead to a level of uncertainty in the comparative effect estimates for OS. However, MSD note that all results from the ITCs for OS favour pembrolizumab and reach statistical significance. The uncertainty in clinical effectiveness impacts on estimates of cost effectiveness of pembrolizumab in the management of R/R cHL after BV. Uncertainty in cost effectiveness is explored through scenario analyses (please see Section B.3.11).

B.2.10. Adverse reactions

Safety analyses were based on the ASaT population up to the data cutoff of 15 March 2021, which corresponds to approximately 5 years after the last patient initiated study treatment.

B.2.10.1 KEYNOTE-087: extent of exposure

Patients were exposed to pembrolizumab for a median of 387.0 days (range: 1 to 1880), resulting in a median of 18.5 administrations (range: 1 to 52; Table 21).[33] Overall, many participants (n=155) remained on pembrolizumab for ≥6 months and

approximately half (n=108) remained on pembrolizumab for ≥12 months. The median duration of exposure in cohort 2 was 254.0 days, with a median of 13.0 administrations and 63.0% of patients remaining on treatment for ≥6 months (Table 21).

Table 21. Summary of drug and clinical trial exposure for cohort 2 and for the full trial population of KEYNOTE-087[33]

Characteristic	Cohort 2 (N=81)		KEYNOTE-087 (N=210)	
Number of days on therapy				
Mean (SD)	393.2 (342.4)		461.4 (355.6)	
Median (range)	254.0 (1 to 1696)		387.0 (1 to 1880)	
Number of administrations				
Mean (SD)	17.5 (12.8)		20.3 (12.8)	
Median (range)	13.0 (1 to 52)		18.5 (1 to 52)	
Duration of exposure				
	n (%)	Person-years	n (%)	Person-years
>0 months	81 (100)	87.2	210 (100)	265.3
≥1 months	80 (98.8)	87.2	206 (98.1)	265.1
≥3 months	73 (90.1)	85.9	194 (92.4)	262.8
≥6 months	51 (63.0)	77.3	155 (73.8)	248.2
≥12 months	35 (43.2)	66.6	108 (51.4)	214.9
Abbreviation: SD, standard deviation.				

B.2.10.2 Summary of adverse events

In general, pembrolizumab was well tolerated by patients with R/R cHL, with a manageable safety profile. The rate of AEs was not unexpected for this heavily pre-treated patient group. Across cohorts, 205 of 210 (97.6%) patients experienced at least one AE (Table 22). Most AEs were of low-grade, as evidenced by the relatively low rate of patients with AEs categorised as Grade 3, 4, or 5 (please see Section B.2.10.5). AEs resulting in death occurred in 3 (1.4%; Table 22) people and were attributed to one each of acute GVHD, post-procedural infection, and septic shock. No death was deemed to be related to a drug-related AE (Table 22).

Table 22. Summary of adverse events for KEYNOTE-087 by cohort for the ASaT population[33]

Characteristic	Cohort 1 (N=69) n (%)	Cohort 2 (N=81) n (%)	Cohort 3 (N=60) n (%)	Total (N=210) n (%)
One or more AE, n (%)	68 (98.6)	80 (98.8)	57 (95.0)	205 (97.6)
Drug-related[a] AE, n (%)	54 (78.3)	52 (64.2)	47 (78.3)	153 (72.9)
Toxicity grade 3–5 AEs, n (%)	22 (31.9)	26 (32.1)	21 (35.0)	69 (32.9)
Toxicity grade 3–5 drug-related AEs, n (%)	12 (17.4)	10 (12.3)	5 (8.3)	27 (12.9)
Non-serious AEs, n (%)	68 (98.6)	79 (97.5)	57 (95.0)	204 (97.1)
Serious AEs, n (%)	15 (21.7)	18 (22.2)	15 (25.0)	48 (22.9)
Serious drug-related AEs, n (%)	8 (11.6)	4 (4.9)	5 (8.3)	17 (8.1)
Died, n (%)	0 (0.0)	2 (2.5)	1 (1.7)	3 (1.4)
Died due to a drug-related AE, n (%)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Discontinued[b] due to an AE, n (%)	8 (11.6)	5 (6.2)	5 (8.3)	18 (8.6)
Discontinued due to a drug-related AE, n (%)	6 (8.7)	4 (4.9)	4 (6.7)	14 (6.7)
Discontinued due to a serious AE, n (%)	5 (7.2)	3 (3.7)	2 (3.3)	10 (4.8)
Discontinued due to a serious drug-related AE, n (%)	3 (4.3)	2 (2.5)	2 (3.3)	7 (3.3)

[a] Determined by the investigator to be related to the drug.

[b] Study medication withdrawn.

Grades are based on NCI CTCAE version 4.0. Non-serious AEs up to 30 days of last dose and serious AEs 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.

Abbreviations: ASaT, all subjects as treated; AE, adverse event; MedDRA, Medical Dictionary for Regulatory Activities; NCI CTCAE, National Cancer Institute Common Terminology Criteria for Adverse Events.

B.2.10.3 Most frequently reported adverse events

The most frequently reported AEs (>15%) across all cohorts included pyrexia (30.0%), cough (26.2%), fatigue (22.9%), diarrhoea (20.5%) and upper respiratory

tract infection (20.5%). An overview of the ten most commonly experienced AEs is available in Table 23, with the full list presented in Appendix F.

Table 23. Overview of the most frequently reported adverse events (>15%) across cohorts from KEYNOTE-087[33]

Adverse effect	Cohort 1 (N=69) n (%)	Cohort 2 (N=81) n (%)	Cohort 3 (N=60) n (%)	Total (N=210) n (%)
Pyrexia	27 (39.1)	19 (23.5)	17 (28.3)	63 (30.0)
Cough	19 (27.5)	22 (27.2)	14 (23.3)	55 (26.2)
Fatigue	15 (21.7)	17 (21.0)	16 (26.7)	48 (22.9)
Diarrhoea	20 (29.0)	12 (14.8)	11 (18.3)	43 (20.5)
Upper respiratory tract infection	23 (33.3)	7 (8.6)	13 (21.7)	43 (20.5)
Nausea	16 (23.2)	11 (13.6)	11 (18.3)	38 (18.1)
Vomiting	16 (23.2)	9 (11.1)	13 (21.7)	38 (18.1)
Nasopharyngitis	12 (17.4)	16 (19.8)	7 (11.7)	35 (16.7)
Arthralgia	14 (20.3)	12 (14.8)	7 (11.7)	33 (15.7)
Hypothyroidism	8 (11.6)	13 (16.0)	12 (20.0)	33 (15.7)
Non-serious AEs up to 30 days of last dose and serious AEs 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. **Abbreviations:** AE, adverse event; MedDRA, Medical Dictionary for Regulatory Activities.				

B.2.10.4 Drug-related adverse events

Of the 210 patients treated in KEYNOTE-087, 153 (72.9%) experienced ≥1 treatment-related AE (Table 24). The most frequently reported drug-related AEs across all cohorts included hypothyroidism (14.3%), pyrexia (11.4%), fatigue (11.0%), and rash (11.0%). An overview of the ten most commonly experienced drug-related AEs is available in Table 24, with the full list presented in Appendix F.

Table 24. Overview of the most frequently reported drug-related adverse events (incidence ≥5% in one or more treatment groups) across cohorts from KEYNOTE-087[33]

Adverse effect	Cohort 1 (N=69) n (%)	Cohort 2 (N=81) n (%)	Cohort 3 (N=60) n (%)	Total (N=210) n (%)

Experienced ≥1 AE	54 (78.3)	52 (64.2)	47 (78.3)	153 (72.9)
Hypothyroidism	6 (8.7)	12 (14.8)	12 (20.0)	30 (14.3)
Pyrexia	12 (17.4)	7 (8.6)	5 (8.3)	24 (11.4)
Fatigue	10 (14.5)	6 (7.4)	7 (11.7)	23 (11.0)
Rash	10 (14.5)	5 (6.2)	8 (13.3)	23 (11.0)
Diarrhoea	9 (13.0)	4 (4.9)	4 (6.7)	17 (8.1)
Headache	10 (14.5)	3 (3.7)	3 (5.0)	16 (7.6)
Nausea	7 (10.1)	2 (2.5)	6 (10.0)	15 (7.1)
Arthralgia	4 (5.8)	5 (6.2)	4 (6.7)	13 (6.2)
Cough	3 (4.3)	4 (4.9)	6 (10.0)	13 (6.2)
Pruritus	4 (5.8)	5 (6.2)	4 (6.7)	13 (6.2)
Non-serious AEs up to 30 days of last dose and serious AEs 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. **Abbreviations:** AE, adverse event; MedDRA, Medical Dictionary for Regulatory Activities.				

B.2.10.5 Grade 3 to 5 adverse events

Across all cohorts of KEYNOTE-087, most patients (n=136; 64.8%) experienced AEs that were a maximum toxicity of Grade 1 or 2.[33] Considering Grade 3–5 events, AEs of Grade 3, 4, and 5 were reported by 26.2%, 5.2%, and 1.4% of participants, respectively. Twenty-seven patients (12.9%) experienced ≥1 Grade 3 or 4 AE that was considered related to pembrolizumab, but no drug-related Grade 5 AE was reported. The most commonly reported Grade 3–5 AE was anaemia (69 patients; Table 25), whereas the most frequent drug-related Grade 3 or 4 AE was neutropenia (5 patients). There were no meaningful differences in rates of Grade 3–5 AEs across cohorts. An overview of the ten most commonly experienced Grade 3–5 AEs (any cause and drug-related by organ class) is available in Table 25, with the full list presented in Appendix F.

Table 25. Overview of Grade 3–5 adverse effects (incidence >0% in one or more treatment groups) experienced across cohorts in KEYNOTE-087[33]

Organ class	Cohort 1 (N=69) n (%)	Cohort 2 (N=81) n (%)	Cohort 3 (N=60) n (%)	Total (N=210) n (%)

Any cause				
Experienced ≥1 AE Grade 3–5	22 (31.9)	26 (32.1)	21 (35.0)	69 (32.9)
Anaemia	4 (5.8)	3 (3.7)	0 (0.0)	7 (3.3)
Neutropenia	3 (4.3)	3 (3.7)	0 (0.0)	6 (2.9)
Pneumonia	4 (5.8)	1 (1.2)	0 (0.0)	5 (2.4)
Diarrhoea	3 (4.3)	1 (1.2)	0 (0.0)	4 (1.9)
Herpes zoster	0 (0.0)	2 (2.5)	1 (1.7)	3 (1.4)
Leukopenia	1 (1.4)	2 (2.5)	0 (0.0)	3 (1.4)
Pyrexia	2 (2.9)	0 (0.0)	1 (1.7)	3 (1.4)
Thrombocytopenia	0 (0.0)	2 (2.5)	1 (1.7)	3 (1.4)
Acute graft versus host disease	0 (0.0)	1 (1.2)	1 (1.7)	2 (1.0)
Alanine aminotransferase increased	1 (1.4)	1 (1.2)	0 (0.0)	2 (1.0)
Drug-related				
Experienced ≥1 AE Grade 3–5	12 (17.4)	10 (12.3)	5 (8.3)	27 (12.9)
Blood and lymphatic system disorders	3 (4.3)	2 (2.5)	1 (1.7)	6 (2.9)
Cardiac disorders	3 (4.3)	0 (0.0)	0 (0.0)	3 (1.4)
Gastrointestinal disorders	1 (1.4)	2 (2.5)	0 (0.0)	3 (1.4)
General disorders and administration site conditions	1 (1.4)	1 (1.2)	1 (1.7)	3 (1.4)
Hepatobiliary disorders	1 (1.4)	0 (0.0)	0 (0.0)	1 (0.5)
Immune system disorders	0 (0.0)	2 (2.5)	0 (0.0)	2 (1.0)
Infections and infestations	2 (2.9)	1 (1.2)	2 (3.3)	5 (2.4)
Investigations	2 (2.9)	1 (1.2)	0 (0.0)	3 (1.4)
Metabolism and nutrition disorders	1 (1.4)	0 (0.0)	0 (0.0)	1 (0.5)
Musculoskeletal and connective tissue disorders	2 (2.9)	1 (1.2)	0 (0.0)	3 (1.4)
Non-serious AEs up to 30 days of last dose and serious AEs 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. Grades are based on NCI CTCAE version 4.0.				

Abbreviations: AE, adverse event; MedDRA, Medical Dictionary for Regulatory Activities; NCI CTCAE, National Cancer Institute Common Terminology Criteria for Adverse Events.

B.2.10.6 Other serious adverse events

Overall, 48 patients (22.9%; Table 26) experienced a SAE during study treatment and through 90 days after the last dose of pembrolizumab (Table 26). The most commonly reported SAE was pneumonia (2.9%). Three cases of acute GVHD were recorded, one of which was fatal. An overview of the ten most common SAEs is available in Table 26, with the full list presented in Appendix F.

Table 26. Overview of serious adverse effects incidence ≥1% in one or more treatment groups) experienced across cohorts in KEYNOTE-087[33]

SAE	Cohort 1 (N=69) n (%)	Cohort 2 (N=81) n (%)	Cohort 3 (N=60) n (%)	Total (N=210) n (%)
Experienced ≥1 SAE	15 (21.7)	18 (22.2)	15 (25.0)	48 (22.9)
Pneumonia	4 (5.8)	1 (1.2)	1 (1.7)	6 (2.9)
Pneumonitis	1 (1.4)	1 (1.2)	2 (3.3)	4 (1.9)
Pyrexia	0 (0.0)	1 (1.2)	3 (5.0)	4 (1.9)
Acute graft versus host disease	1 (1.4)	1 (1.2)	1 (1.7)	3 (1.4)
Bronchitis	0 (0.0)	1 (1.2)	1 (1.7)	2 (1.0)
Herpes zoster	0 (0.0)	2 (2.5)	0 (0.0)	2 (1.0)
Pericarditis	2 (2.9)	0 (0.0)	0 (0.0)	2 (1.0)
Acute kidney injury	1 (1.4)	0 (0.0)	0 (0.0)	1 (0.5)
Acute sinusitis	0 (0.0)	1 (1.2)	0 (0.0)	1 (0.5)
Anaemia	1 (1.4)	0 (0.0)	0 (0.0)	1 (0.5)

Non-serious AEs up to 30 days of last dose and serious AEs 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.
Abbreviations: MedDRA, Medical Dictionary for Regulatory Activities; SAE, serious adverse event.

B.2.10.7 Adverse events leading to discontinuation of pembrolizumab

Eighteen (8.6%; Table 27) patients enrolled in KEYNOTE-087 discontinued pembrolizumab due to one or more AEs. The most commonly reported AEs that resulted in stopping pembrolizumab were pneumonitis (n=7; 3.3%) and infusion-

related reaction (n=2; 1.0%). Rates of AEs leading to cessation of pembrolizumab were generally similar across cohorts. Fourteen patients (6.7%; Table 27) discontinued pembrolizumab due to a drug-related AE. The most commonly reported drug-related AEs resulting in treatment discontinuation were pneumonitis (n=7; 3.3%) and infusion-related reaction (n=2; 1.0%). An overview of AEs leading to discontinuation by organ class is available in Table 27, with the full list presented in Appendix F.

Table 27. Overview of adverse events (incidence >0% in one or more treatment groups) leading to discontinuation of pembrolizumab across cohorts of KEYNOTE-087[33]

Organ class	Cohort 1 (N=69) n (%)	Cohort 2 (N=81) n (%)	Cohort 3 (N=60) n (%)	Total (N=210) n (%)
AEs resulting in discontinuation				
Experienced ≥1 AE	8 (11.6)	5 (6.2)	5 (8.3)	18 (8.6)
Cardiac disorders	1 (1.4)	0 (0.0)	0 (0.0)	1 (0.5)
Immune system disorders	0 (0.0)	1 (1.2)	0 (0.0)	1 (0.5)
Infections and infestations	1 (1.4)	0 (0.0)	1 (1.7)	2 (1.0)
Injury, poisoning and procedural complications	1 (1.4)	1 (1.2)	0 (0.0)	2 (1.0)
Musculoskeletal and connective tissue disorders	2 (2.9)	0 (0.0)	0 (0.0)	2 (1.0)
Drug-related AEs resulting in discontinuation				
Experienced ≥1 AE	6 (8.7)	4 (4.9)	4 (6.7)	14 (6.7)
Cardiac disorders	1 (1.4)	0 (0.0)	0 (0.0)	1 (0.5)
Immune system disorders	0 (0.0)	1 (1.2)	0 (0.0)	1 (0.5)
Infections and infestations	1 (1.4)	0 (0.0)	0 (0.0)	1 (0.5)
Injury, poisoning and procedural complications	1 (1.4)	1 (1.2)	0 (0.0)	2 (1.0)
Musculoskeletal and connective tissue disorders	2 (2.9)	0 (0.0)	0 (0.0)	2 (1.0)

Non-serious AEs up to 30 days of last dose and serious AEs 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.
Abbreviations: AE, adverse event; MedDRA, Medical Dictionary for Regulatory Activities.

B.2.10.8 Adverse events of special interest

Of the 210 patients treated in KEYNOTE-087, 70 (33.3%; Table 28) experienced ≥1 AEOSI. Most patients (62; 29.5%) experienced AEOSIs that were a maximum toxicity of Grade 1 or 2 severity, and there were no Grade 5 AEOSIs. One (0.5%) patient had Grade 4 myocarditis and Grade 3 necrotizing myositis, and 8 (3.8%) experienced Grade 3 AEOSIs. Sixty-seven (31.9%) patients reported an AEOSI that was considered related to pembrolizumab, 7 (3.3%) of whom experienced Grade 3 or 4 treatment-related AEOSIs. Fourteen (6.7%) patients discontinued pembrolizumab due to an AEOSI, 13 (6.2%) of whom stopped treatment due to a drug-related AEOSI. No death was attributed to an AEOSI. The incidence and type of AEOSIs remained generally consistent in KEYNOTE-087 over time and no new AEOSI was identified. The most common AEOSIs reported during the study were hypothyroidism (n=33; 15.7%), infusion-related reaction (n=11; 5.2%), and pneumonitis (n=10; 4.8%). An overview of the most commonly reported AEOSIs is available in Table 28, with the full list presented in Appendix F.

Table 28. Overview of adverse events of special interest (incidence >0% in one or more treatment groups) across cohorts of KEYNOTE-087[33]

AEOSI by category	Cohort 1 (N=69) n (%)	Cohort 2 (N=81) n (%)	Cohort 3 (N=60) n (%)	Total (N=210) n (%)
Experienced ≥1 AEOSI	21 (30.4)	26 (32.1)	23 (38.3)	70 (33.3)
Hypothyroidism	8 (11.6)	13 (16.0)	12 (20.0)	33 (15.7)
Infusion reactions	6 (8.7)	7 (8.6)	6 (10.0)	19 (9.0)
Pneumonitis	3 (4.3)	4 (4.9)	4 (6.7)	11 (5.2)
Hyperthyroidism	1 (1.4)	4 (4.9)	3 (5.0)	8 (3.8)
Colitis	2 (2.9)	1 (1.2)	0 (0.0)	3 (1.4)
Uveitis	3 (4.3)	0 (0.0)	0 (0.0)	3 (1.4)
Myositis	2 (2.9)	0 (0.0)	0 (0.0)	2 (1.0)
Skin	1 (1.4)	1 (1.2)	0 (0.0)	2 (1.0)

Non-serious AEs up to 30 days of last dose and serious AEs 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.
Abbreviations: AEOSI, adverse event of special interest; MedDRA, Medical Dictionary for Regulatory Activities.

B.2.11. Ongoing studies

KEYNOTE-087 is ongoing, with the next database lock anticipated in [redacted]. One additional trial in R/R cHL that is anticipated to provide additional evidence in the next 12 months was identified, and a summary is presented in Appendix M. The ongoing trial does not fully align with the population of interest specified in the decision problem. KEYNOTE-087 was also identified as an ongoing study.

B.2.12. Interpretation of clinical effectiveness and safety evidence

Clinical expectations for patients with R/R cHL treated with SoC after BV are low.[5] After rounds of unsuccessful chemotherapy and BV, few patients are expected to get to SCT. OS in the most relevant studies identified was short — 4-year OS in the Cheah (2016) and Eyre (2017) studies was ~40%, both of which included patients with a more favourable prognosis than those who would be eligible for pembrolizumab in the proposed setting. The cohort from Eyre (2017) who failed BV and did not get an SCT, which is perhaps the subgroup most analogous to the population of interest to the decision problem, had a 2-year OS of just 25%.

After a median follow-up of ~5 years, the evidence on clinical effectiveness derived from cohort 2 of KEYNOYE-087 underscores the benefit of pembrolizumab as a treatment for the management of those with R/R cHL who are ineligible for autoSCT after BV, a population that typically has a poor prognosis and has exhausted available treatment options. At 5 years of follow-up, OS was at >70%, estimated mean OS for cohort 2 was 53.7 months (SE 1.87 months). OS data from pembrolizumab's time in the CDF substantiate the findings from KEYNOTE-087, with 66.0% of people alive after a median follow-up of 19.2 months, during which time median OS was not reached with OS at 4-years being ~56%. MSD acknowledge that cohort 2 from KEYNOTE-087 is younger and fitter than the population forming the SACT dataset and consider that the SACT dataset is more generalisable to patients in England who would be eligible for treatment with pembrolizumab in this setting.

MSD consider that the benefits of pembrolizumab in the treatment of R/R cHL may extend beyond improving clinical outcomes after initial treatment. As highlighted by clinical experts consulted at an MSD UK advisory board, there is some evidence, albeit from retrospective analyses, that treatment with checkpoint inhibitors

chemosensitises some patients to their next treatment,[45-51] meaning they could potentially achieve the level of response required to be considered for SCT.

About 30% of patients enrolled in cohort 2 of KEYNOTE-087 in the SACT dataset proceeded to SCT, which is potentially curative. A retrospective study of heavily pretreated patients with R/R cHL (N=78) found that treatment with PD-1 inhibitor prior to autoSCT was associated with an 18-month PFS of 81% (95% CI: 69% to 89%) and OS of 96% (95% CI: 87% to 99%).[28] Interestingly, patients who responded to PD-L1 inhibitor had particularly favorable outcomes, with an 18-month PFS of 88%. A second small retrospective study in patients with R/R cHL (N=13) who received autoSCT as consolidation therapy after checkpoint inhibitor found that 11 (84.6%) patients obtained a CR. After a median follow-up of 3.3 years, only one patient who achieved CR experienced relapse, which occurred 3.9 months after autoSCT.[52] Similar outcomes have been reported for those undergoing alloSCT after PD-1 inhibitor, with one study (N=209) recording 2-year OS of 82% (95% CI: 76% to 87%).[53]

MSD carried out ITCs versus SoC for OS using data derived from cohort 2 of KEYNOTE-087 and from the SACT dataset, all of which generated effect estimates in favour of pembrolizumab and all of which reached statistical significance. Due to the data available, only unanchored ITCs could be carried out. MSD acknowledge the limitations associated with the ITCs but consider they support MSD's position that pembrolizumab is a clinically effective treatment option for a patient group with few choices.

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 most AEs experienced were low grade, and did not result in study discontinuation.

B.3. Cost effectiveness

Summary of key cost effectiveness information

Conclusions of cost-effectiveness analysis:

- Pembrolizumab significantly improves both OS and the probability of receiving a potentially curative SCT vs. SoC alone. The base case ICER is [redacted] gained vs. SoC. The model's conclusions are robust to sensitivity analysis with all scenarios having ICERs below [redacted] gained.

Model structure:

- The model uses a landmark structure. It has 2-states (alive/dead) prior to the landmark and 3 states (cured-SCT, no/failed SCT and dead) beyond it.

Model inputs:

Patient population inputs:

- Patients with R/R cHL who have had brentuximab vedotin and have not had an SCT. Outcomes on pembrolizumab are drawn from the SACT database.

Clinical efficacy inputs:

- OS HR prior to the landmark (source: various Indirect Treatment Comparisons)
- Differential probability of SCT (source: Structured Expert Elicitation [SEE])
- Differential probability that SCT is curative (source: SEE)
- Differential OS post landmark for patients in No/Failed SCT health state (source: KEYNOTE-204)

Utility inputs:

- The KEYNOTE-204 trial (pembrolizumab vs. BV in SCT-3L+ cHL)
- General population utility for cured patients

Costs and resource use inputs:

- NHS reference costs
- Various cHL studies, NICE appraisals and clinical expert advice

Base-case results and sensitivity analyses:

- The model's base case ICER is [redacted] gained.
- Extensive sensitivity analyses have been conducted, including pessimistic scenarios where a number of conservative assumptions have been combined. Pembrolizumab remains cost-effective in all these analyses.

Cost effectiveness conclusions:

- The model demonstrates pembrolizumab to be a cost-effective treatment in this indication with a base case ICER significantly below the cost-effectiveness threshold. The ICER remains below the threshold in all plausible sensitivity analyses. The cost-effectiveness is principally driven by the OS HR prior to the landmark, the magnitude of which is significant in all indirect treatment comparisons, even though all contain *a priori* biases against pembrolizumab. A number of other treatment effects can be removed entirely without much effect on the ICER.

B.3.1. Published cost-effectiveness studies

An SLR was conducted to identify published cost-effectiveness studies for pembrolizumab or other R/R cHL therapies to inform the original TA540 submission, but found no economic evaluations for the population of interest. In addition, there were no NICE technology appraisals in cHL available at the time and therefore a de novo model was required. For this CDF exit, a comprehensive SLR update was conducted, in line with the NICE methods guide,[54] to identify and summarise new published cost-effectiveness analyses. The SLR update was conducted on 20th February 2023, based on the methodology of the TA540 submission.[2] Given the small patient population in this late line of therapy (4L), only one economic evaluation was identified (Jones 2017).[55] Jones et al. (2017)[55] conducted an economic evaluation of Nivolumab versus SoC (derived using Cheah et al. 2016[4]). The study was published as an abstract only and therefore did not provide enough

data to inform the choice of model structure or inputs. Further information on the SLR methodology, search strategy and results is provided in Appendix G.

B.3.2. Economic analysis

B.3.2.1 Patient population

As stated in B.1, the patient group contained in the model is as per the NICE final scope i.e. patients with R/R cHL who have had BV and cannot have autoSCT.[3] The marketing authorisation for pembrolizumab is broader than this, including all patients R/R cHL who have failed autoSCT or following at least two prior therapies when autoSCT is an unsuitable treatment option.[7] That the population is narrower in this appraisal is due to two factors. Firstly, treatment with pembrolizumab for patients who have failed autoSCT or who cannot have auto SCT and who have not been treated with BV was considered in TA772,[26] which recommended pembrolizumab in this population. Second, the NICE Appraisal Committee's decision in TA540, was to recommend pembrolizumab via the CDF for a subpopulation only i.e. as an option for treating R/R cHL in adults who have had BV and cannot have autoSCT.[2] Within the same guidance, pembrolizumab was not recommended for patients with R/R cHL in adults who have had BV and autoSCT. Therefore, the population of interest considered in this economic evaluation is as per the NICE final scope.[3]

The baseline characteristics of the patients in the model are presented in Table 29.

Table 29. Baseline patient parameters in the KN-087 model

Characteristic	Mean	Source
Baseline age (years)	51	SACT[34]
Proportion female	0.40	SACT[34]
Weight (kg)	73.73	KEYNOTE-087 Cohort 2 ASaT population[33]
Body Surface Area (BSA)(m^2)	1.85	KEYNOTE-087 Cohort 2 ASaT population[33]

B.3.2.2 Model structure

Unlike the advanced solid tumour setting, where Partitioned Survival Models (PSMs) dominate, a variety of model structures of varying complexity have been used in haematology-oncology submissions to NICE. There have been four NICE Technology Appraisals in cHL, TA540[2] (the original appraisal for this topic), which used a landmark semi-Markov hybrid model, TA772 (pembrolizumab for treating R/R cHL after SCT or at least two previous therapies), which used a PSM and TA462[27] (nivolumab after BV and SCT), which used a Markov model incorporating multiple "special case" transitions within the overall umbrella of three health states and TA524[26] (BV in the 3L setting), which used a semi-Markov model with six (depicted in the model diagram) or eight (depicted in the Markov trace) health states that attempted to capture auto and allo SCT separately. A range of other model structures have been used in other haematology-oncology NICE submissions where the intervention was a bridge to SCT or where SCT was distal in the treatment pathway. In TA541[56] (an acute lymphoblastic leukaemia indication) included a 4-state Semi-Markov model whereby 'SCT and post SCT' were modelled as a separate health state. This included all patients who received SCT irrespective of response status. Tunnel states were incorporated to capture wait time to SCT. In TA893[57] and TA554[58] (both B-cell acute lymphoblastic leukaemia indications) and TA677[59] (R/R mantle cell lymphoma) a 3-state PSM was used. After reviewing these appraisals, we concluded that there is no one preferred or standard model structure in cHL or within the health economics of haematology-oncology more broadly. It is important to emphasise that TA772 is the only cHL appraisal that has been underpinned by a parallel RCT and that the treatment effects in all other models have been derived from a combination of indirect comparisons, linked-evidence surrogate outcomes approaches and clinical opinion, as will be the case here.

B.3.2.3 Critique of original TA540 model structure

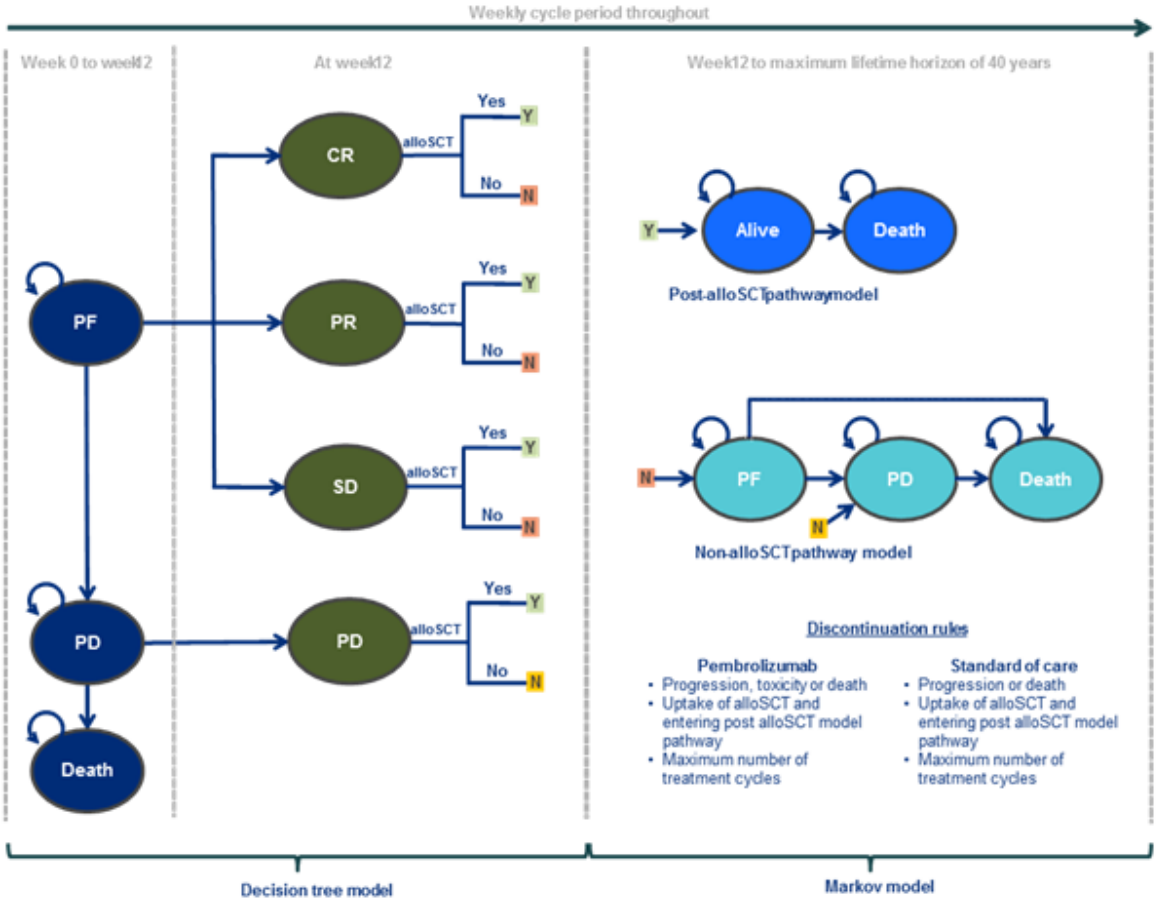
We reviewed the model MSD submitted to NICE for TA540 to see if it could simply be updated with the latest clinical data.

The company's original model is a semi-Markov hybrid landmark model. It can be effectively thought of as a multi-section model that consists of a PSM up to the

landmark, a decision tree that divides patients into those who receive SCT and those who don't and two separate Markov models that capture the relevant long term outcomes for the two groups.

For the initial PSM up to the landmark, PFS and OS hazard ratios were derived from indirect comparisons and applied to the SoC arm. At the landmark, a decision tree was applied to calculate the proportion of patients who underwent SCT. This decision tree was calculated by multiplying the probability of patients being in states of Complete Response, Partial Response, Stable Disease or Progressed Disease (treatment effect taken from a Matching-Adjusted Indirect Comparison using the KEYNOTE-087 data and comparator studies' patient characteristics) along with probability that a patient would get an alloSCT in each of these response categories, derived from a clinician survey (assumed to be consistent in both arms). Those who did not get an SCT proceeded down a pessimistic 3-state semi-Markov model and those who did proceeded down an optimistic 2-state Markov model. No treatment effect was applied to those who got an SCT but an indefinite PFS hazard ratio was applied to those who did not.

Figure 11. Original model from TA540



There are some important limitations with this model structure, some of which were highlighted in the EAG's critique but others that have become clearer as our understanding of how pembrolizumab is being used in cHL has evolved in the six years since this model was built.

The first important limitation of this model is structural. In the original appraisal, two landmarks were examined; 12 and 24 weeks, and the model's ICER was found to be very sensitive to this structural variation. This appears to be driven by the proportion of people who progress and die between the 12-week and 24-week landmarks; as the landmark is pushed out, there are far more of these in the SoC arm and therefore far fewer patients who become eligible for SCT. For example, in the base case 12-week model, 40% of patients on pembrolizumab and 30% of patients on SoC got an SCT but in the 24-week model these numbers had dropped to 34% and 14% respectively. The landmark represents the median time to SCT, and all SCTs occurring at the median is also a limitation; it is clear from the SACT data that, while

24 weeks might be a reasonable estimate of the median time to SCT (among patients who actually had one), events occur along a continuum from a few weeks to as much as two years, although the vast majority occur within nine months of treatment initiation. Another limitation of this structure is that patients may only have an SCT on response to therapy received prior to the landmark whereas it is clear from the SACT data that a number of patients receive pembrolizumab followed by another round of chemotherapy then SCT. Clinicians consulted at an MSD UK advisory board confirmed that after treatment with checkpoint inhibitors they would expect some patients to become chemosensitised and that treatment with a further line of chemotherapy is a viable treatment strategy to bridge to SCT, despite these patients having not responded to chemotherapy in earlier lines.

The second issue is the way that the proportion of patients receiving an SCT in both arms is estimated. The proportion of patients receiving an SCT is affected by both the landmark timepoint and the PFS and OS treatment effects and it is not clear that it should be. In the pembrolizumab arm, we now know from the SACT data what proportion receive SCT after treatment with pembrolizumab and do not need to rely on a linked-evidence approach. We also know that some of these patients receive their SCT on chemotherapy following pembrolizumab rather than as response to initial therapy. In the 12-week model, the proportion estimated to be receiving SCT on SoC via this method is much higher than is usual in clinical practice. These are patients who have failed multiple lines of chemotherapy and BV (itself a chemotherapy-based regimen) and not achieved enough of a response to progress to SCT. Clinicians consulted at the MSD UK advisory board stated that these patients have chemo-insensitivity and very few of them would get a good enough response to fourth line chemotherapy to receive an SCT. The base case 24-week model estimates more reasonable proportions in both arms although this is coincidental rather than data-driven and will be sensitive to changes in other parameters like OS and PFS treatment effects. The evidence on the proportion of patients getting an SCT in SoC is discussed in Section B.3.3.2.

The third potential issue is that the model structure explicitly assumes that PFS is a good surrogate for OS after the landmark (the PD-to-death transition probability is the same in both arms). It is not certain that this is true in R/R cHL. PFS in

KEYNOTE-087 is far shorter than OS, for example, demonstrating that patients in that study spend the majority of their within-trial life years in the implied PD health state (PFS is just 20% at 2.5 years whereas OS is close to 90%). That does not necessarily mean that there shouldn't be a constant transition probability between PFS and death but that the link is, at least, not immediate. Progression was also not associated with changes in utility in KEYNOTE-087. We concluded that a model structure where OS relies on PFS was not necessarily supported by the data.

Overall, we concluded that a model structure that could overcome these structural limitations was now possible, given the availability of real-world data on pembrolizumab's use within this setting in the NHS and greater familiarity in the clinical community allowing parameters where evidence is lacking to be estimated with greater confidence by clinicians.

B.3.2.4 The company's new model

B.3.2.4.1 General principles

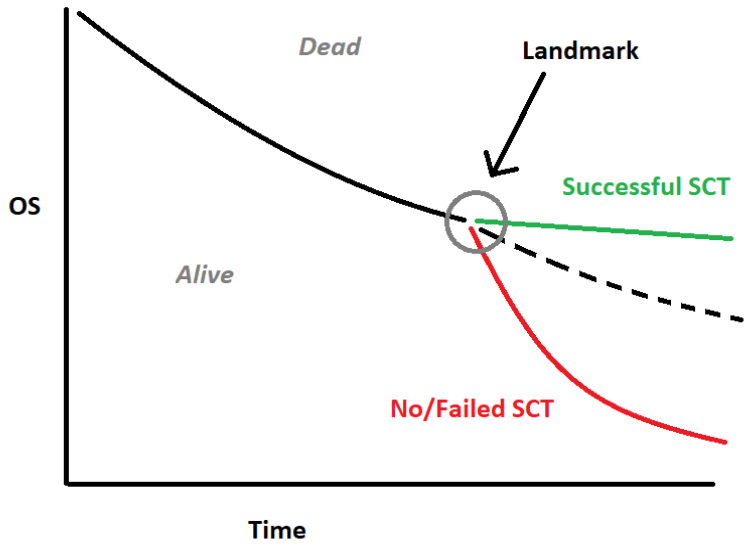
When considering what model structure to adopt, we were mindful that this is an evidence-light area. No parallel RCTs have ever been conducted in the 4L+ setting and KEYNOTE-204 is the only one to have been conducted in the 3L+ setting. No new clinical or evidence that is directly relevant to the decision problem has been published since the original appraisal meaning that no evidence directly applicable to the SoC is available. We were therefore wary of building a complex model that required too many assumptions.. As such, we sought to minimise the number of health states and transition probabilities, while still capturing the most important outcomes.

From reviewing the other NICE cHL appraisals, the SACT data and consulting NHS clinicians we concluded that there are certain important differences between pembrolizumab and standard care that should be captured in our model:-

- Adding pembrolizumab to the pathway is likely to increase overall survival.
- Adding pembrolizumab to the pathway is likely to increase the number of SCTs.

- Adding pembrolizumab to the pathway is likely to increase the probability that an SCT is curative because PD-1 inhibitors are known to increase chemosensitivity (SCT is a chemotherapy-based intervention) and because pembrolizumab is likely to elicit better and more durable remission than SoC.
- Patients who cannot have or do not want an SCT can continue to receive pembrolizumab and may benefit from the treatment for many years.
- Patients on pembrolizumab typically have better quality of life than patients on standard care.
- PFS may not be a reliable surrogate for OS or for SCT in R/R cHL.

Figure 12. New Economic Model Structure



**NB: Dotted line shows the weighted average OS from the two post-landmark groups*

Aligned with the general principles, we adopted the landmark model illustrated in Figure 12 inclusive of the following treatment effects in order of anticipated influence on (cost-)effectiveness:-

- OS HR up to the landmark
- Differential probability of SCT
- Differential probability that SCT is curative
- Differential HRQoL prior to the landmark
- OS HR post the landmark in the No/Failed SCT health state

- Differential HRQoL after the landmark for patients in the No/Failed SCT health state

This model is conceptually very similar to the one that was considered by the committee at CDF entry but with attempts to address the problems outlined above and make the best use of the evidence that is now available. The rationale for this structure is explained in the following section.

B.3.2.4.2 Omission of PFS

PFS is omitted from the model. The principal reasons for this were that we had no PFS data recorded in SACT and no reliable way to estimate PFS for the SACT cohort. Since SACT was likely to be the primary source of OS in the model, we felt this was important. On review of the pembrolizumab datasets, PFS appeared not to be a reliable surrogate for either OS or whether patients receive an SCT. This is apparent for OS because there is a large gap between the PFS and OS curves observed in KEYNOTE-087 (2.5 year PFS is ~25% in KN087, where OS is nearly 90%). For SCT, it was clear that PFS was not a reliable indicator because we observed in the SACT data that many patients could receive a potentially curative SCT after progression on pembrolizumab. Omitting this health state had the benefit that there were fewer uncertain treatment effects to estimate and fewer assumptions to make. It had the drawbacks that the use of subsequent treatments could not easily be tied to progression and that we might have lost some nuance in terms of the way progression affects quality of life, although progression was not associated with changes in utility in KEYNOTE-087. This does not obviously bias the model in one direction or another and any related under or overestimation of QALYs is likely to be small in comparison to treatment effects that influence OS or SCT, which are more fully captured.

B.3.2.4.3 Landmark and SCT considerations

It was clearly important to capture SCT in the model but OS post SCT is associated with a time-dependent survival curve (many patients die or relapse in the first two years but then a significant plateau of cured patients emerges), which is computationally complex to include in downstream health states, requiring the use of specialist modelling software or, if using MS Excel, macros that increase run-time

and opacity. Mindful of the ERG's comments on the opaqueness of the company's model for TA462,[27] which attempted to capture the complex shape of downstream transitions in this way, we were keen to adopt a simple transparent structure which reflected the main outcomes of interest.

We observed that all SCT events had taken place by two years in the SACT data and the vast majority by nine months.[34] On reviewing evidence on SCT outcomes in this population, we saw almost completely flat PFS and OS curves after two years post SCT. We therefore concluded that all SCT events of interest (SCTs, SCT-related deaths and relapses) would have taken place by four years. By this time, the cohort in the model could effectively be divided into patients who had been cured by an SCT, patients who had received an SCT but relapsed and patients who had never received an SCT. Like the previous structure, a decision tree was implemented at the landmark, albeit a more simple one. The probability of SCT was multiplied by the probability of SCT being curative to determine the proportion of patients in the cured-SCT state and the remaining alive patients were all assigned to the No/Failed SCT state. This structure had the advantage that the short term outcomes for patients who had an SCT and relapsed did not have to be modelled explicitly; they were captured within the whole group before the landmark and within the No/Failed group afterwards.

At the UK MSD advisory board we validated the assumption that all SCT events of interest would have taken place by four years with 6 UK clinicians currently treating advanced cHL in their clinics and they confirmed it to be reasonable.[5] The experts estimated probabilities of patients having an SCT and SCTs being curative as part of a structured expert elicitation exercise (See Section 3.3.4 and Appendix N).

Allowing the OS curves to continue to four years has the major advantage that the SCT related events, their treatment effects and short term outcomes do not have to be estimated explicitly but rather are implicit parts of the OS curves.

This model structure has the advantage that moving the landmark does not influence the proportion of patients who actually receive an SCT. We set up the model to examine alternative landmarks in sensitivity analysis.

B.3.2.4.4 Transitions beyond the landmark

Beyond the landmark there are effectively two health states; those who are cured of their cHL and those who are in the No/Failed SCT health state.

The cured SCT group are homogenous within the model and are assigned general population quality-of-life using the study by Ara and Brazier (2010)[60] in line with a previous R/R cHL appraisal (TA772)[1] and general population mortality rates in both arms in the base case. We make no distinction in quality of life between patients who have been cured by allo or auto SCT or by arm.

The No/Failed SCT group are also assumed to be relatively homogenous in that they have received all treatments of interest and are anticipated to have relatively short overall survival on SoC although a proportion of patients in each group will have received a failed SCT. Assigning accurate transition probabilities for this group was more difficult so we undertook several sensitivity analyses.

Table 30. Features of the economic analysis

	Previous evaluations				Current evaluation	
Factor	**TA462**	**TA524**	**TA540**	**TA772**	**Chosen values**	**Justification**
Time horizon	40 years	70 years	40 years	40 years	40 years	The simulated cohort evaluated in the economic model is set to 40 years as this approximates a lifetime in this patient population.
Cycle length	Monthly	Weekly	Weekly	Weekly	Weekly	Weekly cycles were chosen in the model as this was considered more likely to accurately capture costs and treatment administration schedules than monthly cycles. The cycle length is consistent with the previous pembrolizumab submission, TA540[2] in addition to TA524,[26] and TA772[1] and is considered to be sufficiently short to allow an accurate estimation of the event timings while not adding unnecessary complexity of daily cycles.

Discount rate for utilities and costs	3.5%	Not reported	3.5%	3.5%	3.5%	Consistent with the NICE reference case.
Perspective	NHS and Personal Social Services perspective	NHS	NHS and Personal Social Services perspective	NHS and Personal Social Services perspective	NHS and Personal Social Services perspective	Consistent with the NICE reference case.
Treatment waning effect	Not reported	Not reported	Not reported	Not applicable as OS conservatively assumed equal. Not imposed on PFS.	Not applied in base case. 3-5 years post pembrolizumab cessation for mortality treatment effect in sensitivity analysis.	No treatment waning assumption imposed in TA540 and not on PFS for TA772. 3-5 years post cessation is in line with a NICE committee's most recent stated preferences (TA885).

Source of utilities	Nivolumab arm derived from EQ-5D-5L questionnaire from CheckMate 205 and converted to the EQ-5D-3L tariff. SoC arm was derived using the Swinburn et al. (2015[61]) and reweighted through response rates from the Cheah et al. (2016)[4] study.	Utilities were sourced from published sources including BV clinical studies (Swinburn 2015[61]), and a published study of utility post ASCT (van Agthoven 2001[62])	Derived from EQ-5D-3L questionnaire from the KEYNOTE-087 trial.[2]	Utilities were sourced from KEYNOTE-204 trial (based on EQ-5D-3L questionnaires collected during the trial)	Derived from EQ-5D-3L questionnaire from the KEYNOTE-204 trial.	Consistent with the NICE reference case. Although one line of treatment earlier, KEYNOTE-204 utilities were prioritised over KEYNOTE-087 due to the greater sample size and the information vs a comparator treatment. KEYNOTE-087 values were used in sensitivity analysis.

Source of costs	PSSRU NHS Reference costs MIMS BNF	Clinical expert opinion advised the medical and administration costs	PSSRU, NHS Reference costs eMIT BNF	Drug costs were sourced from BNF and eMit. Drug administration costs and AE costs from NHS Reference costs 18/19 Disease management costs and terminal care costs from previous TAs based on PSSRU and NHS reference costs	PSSRU,[63] NHS Reference costs (2021/22)[64] eMIT[65] BNF[66]	Consistent with the NICE reference case.

B.3.2.5 Intervention technology and comparators

The intervention (i.e. pembrolizumab) was applied in the model as per the licensed dosing regimen (i.e. administered intravenously at a fixed dose of 200mg over 30 minutes every 3 weeks [Q3W]) to a maximum of 35 treatment cycles. Pembrolizumab is also able to be given every 6 weeks and we examined this in scenario analysis.

The final scope specifies the following treatments as relevant comparators:[3]

- Single or combination chemotherapy including drugs such as gemcitabine, vinblastine and cisplatin
- Best supportive care

Data on patients treated with pembrolizumab was be drawn from the SACT dataset collected as part of the CDF arrangement and from the latest available data from the KEYNOTE-087 study.

Which interventions make up the standard of care in this population and what study data could be used to represent the relevant outcomes is far less certain.

MSD note the comparators in the Final Scope have not changed from the original submission TA540. As highlighted in TA540 and relevant to the current CDF exit, the clinical pathway for R/R cHL patients is still subject to considerable heterogeneity and uncertainty given the low number of patients in 4L, who typically have a relatively short survival.[2] There is also uncertainty for patients in the downstream setting, where no SoC exists and use of investigational therapies outside of the NICE scope is common.

Given the considerable variation in which treatments patients with R/R cHL receive in SoC, the same approach was adopted as in TA540; identifying studies in the literature to inform either single or combination chemotherapies as well as asking clinicians at a UK advisory board. As stated in section B.2.1, no further studies were identified in the clinical SLR update. In the TA540 submission, the comparator composition was solely based on the Cheah et al. (2016) study. In this study, 100% (n=100) of patients had received prior BV and 71% (n=71) had prior autoSCT, which meant the study better represented the post-autoSCT, post-BV HL population,

"population 1", from the original appraisal/the KEYNOTE-087 trial. Following the FAD of TA540, the NICE Appraisal Committee recommended pembrolizumab only for patients who received BV and cannot have autoSCT. This means the population in Cheah (i.e. treated with BV and received prior autoSCT) does not match the population for as per the NICE final scope for ID5084.[3] Importantly, the fact that so many of these patients have received an SCT suggests a greater level of chemo-sensitivity and a greater level of fitness in their clinical histories than would be expected in the cohort for this appraisal, who have all either never achieved a good enough response to chemotherapy or never been fit enough for an SCT. The Cheah patients were also over 20 years younger on average than the SACT cohort, with a median age of 32 years vs. a mean age of 51 years. The clinical advisors at the UK advisory board agreed that this study would overestimate outcomes versus the KEYNOTE-087 trial population. This suggests that, a-priori, we should expect this study's outcomes to be better than our SoC cohort, all of whom have never been able to have an SCT.

In the original appraisal, the NICE committee welcomed the inclusion of the retrospective analysis by Eyre et al. (2017) as an alternative surrogate for outcomes on SoC. This study is also associated with a number of indirectness problems. Of note, this is a study of outcomes on BV, a more effective treatment option than the SoC in this appraisal, and 100% of these patients were considered clinically fit for transplant (vs. 61% in the SACT cohort). Like Cheah, the Eyre patients were also over 20 years younger on average than the SACT cohort, with a median age of 32 years vs. mean 51 years. The clinical advisors agreed that this study would overrepresent outcomes versus KN-087 because the patient population is younger, has a better performance status and is receiving an effective active treatment in BV. Eyre et al. (2017)[67] is from a UK real world setting where 100% of patients received BV and 15 (15%) and 19 (19%) of patients went on to receive autoSCT and alloSCT respectively. The Eyre study included a subgroup of patients who failed to reach SCT (n=38, 39%) which might be a closer match to the population of interest for this appraisal. We considered that, caveats aside, data from this study might be informative for either treatment options or outcomes.[67]

When attempting to elicit the regimens that made up SoC, the Cheah comparator composition from TA540 was amended to include some regimens from Eyre et al. (2017). In Eyre, treatment details were collected using hospital records from 9 UK centres by the treating physician. The combined range of regimens were subsequently validated and amended by the MSD Clinical Advisory Board to reflect current UK clinical practice. During the Advisory Board, clinicians were unable to confidently estimate proportions for SoC regimens in 4L, given there is no standard for cHL in this treatment line and pembrolizumab has been available in the CDF for a number of years. The experts mentioned several reasons for this, including heterogeneity of chemotherapy options available in centres, rarity of the condition and patient factors (e.g. fitness levels, age, patient preference).[5] For this reason, we made the assumption that proportions were equal across all 4L treatments and decided to vary this assumption in sensitivity analysis. The SoC composition is summarised in Table 31, inclusive of the following assumptions and amendments:

- Investigational agents were removed as these treatments are not typically relevant in NICE submissions.
- The 'other' group of treatments from Cheah does not provide enough detailed information to allocate costs. Therefore, this has been excluded from the SoC composition.
- Second autoSCT is not considered to be a relevant comparator in this patient population as assumed in the TA540 submission and in the scope for this appraisal (based on the July 2017 Advisory Board have stated patients with R/R cHL would rarely receive this). Therefore, this has been excluded from the SoC composition.[68]
- BV retreatment after its NICE recommended place in the care pathway is not explicitly recommended by NICE. Therefore, composition of SoC has been re-weighted excluding this therapy.
- Gemcitabine-based chemotherapies and 'platinum based' regimens were split based on the uptake from the Advisory Board (previously in TA540, these were pooled regimens with 'Other alkylators'). "Other alkylators" were removed from treatments identified by clinicians at the Advisory Board, as this was assumed to be likely to refer to chlorambucil, given as part of the DECC regimen.
- 'Platinum based' therapies was replaced by ifosfamide, carboplatin, etoposide (ICE) and mitozantrone, cyclophosphamide, prednisolone, mitoxantrone,

cyclophosphamide, etoposide, bleomycin and oncovin (PMitCEBO) based on the Advisory Board validation.

- Oral chemotherapy i.e. dexamethasone, etoposide, chlorambucil, lomustine (DECC) and, Mini-BEAM and radiotherapy was included based on the Eyre et al. (2017) study.

Table 31. SoC treatment composition

Treatments	%	Source
Bendamustine	14.29	Cheah et al. (2016)[4] and Eyre et al. (2017)[67] and MSD Advisory Board[5]
ICE (ifosfamide, carboplatin, etoposide)	14.29	
Weekly chemotherapy (PMitCEBO)	14.29	
Gemcitabine-based (IGEV, GEM-P, GDP, GVD)*	14.29*	
Oral chemotherapy (DECC; dexamethasone, etoposide, chlorambucil, lomustine)	14.29	
Radiotherapy	14.29	
Mini-BEAM	14.29	
Abbreviations: BV, brentuximab vedotin; DECC, dexamethasone, etoposide, chlorambucil, lomustine; ICE, ifosfamide; carboplatin, etoposide; Mini-BEAM, carmustine, etoposide, cytarabine, melphalan; PMitCEBO, prednisolone, mitoxantrone, cyclophosphamide, etoposide, bleomycin and oncovin. **This weight has been divided evenly between the four gemcitabine-based regimens*		

It was assumed the likelihood of patients receiving BSC (i.e. no active treatment) was minimal as, based on TA540 and feedback at the UK Clinical Advisory Board, patients will likely receive therapy where feasible. The consensus among the clinicians from the Advisory Board was that pembrolizumab via the CDF is the current standard of care for these patients.[5] The clinical experts advised if pembrolizumab did not exist in the clinical pathway, patients would receive the treatments according to Table 31.[5] As such, although BSC was listed as a separate comparator in the scope, we have not included it as such in the economic analysis, instead it has been applied as a subsequent therapy, as in TA540.

B.3.3. Clinical parameters and variables

As discussed in Section B.2.2, there is both clinical trial data and real-world evidence available to inform the model effectiveness parameters for the outcomes on pembrolizumab for patients in R/R cHL The main clinical trial of interest was KEYNOTE-087, a single arm trial of pembrolizumab in patients with R/R cHL using cohort 2 for those patients who have previously received BV and are ineligible for autoSCT. Alternatively, clinical outcomes data was collected for patients receiving pembrolizumab in this population through SACT. The SACT data provides the largest source of real-world data for this indication and is therefore considered the best source of evidence to reflect the outcomes of patients on pembrolizumab in UK clinical practice. The SACT data were therefore the preferred source of clinical data for the pembrolizumab arm of the model and were used as the base case inputs for several parameters, as presented in Table 41.

As KEYNOTE-087 was a single-arm, there was no randomised clinical trial data available that provided a direct comparison of pembrolizumab versus standard of care in this indication. Additionally, no further studies were identified in the clinical SLR that provide outcomes evidence on SoC specific to the population of interest in this appraisal. Therefore, a series of indirect treatment comparisons and a structured expert elicitation were required to inform the clinical outcomes of patients in the SoC arm of the economic model.

The economic model reflects the differences between the pembrolizumab and SoC pathways using a number of explicit or implicit treatment effects:-

- Overall survival hazard ratio up to the landmark
- Probability that a patient gets SCT
- Probability that SCT is curative
- Overall survival hazard ratio after the landmark for patients in the NO/Failed SCT health state
- Differential HRQoL before the landmark
- Differential HRQoL after the landmark in the No/Failed SCT group
- Differential adverse event rates

In this section we detail the calculations and options for all of these effects. The sources for these base case clinical parameters for both treatments are summarised in Table 41.

B.3.3.1 Overall survival

B.3.3.1.1 Overall survival on pembrolizumab up to landmark

OS data for pembrolizumab up to the four-year landmark is available from both the SACT dataset and KEYNOTE-087 and no survival extrapolations are needed for this parameter. For the purposes of the economic model, we fit standard one-piece parametric survival curves to the SACT and KEYNOTE-087 data. This would enable cycle by cycle transitions to better represent a theoretical cohort rather than following the KM curve would and, crucially, would allow the landmark time to be extended outwards in sensitivity analysis. We noted that there was little difference between any of the options in terms of visual or statistical fit and selected the log-logistic model as a central estimate and because it had the property of declining hazards, which is logical given the presence of curative SCTs and is typical in pembrolizumab trials because of the highly variable outcomes between responders and non-responders to immunotherapy.

Figure 13. SACT OS and parametric fits

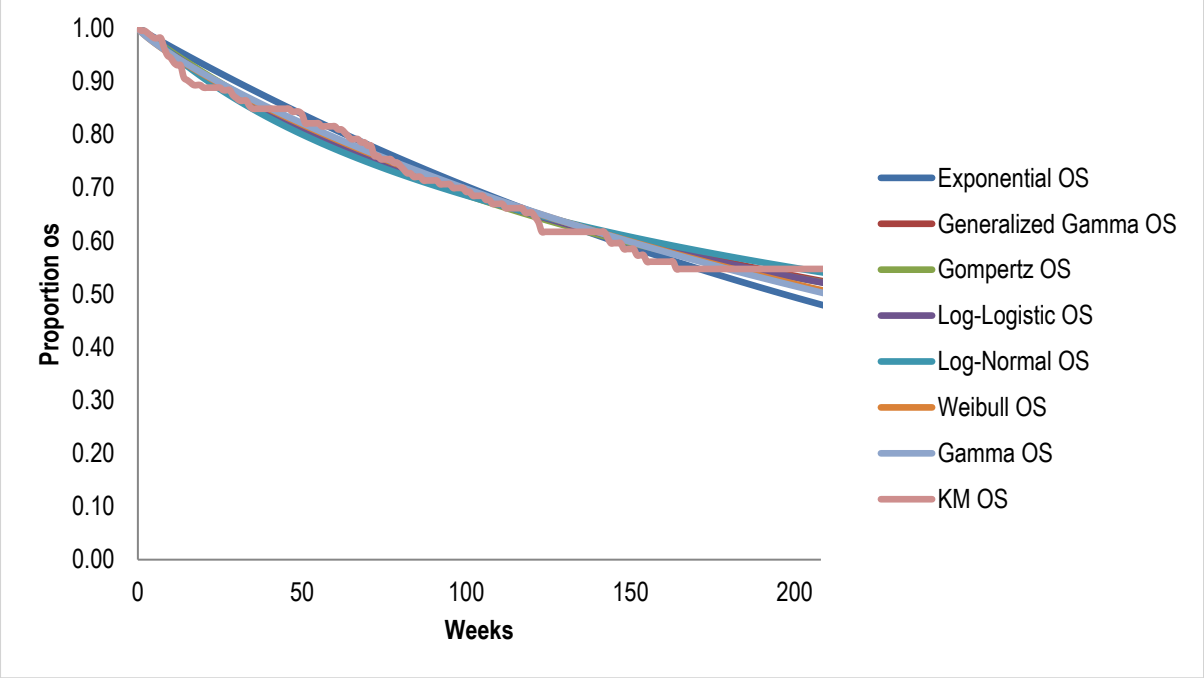


Table 32. Model fit statistics for parametric curves fit to SACT OS

Functional Form	AIC	BIC	Average of AIC and BIC
Exponential	972.398	975.769	974.084
Weibull	972.039	978.780	975.409
Log-logistic	971.615	978.357	974.986
Log-normal	972.077	978.818	975.448
Gompertz	971.727	978.469	975.098
Generalized Gamma	973.198	983.310	978.254
Gamma	972.278	979.020	975.649
Abbreviations: AIC, Akaike Information Criterion; BIC, Bayesian Information Criterion; OS, overall survival; SACT, Systematic Anti-Cancer Therapy			

Given that the SACT dataset is much larger (N=215 vs N=81) and represents how pembrolizumab is used in the UK NHS (i.e. as a bridge to transplant where possible), we selected this dataset in the base case analysis. We note that outcomes in the Real World Data are worse than in the KEYNOTE-087 trial, likely due to higher age,

worse patient fitness and presence of comorbidities. OS data from KEYNOTE-087 was used in the model in sensitivity analysis.

B.3.3.1.2 Overall survival hazard ratio

Overall survival on the SoC up to the 4-year landmark is calculated by applying a constant hazard ratio to the pembrolizumab arm. Given that data have not been published on outcomes on SoC in this indication, we had to rely on a series of imperfect data sources upon which to conduct Indirect Treatment Comparisons (ITCs). No new evidence was identified during the SLR update so we relied on the two studies that were considered during TA540; Eyre (2017), which was a study of BV in exclusively transplant fit patients, and Cheah (2016) a study of standard care (including 30% on clinical trials) largely in patients who had relapsed after SCT and thus had proven chemo-sensitivity within their clinical history. Neither dataset was considered highly representative of SoC in this appraisal and both datasets were considered to have positive bias when compared to expected outcomes on SoC. Sample sizes were relatively small for both studies (N=99 and N=100) and the median ages of patients were ~20 years younger than in the SACT dataset. Within both studies, OS was reported among patients who did not get an SCT. These might be the closest data to our SoC population as these cohorts would be comprised of patients who had failed on BV and did not get a transplant and our SoC arm are patients who have failed BV and >90% will not get a transplant. However, as we did not have baseline characteristics for these patient subgroups, we were not able to carry out a complete assessment on direction of bias vs the pembrolizumab studies and therefore unable to formally conduct an ITC. We note that visual inspection of the difference in outcomes between these groups and the pembrolizumab studies indicates the OS HR would be very large (larger than the OS HRs we have used in the economic model), were an ITC conducted.

Because we lacked access to the individual patient data, ITCs that compared to the SACT data were unadjusted, including only a narrative description of the likely direction of bias. ITCs vs KEYNOTE-087 were able to be adjusted for differences in baseline characteristics i.e. they were Match-Adjusted Indirect Comparisons (MAICs). Details of the ITCs for the pembrolizumab data vs the Cheah and Eyre studies are available in Appendix D.

There is one source of relevant randomised evidence that the company has access to; the as-yet unreported OS outcomes from the KEYNOTE-204 study. This was a study of pembrolizumab vs BV in the 3L setting and included a stratified subgroup who had not had an SCT. This trial was the basis for NICE TA772 although no OS data was available at that point. We detail the outcomes in Confidential Appendix P. Because BV has established clinical effectiveness vs SoC, it is expected that using this estimate in the model is conservative.

Appendix P contains model diagnostics for the ITC (log-cumulative hazard plots, visual assessment of KMs and Schoenfeld residuals). On reviewing these we concluded that there are no major violations of the proportional hazards assumption and that the use of a constant hazard ratio up to the landmark is reasonable, particularly as it is only estimating outcomes within the first four years of the model.

One further source of comparative evidence was identified through hand searching; the Markov trace from the BV vs SoC model used in NICE TA524.[26] This was used to obtain an indicative overall survival hazard ratio for this comparison, for use in Bucher ITCs[69] where BV was the common comparator i.e. where the treatment effect of pembrolizumab vs SoC can be assumed to be the sum of the treatment effects for pembrolizumab vs BV and BV vs SoC.

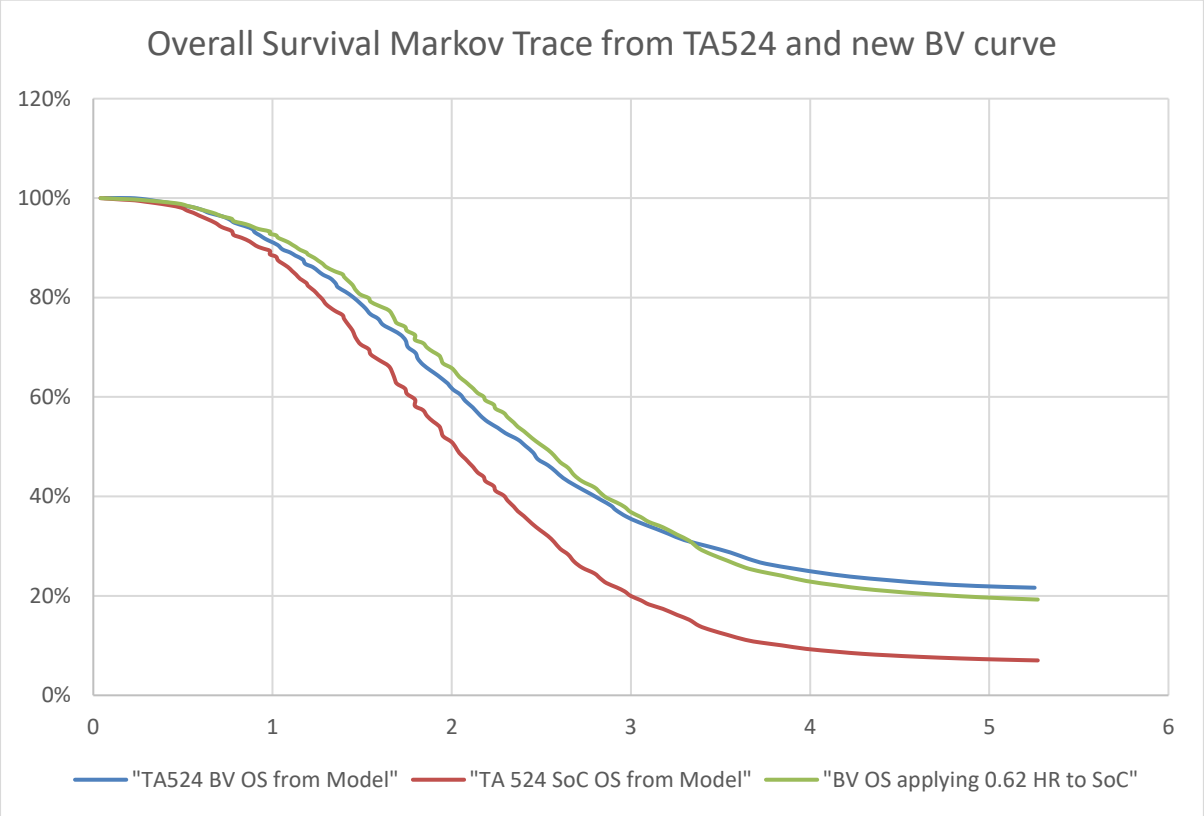
B.3.3.1.3 Treatment effect of BV vs SoC

No anchored evidence on the treatment effect of BV vs SoC was identified in the SLR or during the NICE Technology Appraisal process that led to the approval of BV (TA524).[26] However, some comparative evidence does exist within that appraisal, specifically relating to “population 3” (those who have not had an SCT; the population that was re-evaluated after a period in the CDF). A schematic detailing the health state membership over time in the company’s economic model complete with changes request by the EAG is available (page 132, committee papers[26]). This economic model represents the combination of the best available evidence on the effectiveness of BV vs SoC identified at the time and an implied treatment effect that was accepted by the NICE committee as plausible.

In order to calculate the implied and accepted OS treatment effect from this model, we used WebPlotDigitizer[70] to obtain the proportion alive in both model arms during

the first four years. A single constant hazard ratio was applied to the SoC curve to calculate a new BV curve. The hazard ratio was varied until the new curve fit the original BV trace as closely as possible based on visual assessment. Looking at the shape of the new curve, we concluded that a constant hazard ratio reasonably characterises the implied treatment effect on OS i.e. that the proportional hazards assumption is not obviously violated within the first four years of model time. We concluded the BV curve was best approximated using an overall survival hazard ratio of 0.62. We validated this estimate as plausible at the UK Clinical Advisory board.

Figure 14. Overall Survival Markov Trace from TA524 and new BV curve



This approach has the advantage that, in an evidence light area, it is based on the deliberations and conclusions of the NICE committee but important disadvantages as well. The population that this appraisal related to was earlier in the cHL treatment pathway than in the current appraisal and the model itself is not based on randomised evidence.

Without an appropriate method to calculate a standard error for (the log of) our point estimate, we assigned an arbitrarily large value of 0.244 (so that the upper confidence limit was equal to 1).

We used the values we obtained in this exercise along with the various estimates of treatment effect for pembrolizumab vs BV to conduct simple Bucher ITCs[69] to calculate the implied 4-year OS HR for pembrolizumab vs SoC. For the Bucher ITCs, the hazard ratio for pembro vs SoC was calculated by multiplying the hazard ratio for pembro vs BV and the hazard ratio for BV vs SoC together. The standard error of the log hazard ratio was obtained by taking the square root of the sum of the two variances. The upper and lower confidence limits on the natural scale were the exponentials of the limits on the log scale.

B.3.3.1.4 Summary of Overall Survival HR estimates

Table 33 shows the results of the various options from the ITCs. In all cases apart from Estimate 1 the direction of bias favours the SoC.

Table 33. Summary of Overall Survival HR estimates

Estimate Number	Comparison	HR (CI)	Key Limitations
1	Bucher ITC (KN204 and TA524)	[redacted]	Two 3L studies, assumed s.e. from TA524.
2	KN204 OS	[redacted]	3L study, control arm is BV.
3	Bucher ITC (SACT vs Eyre and TA524)	0.41 (0.22 - 0.77)	100% patients fit for transplant in Eyre study, assumed s.e. from TA524.
4	ITC SACT vs Eyre	0.66 (0.44 - 0.98)	Eyre is 3L BV study, 100% fit for transplant.
5	ITC SACT vs Cheah	0.59 (0.4 - 0.86)	71% had prior transplant, 30% received investigational agents.

6	MAIC KN087 vs Eyre	0.23 (0.12 - 0.42)	4 + KN087 applicability concerns. Comparator BV.
7	MAIC KN087 vs Cheah	0.24 (0.14 - 0.4)	5 + KN087 UK applicability concerns.
* Note: these are parameterized as reciprocal HRs in the economic model because OS data must be anchored to the pembrolizumab arm.			

We selected Estimate 1 as the base case analysis. The reason for this was that it made use of the only relevant source of randomised evidence (KEYNOTE-204) along with a comparison that had already been validated by a NICE committee based on the totality of the evidence (NICE TA524). We felt that, since both these sources of evidence are anchored in some way and not obviously biased in favour of one comparator. Estimate 1 using the Bucher ITC is therefore potentially more reliable than the various unanchored ITCs we conducted. The key limitation is that it relates to third line rather than fourth line patients.

B.3.3.1.5 Survival after being cured by SCT

Survival for the cured SCT group was assumed to be equal to the general population using the 2019-2020 national life tables for England (Office for National Statistics).[71] We discussed the plausibility of this assumption with clinicians at the UK advisory board. The clinicians felt that it is plausible that average OS might be lower than the general population in a population who had previously had cHL and gone through an SCT but were not able to estimate by how much. They were not able to point us to any studies that would allow us to estimate the magnitude of any mortality decrement. As such, we examined a series of arbitrary Standardised Mortality Ratio adjustments in sensitivity analysis e.g. multiplying all cycle-specific OS event rates by 1.5 to examine the effect on the ICER.

B.3.3.1.6 Survival after the landmark for the No/Failed SCT group

No directly applicable data were available to help estimate the longer-term survival of patients who either never received an SCT or had failed one. The closest data available were KM curves on patients who had never received an SCT from the SACT data. Whether the addition of patients who had relapsed after SCT into this

group would mean the SACT estimates were optimistically or pessimistically biased was difficult to know. On one hand, the failed SCT patients could have worse average outcomes than those who had remained on pembrolizumab because they had undergone a difficult intervention that offered no benefit and stopped an effective one early. On the other hand, these patients were likely more clinically fit for SCT in the first place and therefore might have a better baseline expectation in comparison to patients for whom SCTs are not suitable. We fit standard one-piece parametric models to this dataset as shown in Figure 15 and selected the extrapolation with best fit based on AIC, visual fit and that conform to clinical expectations on pembrolizumab as summarised in Table 34. We applied the cycle-specific transition probabilities from year 4 to the patients in this health state.

Figure 15. OS extrapolations one-piece parametric model for after landmark no/failed SCT group (SACT[34])

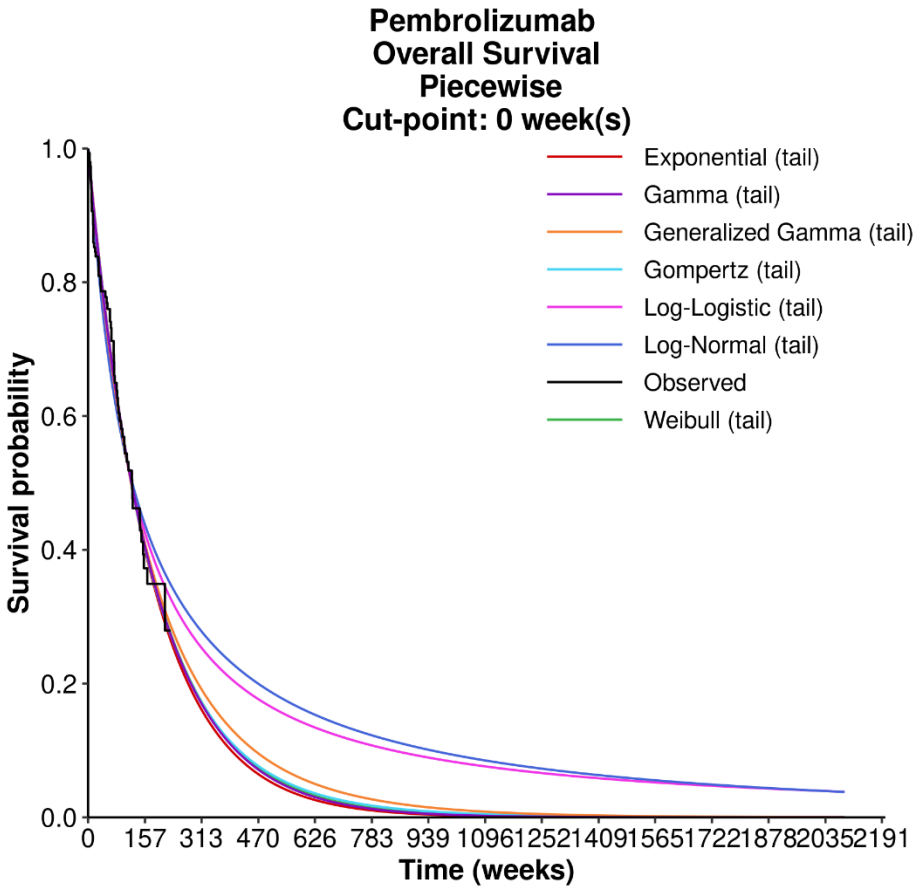


Table 34. Summary of goodness of fit for cohort after the landmark for no/failed SCT (SACT)

Distribution	Parameter A	Parameter B	Parameter C	AIC	BIC
Exponential	-5.14			825.34	828.35
Weibull	5.16	-0.04		827.22	833.24
Gompertz	0	-5.12		827.32	833.34
Log-Logistic	4.79	0.12		828.88	834.9
Log-Normal	4.8	0.47		829.64	835.66
Generalized Gamma	5.11	0.13	0.82	829.2	838.23
Gamma	-0.04	-5.21		827.23	833.25

Source: SACT report.[34]

We note that all curves provide comparable visual fit to the majority of the KM curve. The log-normal and log-logistic are above the KM curve at the tail and provide more optimistic extrapolations. AIC/BIC differences are not large. In the absence of clinical context, our default pick would likely be the exponential curve (similar visual fit to other curves, lowest AIC). The Weibull, gompertz, gamma and generalised gamma curves exhibit a similar pattern of constant or near-constant hazard to that of the exponential curve. This is likely unjustified among patients treated with pembrolizumab where the pattern of decreasing hazards, driven by the cohort becoming increasingly comprised of patients who have achieved durable response has been consistently observed across the clinical trial portfolio. We felt that the most conservative curve that did not exhibit constant or near-constant hazards, the log-logistic, should be the base case, with the exponential curve used in sensitivity analysis.

To calculate the corresponding OS hazard rates for the SoC we applied a hazard ratio to the above SACT-derived transition probabilities. Clinicians at the UK advisory board confirmed that there would be a treatment effect for some years after stopping pembrolizumab but were unable to say how long this would last. We implemented a treatment waning assumption in sensitivity analysis, where hazards in the pembrolizumab arm were gradually waned to become equal to the calculated

hazards in the SoC arm over time, equivalent to 3–5 years post cessation of pembrolizumab, in line with the latest stated preferences of NICE committees (e.g. NICE TA885[72]).

To obtain the relevant hazard ratio, we examined the OS data from KEYNOTE-204 on patients who had never had an SCT within the study. A priori, the use of this data is biased in favour of SoC because BV is likely more effective than SoC but the group in the model also contains a number of patients who have relapsed after SCT and rendering the direction of bias less certain. Overall, we concluded that though there is likely some benefit for pembrolizumab in this cohort, the magnitude and persistence of the treatment effect is somewhat uncertain.

The relevant OS HRs from KEYNOTE-204 are available in confidential Appendix P.

B.3.3.2 Probability of receiving SCT

For the pembrolizumab arm, the number and proportions of patients that receive either an autoSCT or an alloSCT is available from both the KEYNOTE-087 trial data and the SACT data, as presented in Table 35. The probability of patients receiving an SCT in the base case was informed using the SACT data. This was considered to be the best source of evidence as it is the largest dataset of-real world outcomes from NHS clinical practice and would better reflect real-world practice in the UK. A noticeable difference in evidence between KEYNOTE-087 and SACT is the ratio of patients receiving auto vs alloSCT. This difference could be due to setting-specific factors such as pembrolizumab being used specifically as a bridge to SCT among fit patients in SACT but not in KEYNOTE-087. This input was further validated by clinicians to be representative of clinical practice in the UK.

Table 35. Probability of receiving SCT

	SACT, n (%)	**KEYNOTE-087, n (%)**
Total SCTs	65/215 (30.2%)	24/81 (29.6%)
Autologous SCT (of those who received an SCT)	23 (35.4%)	15 (60%)
Allogeneic SCT (of those who received an SCT)	42 (64.6%)	10* (40%)

Abbreviations: SACT: Systemic Anti-Cancer Therapy database; SCT: Stem cell transplant. *One patient received both auto and allo SCT

In the SoC, no evidence was identified for the probability of receiving either an auto or alloSCT for patients with R/R cHL on SoC in the fourth line, and therefore this input was assumed to be the same 60/40 split as in the pembrolizumab arm.

B.3.3.3 Structured expert elicitation

B.3.3.3.1 Background

Several key parameters for estimating a treatment effect for pembrolizumab versus standard of care were not identified in the systematic review. Therefore, a Structured Expert Elicitation (SEE) exercise was conducted to obtain robust estimates that could inform these values. The parameters estimated were:

1. Proportion of people on standard of care alive after 4 years
2. Proportion of people on standard of care having a SCT
3. Proportion of people on standard of care who have an SCT that are cured
4. Proportion of people on checkpoint inhibitors who have an SCT that are cured

B.3.3.3.2 Methods – application of MRC protocol

The NICE methods guide recommends that SEE is carried out in alignment with the MRC reference protocol.[73] The methods used are outlined in Table 36. Summary of methodology (aligned to MRC protocol). Full details of the methodology, including the pre-specified protocol, training materials and evidence pack are available in Appendix N.

Table 36. Summary of methodology (aligned to MRC protocol)

Element	Methods used
Experts	• 6 experts were be included (note: 1 expert was unable to provide responses due to IT issues and participated in the group discussion only) • Experts were recruited for an advisory board and reflect a broad range of UK experience. Willingness to participate in the SEE activity was confirmed prior to the advisory board meeting • Experts' involvement in any previous MSD activities was recorded

Quantities elicited	• All quantities elicited were simple quantities (e.g. proportion of patients who have a stem cell transplant) • Survival is dependent on the proportion of patients who have stem cell transplants and the percentage of patients for whom this is curative. For this exercise survival was treated as an independent variable. This is in line with the guidance in the MRC protocol which states "only ask about independent variables, express dependent variables in terms of independent variables or use dependence elicitation methods"
Approach to elicitation	• Beliefs were elicited individually via completion of a chips and bins exercise using a Microsoft Excel template • Pooled results were presented to the experts at an advisory board. The experts had the opportunity to discuss the pooled results • Between-expert variation was explored explicitly during the consensus session
Method	• The Fixed Interval Method (FIM) was used (Chips and Bins method)
Aggregation	• The output from the Excel template was collated using the STEER R-Shiny app[74] and aggregated using the app. The app utilised the SHELF fitdist[75] function, which fitted distributions to the experts' individual judgements and to perform linear pooling with equal weighting of experts • As the elicited quantities are all proportions, the 'beta' distribution was explored as a first choice as it is naturally bound by 0 and 1, with alternative distributions explored as scenarios if the 'beta' distribution was determined to be a poor fit • The validity of any adjustments to the pooled distribution was explored both internally and with the experts at the advisory board.
Delivery	• A training slide deck (adapted from the STEER example training deck[76]) was provided to the experts. Experts were advised to contact the MSD team should they have any questions related to the training material. • An evidence briefing (Word format) was provided to the experts. This briefing contained evidence that had come to light during the SLR and was supplemented via hand searching where required. Details of the approach to hand searching were reported. • The elicitation exercise was delivered remotely. A training question was included so the experts can familiarise themselves with the Excel template. Experts were advised to contact the MSD team should they have any questions related to the exercise or technical issues with completing the Excel template. • Experts had the opportunity to discuss the validity of the pooled results at a virtual advisory board meeting. The advisory board facilitators ensured that all experts have opportunity to comment on the results.
Training and piloting	• A training slide deck (adapted from the STEER example training deck) was provided to the experts. A training question was included so that the experts can familiarise themselves with the Excel template. Experts were advised to contact the MSD team should they have any questions related to the training material.

	• The clinical validity and clarity of questions was assessed by one of the clinical experts prior to the meeting and the questions refined for clarity based on this feedback
Rationales and documentation	• Rationales for how the experts made their judgements were collected via a free text box in the Excel template were summarised and discussed at the advisory board • Conduct of the SEE followed the pre-specified protocol – any deviations were recorded and rationales for these provided

B.3.3.3.3 Results

The preferred distributions selected by the experts are presented below for each quantity of interest. One expert was unable to submit a response to the SEE questions but contributed to group discussions. Full details of the experts' rationales for their preferences and exploration of alternative distributions are presented in Appendix N. Unless otherwise stated the distributions include responses from all five experts who provided a response to the exercise.

Proportion of patients on standard of care who receive stem cell transplants

The elicitation question was:

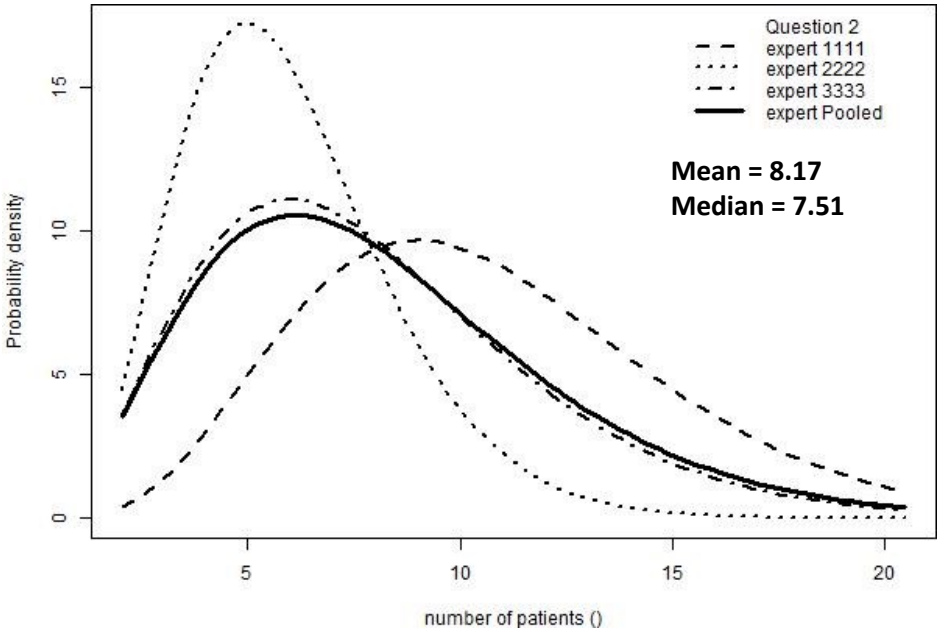
"Imagine a representative cohort of 100 adult patients in the UK (of all ages and fitness) with relapsed or refractory classical Hodgkin lymphoma who have relapsed or not responded to two lines of chemotherapy and brentuximab vedotin and cannot have autologous stem cell transplant. These patients are treated with current standard of care (chemotherapy or best supportive care). How many of the 100 patients would receive an SCT after treatment with standard care?"

The experts' responses fell into two clear groups. To aid discussion of inter expert variability results were presented separately for these groups using a beta distribution. The mean values for stem cell transplant rate in the two groups were 8.17% and 31.47%. After discussion the experts reflected that this difference arose from different interpretations of whether the question referred to theoretically "transplant fit" patients or not. For example, one advisor whose response had been higher indicated that they were considering younger patients with better performance status, and that they would expect approximately 30% of these patients to respond well enough to salvage chemotherapy to consider SCT. This expert indicated that if

they were considering all patients regardless of age/performance status (the population of interest) then their estimate would drop to 10 to 15%.

The experts' preferred distribution is shown in Figure 16. This distribution includes responses from 3 of 5 experts, as the expert group concluded that this group of responses most closely aligned with an interpretation of the question that matched the patient population of interest. All experts agreed that a rate of 5 to 15% for the overall proportion of patients who would receive a stem cell transplant was reasonable. One expert noted that the mean value of 8.17% was in reasonable alignment with a value of 5%, which had been accepted in TA542[26] (for the third line population). Results including responses from all experts using log normal and beta distributions are included as scenario analyses (see Appendix N1.4.1 for full details). Of these alternative distributions experts indicated that the log normal distribution was more plausible as it better reflected the skewed nature of the data.

Figure 16. Experts' preferred distribution for proportion of patients on SOC receiving stem cell transplants



Proportion of patients on standard of care for whom stem cell transplants are curative

The elicitation question was:

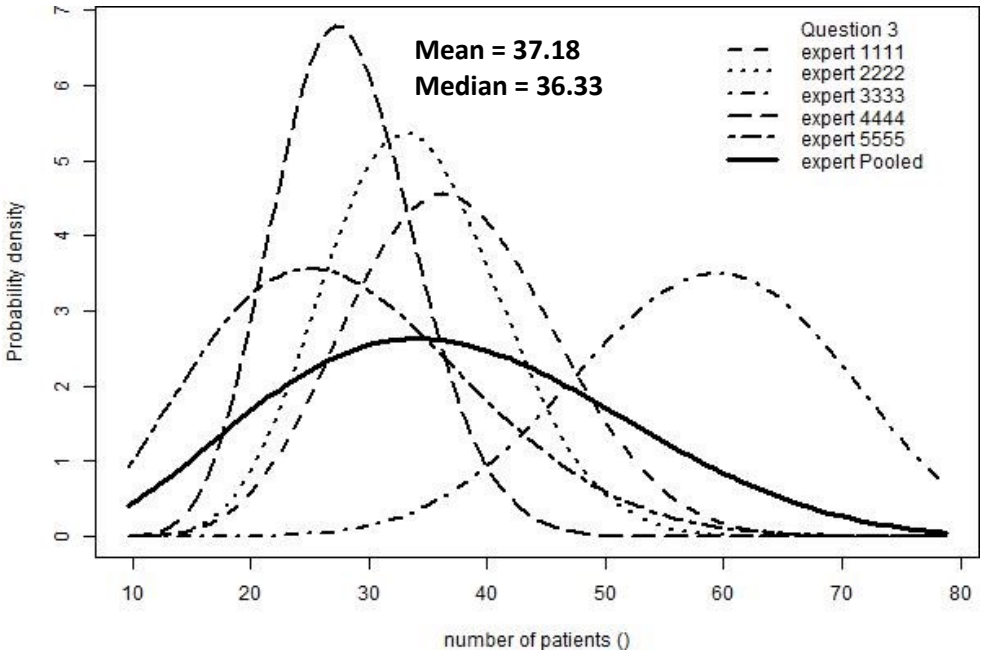
"Imagine a representative cohort of adult patients (of all ages and fitness) in the UK with relapsed or refractory classical Hodgkin lymphoma who have relapsed or not

responded to two lines of chemotherapy and brentuximab vedotin and cannot have autologous stem cell transplant. After treatment with standard of care (e.g. chemotherapy), 100 patients are now able to receive a stem cell transplant. For how many of the 100 patients would the stem cell transplant be curative?"

The experts' responses were typically grouped in the 20-40% range (Figure 17). One response was at a higher value of approximately 60%. This expert stated that if they were completing the exercise again they would move closer to a value of 50%, but not lower than 30-40%. This higher estimate was based on the rationale that cure rates may have improved in comparison to historical data due to improvements in patient selection.

The experts were shown a distribution where the high response had been removed (mean cure rate 31.87%. All experts agreed that the distribution including all responses was more reflective of clinical practice. They drew comparison with data from the AETHERA trial,[77] which they judged to represent patients at high risk of relapse and this be a comparable population. Cure rate in this trial was approximately 40%, which is in line with the pooled estimate. Therefore, the experts' preferred distribution was as shown in Figure 17.

Figure 17. Experts' preferred distribution for proportion of patients on SoC who receive a stem cell transplant that are cured



Proportion of patients on checkpoint inhibitors for whom stem cell transplants are curative

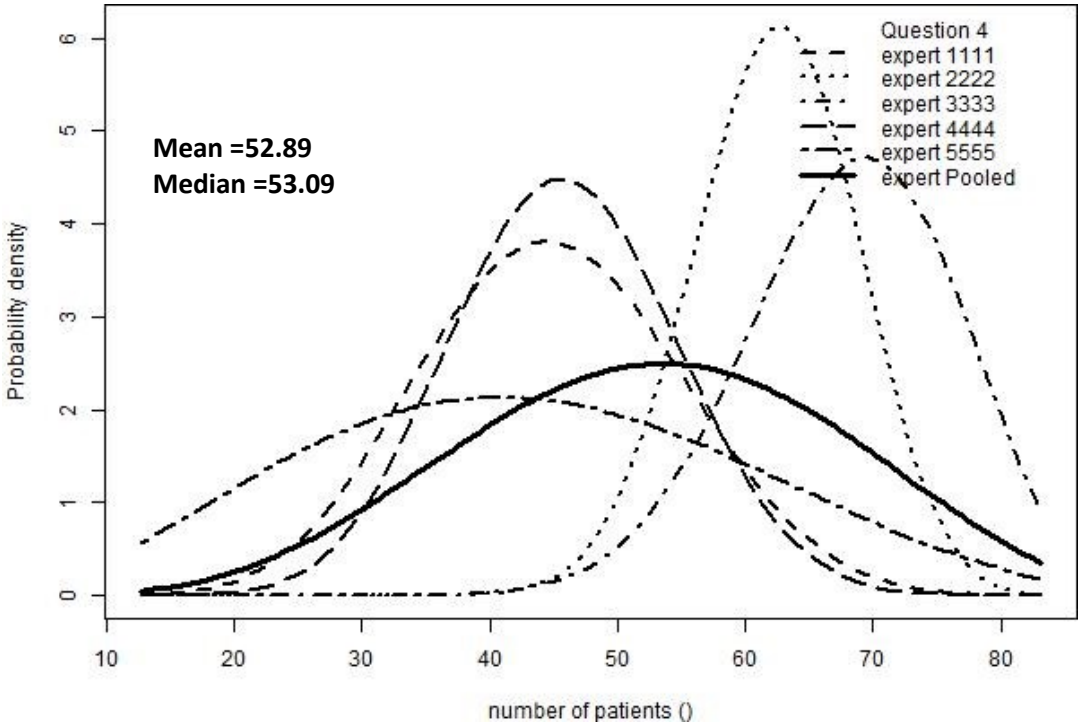
“Imagine a cohort of adult patients in the UK with relapsed or refractory classical Hodgkin lymphoma who have relapsed or not responded to two lines of chemotherapy and brentuximab vedotin and cannot have autologous stem cell transplant. After treatment with a checkpoint inhibitor +/- another line of chemotherapy, 100 patients are now able to receive a stem cell transplant. For how many of the 100 patients would the stem cell transplant be curative?”

The experts indicated that they based on their experiences in clinical practice they would expect an improved cure rate for patients treated with checkpoint inhibitors and agreed that a cure rate of 50-60% was reasonable. Most experts indicated that they had based their results on the Merryman et al study.[53] They highlighted several uncertainties, because:

- Merryman et al is a retrospective study with a small patient numbers
- Results from Merryman may not be generalisable to the patient population of interest because most patients in the study received BV/nivolumab, which is a regimen that is not available in the UK practice
- There is still some uncertainty on the safety profile of checkpoint inhibitors followed by autoSCT

The experts also noted that their estimates may be conservative but that this was based on the uncertainties highlighted above. They agreed that the uncertainty was appropriately captured by the pooled distribution shown in Figure 18.

Figure 18. Experts' preferred distribution for proportion of patients on checkpoint inhibitors who receive a stem cell transplant that are cured



The experts also highlighted that the improved cure rate was not solely due to improved responses with checkpoint inhibitor treatment. They noted that treatment with checkpoint inhibitors could chemosensitise patients, increasing the chances that subsequent lines of chemotherapy would produce stronger responses, which would lead to higher cure rates.

Proportion of patients on checkpoint inhibitors who are alive after 4 years

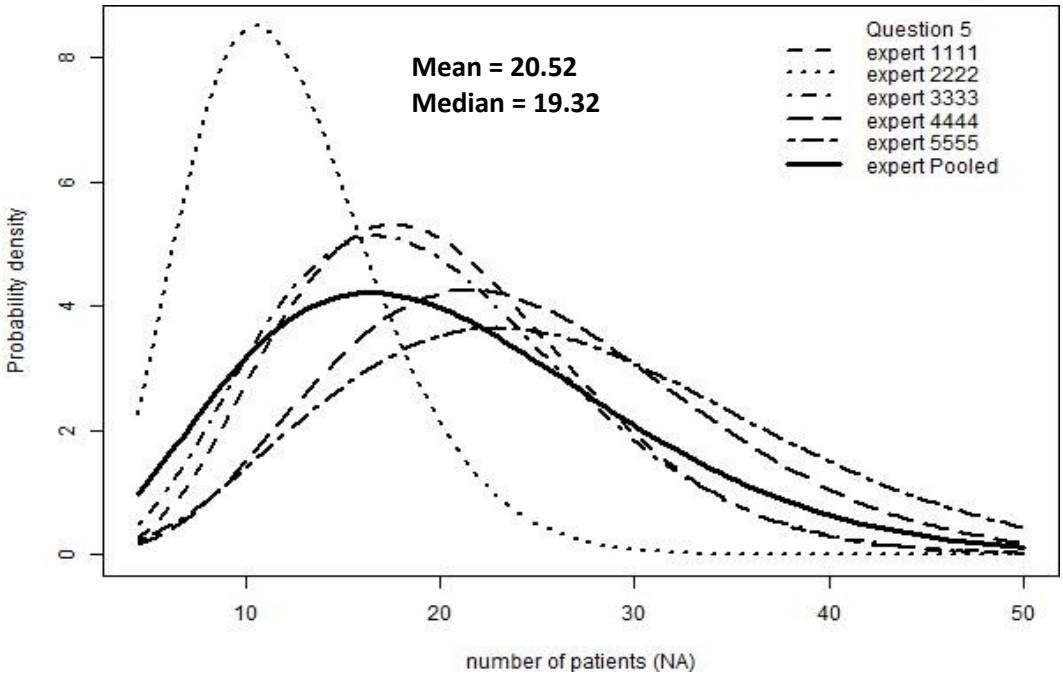
The elicitation question was:

"Imagine a representative cohort of 100 adult patients in the UK (of all ages and fitness) with relapsed or refractory classical Hodgkin lymphoma who have relapsed or not responded to two lines of chemotherapy and brentuximab vedotin and cannot have autologous stem cell transplant. These patients are treated with current standard of care (chemotherapy or best supportive care). How many of the 100 patients will still be alive 4 years after initiation of standard of care?"

The experts agreed that prognosis of patients in this cohort was poor. The expert who had selected the lowest survival rate indicated that they had been considering a patient population where all patients were ineligible for transplant. The experts were

presented with an alternative distribution where this response was removed (mean 22.58, median 21.58). The experts agreed that a survival rate of around 15% would align with expectations and that the distribution including all the experts' responses better reflected clinical reality.

Figure 19. Experts' preferred distribution for proportion of patients on SoC who are alive after 4 years



A summary of the results is provided in Table 37.

Table 37. Summary of results

Parameter	Mean (%)	Median (95% CI)
Proportion of patients on standard of care receiving a stem cell transplant		
Experts' preferred distribution: Expert group 1	8.17	7.51 (2.63, 15.92)
Alternative 1: All experts (log normal distribution)[a]	18.23	12.86
Alternative 2: All experts (beta distribution)[a]	17.48	14.07 (13.99, 45.30)
Proportion of patients on standard of care for whom a stem cell transplant is curative		
Experts' preferred distribution: All experts	37.18	36.33 (15.07, 62.26)
Proportion of patients on checkpoint inhibitors for whom a stem cell transplant is curative		
Experts' preferred distribution: All experts	52.89	53.09 (28.00, 77.12)

Proportion of patients on standard of care alive after 4 years		
Experts' preferred distribution: All experts	20.52	19.32 (6.91, 38.30)
a) for full details of the alternative distributions presented to the experts please see Appendix N1.4.1 (of the two alternative distributions the experts considered the lognormal to be more plausible)		

B.3.3.3.4 Discussion

In addition to providing parameters for use in the economic model, several key insights were generated during the discussion of the results. Notably:

- The estimate for proportion of patients on SoC receiving stem cell transplant (mean 8.17%) was in reasonable alignment with the estimate of 5.3%, which had been accepted in NICE TA524.[26]
- The experts highlighted that the improved cure rate they had estimated for patients treated with checkpoint inhibitors was not solely due to improved response to checkpoint inhibitor treatment compared with standard of care. They noted that treatment with checkpoint inhibitors could chemosensitise patients, increasing the chances that subsequent lines of chemotherapy (including SCT related regimens) would produce stronger responses, which would lead to higher cure rates.
- Prior to the meeting, every expert independently estimated a higher cure rate for SCT after exposure to pembrolizumab than for SCT after standard care alone. There was unanimous agreement at the meeting that although a study investigating this comparison has not been done, the experts believed it to be clinically plausible.

Full discussion of the strengths and limitations of the SEE exercise are discussed in Appendix N. Key strengths were:

- The exercise captured the key parameters required and uncertainty around these, either directly providing model inputs or validating outputs of the economic model.
- Methodology used was aligned to the MRC reference protocol and aimed at fully capturing the extent of uncertainty in the parameter estimates and minimising risk of bias
- Group discussion provided meaningful insights into inter-expert variability. These insights may not have been successful captured by other means (for

example, asking experts to refine their judgement individually based on pooled responses, Delphi panels).

The key limitations were:

- As a group discussion stage was included experts were not asked individually if they wished to change their judgements having observed the pooled responses. The approach of allowing experts to individually adjust their responses has less risk of introducing bias than conducting a group discussion, but may have failed to identify reasons for between-expert variability and would not allow experts to consider these reasons when deciding on whether to adjust their response.
- Experts were asked about overall survival at a single timepoint, which does not allow for time-varying hazards to be explored. This was a pragmatic decision aligned with the chosen model structure and designed to limit the number of questions in the exercise to a manageable quantity.

B.3.3.4 Proportion experiencing adverse events

The cost and HRQoL burden related to adverse events is captured in the economic analyses. The AEs were applied as a one-time cost and disutility in the first cycle of the model. Section B.2.10 details the collection and monitoring of these events in KEYNOTE-087. We incorporated the same AEs as in TA540 [2]. For that appraisal, the specific AEs of interest were originally derived from all cause AEs for grade 3+ as summarised in Table 28; Appendix F. The AEs of interest were selected based on a previous Hodgkin's lymphoma appraisal (TA462) and subsequently validated by clinical experts.[78, 79] At the TA540 Advisory Board, no additional AEs were identified for inclusion in the model, and we note that all other grade 3+ AEs in KEYNOTE-087 occurred in ≤2 patients. As stated in B.2.10.2, pembrolizumab's positive safety profile is well established. Table 38 includes the probability of experiencing AEs of interest for pembrolizumab in cohort 2 of the trial.

Table 38. KEYNOTE-087 adverse events all-cause grades 3+ (cohort 2 only)

Adverse event	Cohort 2 (n=81)	
	Number of events	%

Anaemia	3	3.7%
Diarrhoea	1	1.2%
Dyspnea*	0	0.0%
Fatigue	2	2.5%
Leukopenia	2	2.5%
Nausea	0	0.0%
Neutropenia	3	3.7%
Pyrexia	0	0.0%
Thrombocytopenia	2	2.5%
Vomiting	0	0.0%
*Set to 0 as not reported in the CSR. This AE was included to be consistent with the TA540 methodology.		

As KEYNOTE-087 is a single arm trial, we needed to identify and source the AEs for SoC. In line with TA540, we assumed the same AEs of interest would be relevant in both arms. The treatment-related AEs of grade 3+ for all the SoC treatment options in the model were identified from the literature, mostly using the same sources as TA540, and are summarised in Table 39. It should be noted that these sources are unlikely to match exactly the population of interest. For this CDF exit, the composition of SoC was different to that assumed in TA540 and TA462 (38.46% chemotherapy 18.46% bendamustine and the remainder investigational agents).[27] The new range of SoC regimens is shown in Table 39 below. AE incidence is weighted by the proportion receiving each of the SoC regimens, which have been determined afresh for this appraisal (see section B.3.6.3). For the DECC regimen, no AE incidence was reported in the literature and therefore could not be applied in the economic analysis. However, given that DECC is a palliative chemotherapy regimen of low intensity, limited AEs would be expected.

The weighted AE incidence for SoC is summarised in Table 40. The individual AE incidences were calculated by dividing the number of events per AE by the overall sample size per study to give a percentage. These percentages were then weighted

by the proportion of patients receiving each regimen to calculate a total weighted AE incidence for the SoC arm for use in the model.

Table 39. Chemotherapy adverse events incidence (number of events)

	Benda-mustine	ICE	PMitCEBO	IGEV*	GEM-P*	GDP*	GVD*	DECC	Mini-BEAM	RT
Sample size	36	168	230	91	21	23	37	38	46	30
Anaemia	5	43	0	17	2	2	6	0	0	0
Diarrhoea	0	0	7	0	0	0	1	0	0	0
Dyspnea	0	0	0	0	0	1	4	0	0	22
Fatigue	1	0	0	0	0	2	4	0	0	0
Leukopenia	0	0	0	0	13	0	6	0	0	0
Nausea	1	0	12	1	0	0	0	0	47	0
Neutropenia	2	25	151	26	15	2	19	0	36	0
Pyrexia	1	0	0	0	0	0	0	0	33	0
Thrombocytopenia	7	49	8	18	10	3	16	0	0	0
Vomiting	0	0	12	0	0	3	1	0	46	0
Source	(80) **TA540**	(81) **TA540**	(82)	(83) **TA540**	(84) **TA540**	(85) **TA540**	(86) **TA540**	(87)	(88)	(89)

* The weight has been divided equally among the 4 gemcitabine-based regimens.

Abbreviations: BEAM: carmustine, etoposide, cytarabine, melphalan; DECC: dexamethasone, etoposide, chlorambucil; ICE: ifosfamide, carboplatin, etoposide; IGEV: ifosfamide, gemcitabine, vinorelbine; GDP: gemcitabine, dexamethasone, cisplatin; GEM-P: gemcitabine,

cisplatin, methylprednisolone; GVD: gemcitabine, vinorelbine, doxorubicin; PMitCEBO: prednisolone, mitoxantrone, cyclophosphamide, etoposide, bleomycin, vincristine; RT, radiotherapy.

Table 40. SoC adverse events incidence

Adverse event	Weighted %
Anaemia	7.5%
Diarrhoea	0.5%
Dyspnea	11.0%
Fatigue	1.1%
Leukopenia	2.8%
Nausea	15.8%
Neutropenia	29.2%
Pyrexia	10.6%
Thrombocytopenia	11.9%
Vomiting	15.6%

B.3.3.5 Summary of base case inputs

Table 41. Source of base case OS and SCT treatment effect parameters for pembrolizumab and SoC

Parameter	Source	
	Pembrolizumab	SoC
Pre-landmark		
OS	SACT: total population	HR derived from Bucher ITC using KEYNOTE-204 and TA524
SCT outcomes		
Probability of SCT	SACT: total population	Structured expert elicitation
Ratio of auto and alloSCT	SACT: total population	Assumed equal to pembrolizumab
Probability of curative SCT	Structured expert elicitation	Structured expert elicitation
Post-landmark		
OS: Cured SCT patients	General population mortality	General population mortality

OS: no SCT or failed SCT	SACT: patients who never got an SCT	HR derived from KEYNOTE-204: patients who never got an SCT

B.3.4. Measurement and valuation of health effects

Each health state in the model is associated with a mean utility specific to that state. Although HRQoL was collected in the KEYNOTE-087 trial (as described in Section B.2.6.1.2), this was not included in the cost-effectiveness analysis. For the base case, the utilities for stable and progressed disease are derived from KEYNOTE-204 as described in B.3.5.2.

B.3.4.1 Health-related quality-of-life data from KEYNOTE-087

As stated in B.2.6.1.2, HRQoL was evaluated in the KEYNOTE-087 trial using two QoL measures, the EORTC-QLQ-C30 and the EQ-5D-3L. The EQ-5D-3L was collected i) 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, ii) on treatment discontinuation and iii) 30 days post treatment discontinuation. In KEYNOTE-087 the health state utility values were mapped to the domain scores of the EQ-5D-3L to a single index value, consistent with the NICE guidance.[54] From the CSR, n=81 patients in cohort 2 were recorded in the ASaT population.[33] N=80 patients from cohort 2 were recorded in the PRO FAS population.[90] The mapped EQ-5D-3L health state utility values for cohort 2 from the PRO FAS population are summarised in Table 42.

Table 42. KEYNOTE-087 EQ-5D-3L health utility values based on PFS – cohort 2 PRO FAS[41]

	KEYNOTE-087 (N=80)				
Health states	**n**	**m**	**Mean**	**Standard error**	**95% CI**
Baseline	76	76	0.727	0.030	(0.667, 0.787)
Progression-free*	79	404	0.837	0.010	(0.818, 0.857)

• Ongoing treatment	78	383	0.836	0.010	(0.816, 0.856)
• Completed or discontinued treatment	15	21	0.860	0.050	(0.757, 0.964)
Progressive disease*	41	118	0.824	0.018	(0.789, 0.859)
Unknown*	9	16	0.837	0.044	(0.743, 0.930)

Key

N=81 patients were in cohort 2 from KEYNOTE-087 study, 80 patients satisfied the PRO FAS definition

n = Number of participants with non-missing EQ-5D score.

m = Number of records with non-missing EQ-5D score.

Health state based on Progression-Free Survival Based on BICR per RECIST 1.1.

Unknown = EQ-5D assessment records of censored participants without documented progression per BICR and with time of EQ-5D assessment after (>) censoring date

* EQ-5D score during baseline is excluded.

Summary statistics are computed based on several records per participant treated as independent observations, except for baseline where there is a single record per participant.

Database Cutoff Date: 15MAR2021

B.3.4.2 Health-related quality-of-life studies

A SLR was conducted to inform TA540 on 12th July 2017 identified two quality-of-life studies presenting utility data for patients with R/R cHL (Swinburn et al.2015,[61] Ramsey et al. 2016[91]). Due to immature KEYNOTE-087 trial data at the time, the Swinburn study was used to inform and adjust the utility decrement for PD patients. Ramsey was not used in the economic model as the study did not provide utility data by response status as per the TA540 model structure. A comprehensive SLR update was conducted on 20th February 2023 to identify new HRQoL studies, however, no new published studies met the PICO. Further information on the SLR methodology, search strategy and results is provided in Appendix H.

B.3.4.3 Adverse reactions

The impact of adverse events on HRQoL was explored in the economic analysis. The health disutility associated with a particular AE was estimated by the health

utility decrement from an AE and the time spent in that AE. Using the same approach as in TA540, this was limited to the AE experienced whilst on initial therapy and did not include the events following 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

The disutility values of AEs used in the base case are presented in the Health-related quality of life data used in the cost-effectiveness analysis section.

B.3.5. Health-related quality of life data used in the cost-effectiveness analysis

B.3.5.1 Discussion of KEYNOTE-087 utilities

We decided not to use the KEYNOTE-087 utilities in the base case economic analysis. KEYNOTE-087 trial is a single arm study with no comparative arm. There are also some generalisability concerns relating to the cohort; despite a relatively low proportion receiving SCT, OS is much longer in this trial than in the SACT dataset suggesting a fitter patient group. MSD note it is good practice to ensure health state utility values across model arms are captured from the same data source and instrument[60] where possible as the estimates should have better internal consistency. The KEYNOTE-204 trial captures utilities for both pembrolizumab and BV using the same instrument (EQ-5D) and provides the only comparative source of evidence on pembrolizumab versus a chemotherapy-based regimen. Additionally, KEYNOTE-087 trial is based on n=81 patients in cohort 2, whereas KEYNOTE-204 has a much larger sample size in the overall population (n=300) and n=134 patients in third line without prior SCT for whom EQ-5D data is available. We note that mean EQ-5D on pembrolizumab is very similar between the two trials (0.834 in KEYNOTE-087 and 0.837 in KEYNOTE-204), which supports the generalisabiltiy of KEYNOTE-204.

B.3.5.2 Health state utility values from KEYNOTE-204

HRQoL was evaluated in KEYNOTE-204 using two measures: EORTC-QLQ-C30 questionnaire (version 3.0) which was used to assess cancer-related quality of life, in addition to the generic health status measure, the EQ-5D-3L.Questionnaires were completed at several time points within KEYNOTE 204: pre-dose at Cycle 1 (baseline), Cycle 3 (Week 6), Cycle 5 (Week 12), Cycle 7 (Week 18), and Cycle 9 (Week 24) and every 12 weeks thereafter until PD or up to 1 year while patients receive treatment. Questionnaires were also collected at discontinuation and at the 30-day Safety Follow-up visit. If discontinuation occurred 30 days from the last dose of study treatment, i.e., at the time of the mandatory 30-day Safety Follow-up visit, PRO questionnaires were not repeated. The primary analysis approach for the prespecified PRO endpoints was based on a quality of- life-related full analysis set (FAS) population, which consists of all randomised participants who received at least 1 dose of study treatment and had completed at least 1 PRO assessment.

Outcomes for patients treated with pembrolizumab demonstrated improvements using both scales: Longer PFS in the pembrolizumab group was accompanied by an improvement in health related QOL compared to BV.

HRQoL data were reported directly from patients using the EQ-5D-3L questionnaire. These scores were mapped to the EQ-5D-3L UK value set. For the time period before the landmark, which relates to all patients regardless of progression or SCT status, mean utilities by KEYNOTE-204 trial arm were assigned to the pembrolizumab and SoC arms. We considered this reasonable because the difference is statistically significant, because BV is a more effective treatment than the SoC in this appraisal and therefore is a conservative surrogate, because it was established during NICE TA772 and TA540 that pembrolizumab has a treatment effect on utility as well as disease progression and because the assumption of a persistent utility treatment effect was validated by clinicians at the UK clinical advisory board.

We also assigned a utility treatment effect to the group who had either never received or relapsed after SCT after the landmark. In the absence of direct evidence, the values that we assigned were the mean utilities among “progressed disease”

patients within KEYNOTE-204. We felt this was reasonable because the data are statistically significant and because clinicians at the UK advisory board confirmed that the quality-of-life benefit patients on pembrolizumab receive can persist for many years after treatment. We examined removing this treatment effect so that utility was equal between the groups in sensitivity analysis. In addition, as stated in B.2.10.2, it is established that pembrolizumab is a well-tolerated treatment, therefore it will be reasonable to assume patients will incur little in the way of treatment-related disutility.

We note that the data are from an earlier line of therapy not directly related to the health states within the economic model and therefore these estimates are a source of uncertainty. This is counter-acted to some extent by BV being a conservative surrogate for SoC.

We assigned patients in the cured-SCT health state after landmark the general population utility. This was estimated using age/sex data and the Ara and Brazier[60] utility equation in each cycle. In both KEYNOTE-204 and KEYNOTE-087, the population who received an SCT were significantly younger than the patients who did not. Clinicians at the UK advisory board confirmed that age is a significant determinate of SCT fitness in UK clinical practice. We have requested the SACT data on the age of those undergoing SCT but did not have access at the time of submission so we estimated it instead. We assumed that the mean difference in age of those undergoing and not undergoing SCT from KEYNOTE-204 (11.9 years) would also be observed in SACT. Knowing that 30% of those in SACT received an SCT, we used a goal seeking function to estimate ages for the two groups, which returned values of 45.7 and 57.6 for those undergoing and not undergoing SCT respectively. Using the Ara/Brazier equation gives a baseline utility of 0.883, which is 0.869 by the 4-year landmark for those undergoing an SCT.

Table 43. EQ-5D Health Utility Scores - UK Algorithm (Full Analysis Set Population)

	Pembrolizumab (n=66)					BV (standard of care) (n=68)				
Health states	**n†**	**m‡**	**Mean**	**SE**	**95% CI**	**n†**	**m‡**	**Mean**	**SE**	**95% CI**
Before landmark	66	320	0.837	0.012	(0.815, 0.860)	67	259	0.742	0.016	(0.712, 0.773
Cured	Calculated via Ara-Brazier general population utility equation (0.864 at landmark)									
Post landmark (no SCT or failed SCT)	27	73	0.807	0.026	(0.756, 0.858)	35	64	0.671	0.033	(0.605, 0.738)
Key † patients, ‡ completed forms.										

B.3.5.3 Post-SCT General population utility

Patients in the "cured-SCT" state at 4 years are assigned age matched general population utility "irrespective of health condition" in line with the Ara and Brazier utility set.[60] General population values for each 5-year age band are available from this study. Clinicians at the UK advisory board informed us that it tended to be the younger and fitter patients within the dataset who would receive and tolerate SCT. In SACT, the mean age of patients getting an SCT was 34.4 and the mean age of patients not getting an SCT was 58.1. A difference was also present in KEYNOTE-087, where the SCT cohort had a median age of 33.0 years vs 40 years in the study at large and in KEYNOTE-204, where the SCT cohort had a mean age of 32.7 versus 44.7 in the study at large. As such, using the general population utility for the ages specific to the whole trial population likely underestimates QALYs among patients cured by an SCT and therefore biases the model's results slightly against the pembrolizumab arm, where more SCTs occur. We conducted a scenario analysis where we dropped the mean age in the model to 34.4 to explore the potential scale of this underestimate.

B.3.5.4 Age-related utility decrements

The Ara and Brazier study (2010)[60] also considers background disutility due to ageing which are applied to all health states in the model as recommended in the

NICE methods guide.[54] Table 44 includes the coefficients from the Ara and Brazier (2010) study. The proportion male and baseline age are informed by the SACT baseline characteristics. The median age from SACT was used to determine decrements in all health states across the model's time horizon.

Table 44. Regression coefficients used for the estimation of age-related disutility

Parameter	Coefficient	Source
Age (years)	-0.0002587	Ara and Brazier et al. (2010)[60]
Age²	-0.0000332	
Male	0.0212126	
Intercept	0.9508566	

B.3.5.5 Disutility due to adverse events

Given the absence of disutilities from KEYNOTE-087 or in any R/R HL study, disutilities were identified in other oncology studies. MSD applied the same methodology as in the TA540 submission of sourcing alternative data inputs from oncology publications in published literature (including 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 are summarised in Table 45.

Table 45. Summary of disutility sources

Source	Disease area	Population (sample size)	Method of valuation	Country
Beusterien (2010)[92]	Chronic lymphocytic leukaemia	General public (n=89)	SG	UK
Doyle (2008)[93]	Non-small cell lung cancer	General public (n=101)	SG & VAS	UK
Lloyd (2006)[94]	Breast cancer	General public (n=100)	SG	UK
Nafees (2008)[95]	Small cell lung cancer	General public (n=100)	TTO	UK

Shingler (2013)[96]	Soft tissue sarcoma	General public (n=100)	TTO	UK
Tolley (2013)[97]	Late-stage chronic lymphocytic leukaemia	General public (n=110)	TTO	UK
PEGASUS-TIMI 54 study (TA420)[98]	Post myocardial infarction	Trial population (n=21,162 [n=118,745 completed questionnaires; 0 to 54 months])	EQ-5D-3L (UK value set)	Global

Of the disutility studies, PEGASUS-TIMI-54 was the only trial identified from the search. The remaining studies were based on general population estimates. Where multiple sources were available, the average was applied across the studies.

The disutilities and duration sources from the literature are presented in Table 46, with a summary of the final model inputs in Table 47. The AE durations were primarily sourced from the TA306[99] and TA476[100] appraisals over the published literature as the trials were based on patient level data. In TA306, the trial (PIX301) was a phase III study in relapsed aggressive non-Hodgkin's lymphoma and TA476[100] (CA046) was a phase III study on locally advanced untreated pancreatic cancer. 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 applied to give an overall AE duration.

Table 46. Adverse event disutilities and durations from the literature

Adverse event	Disutility	Source	Used in	Duration (days)	Source
Anaemia	-0.09	Beusterien (2010)[92]	TA462[27]	16.1	TA306[99]
				(12.4+14.5)/2 = 13.45	TA476[100]
Diarrhoea	-0.08	Beusterien (2010)[92]	TA462[27]	(5.567+5.5)/2 = 5.53	TA476[100]

	-0.0468	Nafees (2008)[95]	TA395[101]		
	-0.103	Lloyd (2006)[94]			
	-0.327	Shingler (2013)[96]			
Dyspnea	-0.0481	PEGASUS-TIMI 54 study (TA420)[98]	TA420[98]	12.7	TA476[100]
Fatigue	-0.07346	Nafees (2008)[95]	TA462[27] TA440[102] TA411[103] TA395[101]	31.5	TA476[100]
	-0.262	Shingler (2013)[96]	TA440[102]	(19.885+19.14)/ 2 = 19.51	TA476[100]
	-0.115	Lloyd (2006)[94]			
Leukopenia	Assumed same as neutropenia			14	TA306[99]
				(10.041+10.4) = 10.22	TA476[100]
Nausea	-0.04802	Nafees (2008)[95]	TA462[27] TA411[103] TA395[101]TA476[100]	6	TA306[99]
	-0.357	Shingler (2013)[96]	--	(11.179+20.933) /2 = 16.06	TA476[100]
	-0.05	Beusterien (2010)[92]			
Neutropenia	-0.08973	Nafees (2008)[95]	TA462[27]	15.1	TA306[99]
	-0.163	Tolley (2013)[97]	TA359[104]	(9.547+9.291)/2	TA476[100]
				9.42	
Pyrexia	-0.11	Beusterien (2010)[92]		12.3	TA306[99]
Thrombocytopenia	-0.108	Tolley (2013)[97]	TA462[27] TA359[104]TA476[100]	23.2	TA306[99]
				(8.057+9.32)/2 = 8.69	TA476[100]

Vomiting	-0.04802	Nafees (2008)[95]	TA462 [27] TA411[103]	2.3	TA306[99]
	-0.357	Shingler (2013)[96]	--	(5.852+10.875)/ 2 = 8.36	TA476[100]
	-0.05	Beusterien (2010)[92]			
	-0.103	Lloyd (2006)[94]			

Table 47. Summary of adverse event disutilities and duration

Adverse event	Disutility	Duration (days)	Disutility per event duration
Anaemia	-0.0900	14.8	-0.0036
Diarrhoea	-0.1392	5.5	-0.0021
Dyspnea	-0.0481	12.7	-0.0017
Fatigue	-0.1502	25.5	-0.0105
Leukopenia	-0.1264	12.1	-0.0042
Nausea	-0.1517	11	-0.0046
Neutropenia	-0.1264	12.3	-0.0043
Pyrexia	-0.1100	12.3	-0.0037
Thrombocytopenia	-0.1080	15.9	-0.0047
Vomiting	-0.1395	5.3	-0.0020

Pembrolizumab has a well-tolerated and manageable safety profile in cHL as demonstrated in KEYNOTE-087 and is also favourable compared with BV in as supported by the KEYNOTE-204 trial. In addition, the mean utilities outlined in

Table 43 theoretically already account for the toxicity of pembrolizumab and BV, so including specific AE-associated disutilities may lead to a small amount of double counting. To explore this potential uncertainty, a scenario analysis was conducted where it was assumed AE disutility was set to 0 across all therapies.

B.3.5.6 Disutility associated with SCT procedure

SCT is a difficult procedure to undergo and is associated with prolonged hospital stays that have a significant but mostly temporary impact on HRQoL. We therefore felt it appropriate to try to capture this in the economic model.

HRQoL was an outcome of interest in both the clinical and economic SLRs but, as no relevant evidence was identified during those processes, a targeted review was conducted to identify more recent and relevant utility studies focusing on the complications following SCT treatment. We reviewed recent NICE technology appraisals, searched references we found there and performed internet searches for relevant papers. There were a paucity of directly relevant data in other cHL appraisals. No assumptions were made in TA772[(1)] and in TA540,[(2)] data from a study of alloSCT complications in ALL and AML was used to estimate disutilities for Graft vs Host Disease. In TA524,[(26)] we identified a table of potentially relevant EQ-5D data collected from cHL and NHL patients undergoing SCTs. We calculated a decrement for SCT for the time periods reported in this table and an associated total QALY disbenefit from the SCT procedure. The absolute QALY decrement was calculated using a weighted average of the decrements for each time point by the time point (in years). We capped the longer term disbenefit at 2 years as experts at the UK clinical advisory board told us that all SCT-related events of interest would have been resolved by this point.[(5)]

A summary of the utilities from TA524 for each time point and the decrement differential between each time point and absolute QALY decrement is summarised below in Table 48.

Table 48. Absolute QALY decrement for SCT (adapted from TA524)

Time period	EQ-5D	Decrement	Years
Utility Event Free Survival (cHL/NHL)	0.82		

Utility 0-14 days after SCT	0.42	0.4	0.038
Utility 14 days to 3 months	0.6	0.22	0.212
Utility 3 months+ (capped at 2 years)	0.77	0.05	1.750
Absolute QALY decrement			**0.149**

This calculation is fairly crude, not explicitly accounting for auto/allo splits or proportions of patients expected to experience acute or chronic Graft vs Host Disease, for example, but at least is based on data that have been used to quantify for the HRQoL impact of SCT in a similar NICE appraisal. This parameter can be arbitrarily varied in sensitivity analysis but we note that it is largely relevant to the cost-effectiveness of SCT rather than pembrolizumab. Since helping patients get to SCT is one of the primary goals of therapy in this indication, and society has already decided SCT is worth paying for, it would be perverse if the inflation of this parameter negatively influenced the decision to recommend pembrolizumab.

B.3.5.7 Parameters used in the cost effectiveness analysis

The full list of variables used in the cost effectiveness analysis is presented in Appendix O.

B.3.6. Resource identification, measurement and valuation studies

A SLR was conducted in support of TA540 on 12th July 2017 to identify cost and resource use in the treatment and ongoing management of R/R cHL patients in the economic analysis. An update of the SLR was conducted on 20th February 2023 using the same methodology to identify new evidence relevant to the population of interest. The original SLR on 12th July 2017 identified 14 unique studies from 17 publications, of which only one (Radford 2014[21]) was UK-specific. The updated SLR on 20th February 2023 identified 8 studies, but none were UK-specific. Therefore, no cost or resource use data from the SLR was considered for this economic analysis. Further information on the SLR methodology, search strategy and results is provided in Appendix I section.

B.3.6.1 Use of NHS reference costs of 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. It is well established that the administration cost of pembrolizumab is equivalent to NHS reference cost code SB12Z[64] for the "administration of a simple chemotherapy", with an infusion lasting 30 minutes.

B.3.6.2 Input from clinical experts

The costing approach detailed here was previously validated with clinical experts in previous HTA submissions of pembrolizumab.[105, 106]

B.3.6.3 Intervention and comparators' costs and resource use

B.3.6.3.1 Pembrolizumab

Pembrolizumab is supplied as 100mg vials and the cost effectiveness model assumes a fixed dose of 200mg every 3 weeks (Q3W). This is aligned with the licensed dose of pembrolizumab as well as the dosing in KEYNOTE-087. The list price of a 100mg vial is £2,630.[107] Based on 2 x 100 mg vials applying list price, the drug acquisition cost of pembrolizumab per cycle is £5,260. Pembrolizumab is available to the NHS for a reduced price via a Commercial Access Agreement, which is a simple discount. For details of the CAA currently in place with the discount please refer to Table 2. All analyses will apply the PAS discount so that the true cost-effectiveness of pembrolizumab can be openly discussed in the results section.

In addition, pembrolizumab monotherapy is also licensed at a fixed dose of 400mg every 6 weeks (Q6W) as per the EMA and this schedule was permitted in the SACT Clinical Treatment Criteria.[8] To assess this impact, the Q6W regimen will be presented in scenario analyses.

<u>Treatment duration</u>

As per the marketing authorisation, patients treated with pembrolizumab are treated until disease progression is confirmed, if unacceptable toxicities occur or if they reach NHS England's 35-cycle stopping rule. In addition, patients also discontinue in the following situations (whichever is soonest):

- Patient receives an SCT
- Loss of clinical benefit
- Excessive toxicity
- Patient choice to discontinue treatment

Time on Treatment (ToT) was captured in the economic model. The most relevant ToT data for the model come from SACT rather than KEYNOTE-087. This is because pembrolizumab was often used as a bridge to transplant and because the patient mix is more reflective of UK practice. Table 49 summarises the modelled drug acquisition costs for pembrolizumab per cycle.

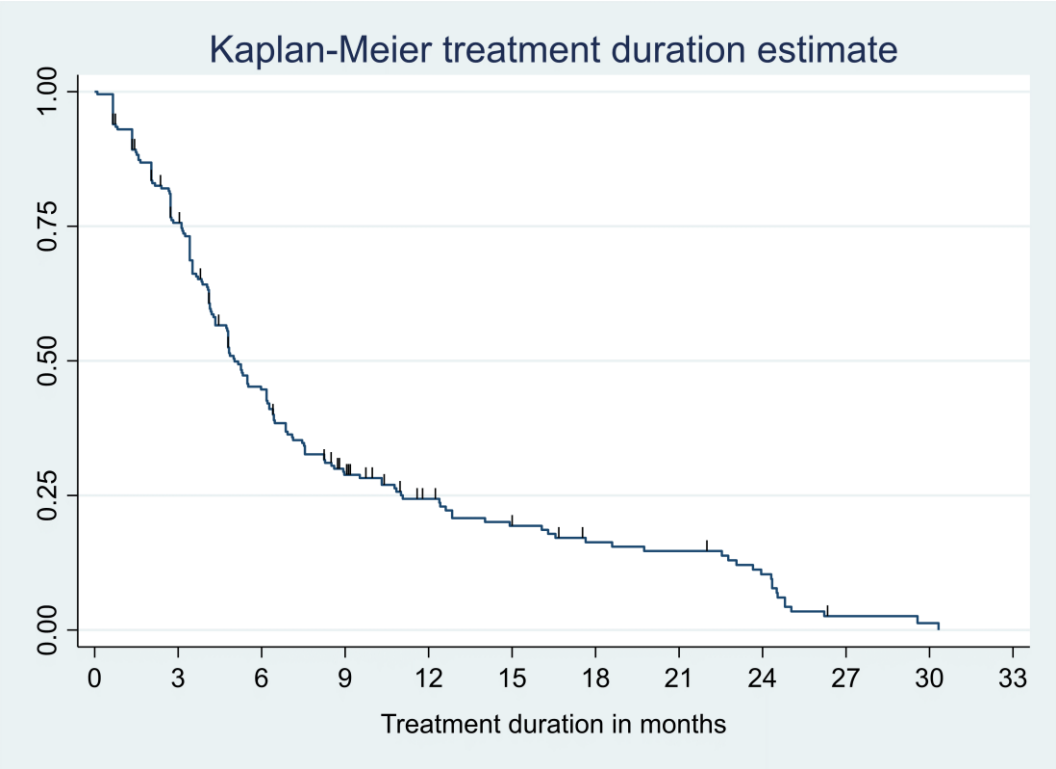
Table 49. Pembrolizumab dosage, administration, treatment duration, vial size and list price.

Treatment	Dosage (mg)	Admins per cycle	Cycle length (days)	Max cycles	Vial size (mg)	Vial (1) price, £/vial	Vials used	Cost (£) per cycle
Pembrolizumab	200	1	21	35	100	2,630	2.00	5,260

Administration costs for pembrolizumab

Pembrolizumab is administered Q3W or Q6W as an intravenous infusion over 30 minutes. The HRG code SB12Z Deliver Simple Parenteral Chemotherapy at First Attendance was sourced from NHS reference costs[64] was applied to every cycle of pembrolizumab. Based on the 2021/22 cost year, the administration of SB12Z using the total HRG cost for every cycle is £286.71.59. The ToT KM curve from the SACT report is shown in Figure 20.

Figure 20. Treatment duration for pembrolizumab (N=215)



Overall, cost per cycle, inclusive of PAS discount and administration cost was assigned to each patient still on treatment, as dictated by the ToT curve from SACT. No “relative dose intensity” (RDI) is available from SACT. In pembrolizumab trials, this usually captures patients who are taking a treatment break but resume afterwards and is usually the cause of a small tail on the ToT KM curve after 2 years. Our approach to costing is in NICE Technology Appraisals is to either assume 100% RDI and a hard stop at 35 cycles or to use the whole KM curve and apply an RDI if available. Because of the lack of RDI data, only the first approach is available.

B.3.6.3.2 Standard of care

Table 50 summarises the regimens included in the blended comparator arm of the model. The sources for these are based on Cheah et al. (2016),[4] Eyre et al. (2017) and the Clinical Advisory Board conducted by MSD.[5] For this CDF exit, the approach for SoC composition departed from TA540 which weighted the SoC into chemotherapy (38.46%), bendamustine (18.46%) and investigational agents (43.1%) that was originally based on the TA462[27] appraisal. [5] For this CDF exit, the approach for SoC composition departed from TA540, which weighted the SoC into chemotherapy (38.46%), bendamustine (18.46%) and investigational agents

(43.1%). This was originally based on the TA462[27] appraisal. As stated in section B.3.2.5 on "Intervention technology and comparators", the clinical experts could not confidently provide proportions for SoC, therefore an assumption was made to distribute the SoC treatments equally. The proportions were varied in sensitivity analysis.

Table 50. Composition of SoC

Treatments	%	Source
Bendamustine	14.29	Cheah et al. (2016)[4], Eyre et al. (2017)[67] and MSD Advisory Board[5]
ICE (ifosfamide, carboplatin, etoposide)	14.29	
Weekly chemotherapy (PMitCEBO)	14.29	
Gemcitabine-based (IGEV, GEM-P, GDP, GVD)*	14.29*	
Oral chemotherapy (DECC; dexamethasone, etoposide, chlorambucil, lomustine)	14.29	
Radiotherapy	14.29	
Mini-BEAM	14.29	
Abbreviations: BV, brentuximab vedotin; DECC, dexamethasone, etoposide, chlorambucil, lomustine; ICE, ifosfamide, carboplatin, etoposide; Mini-BEAM, carmustine, etoposide, cytarabine, melphalan; PMitCEBO, prednisolone, mitoxantrone, cyclophosphamide, etoposide, bleomycin and oncovin. **the weight has been divided equally among the 4 gemcitabine based regimens*		

The dosing and cycle details for the SoC treatment regimens and components of SoC are summarised in Table 51.

Table 51. Dosing and cycle descriptions for SoC

Regimen/Treatment		Dosing	Cycle
Bendamustine	-	120mg/m² on 2 days per cycle	Cycle length of 28 days, to a maximum of 6 cycles
ICE	Ifosfamide	5000mg/m² on 1 day per cycle	Cycle length of 14 days, to a

	Mesna	5000mg/m² on 1 day per cycle	maximum of 3 cycles
	Carboplatin	800mg on 1 day per cycle	
	Etoposide	100mg/m² on 3 days per cycle	
Weekly chemotherapy (PMitCEBO)	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	300mg/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)	
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/m^2 on 2 days per cycle	
Oral chemotherapy (DECC)	Dexamethasone	6mg/m^2, once a day for 5 days per cycle	Cycle length is 6 weeks (if well tolerated then can reduce to 28 days), for maximum 6 cycles
	Etoposide	150mg/m^2, once a day for 3 days per cycle	
	Chlorambucil	15mg/m^2, in 3 divided doses for 4 days per cycle	
	Lomustine	80mg/m^2, on 1 day per cycle	
	Ondanestron	8mg, twice a day for 5 days	
Mini-BEAM	Carmustine	60mg/m^2, on 1 day per cycle	Cycle length of 21-28 days depending on blood count recovery, duration of 3 cycles
	Etoposide	100mg/m^2, twice a day for 4 days per cycle	
	Cytarabine	150mg/m^2, once a day for 4 days per cycle	
	Melphalan	30mg/mg^2, on 1 day per cycle	
Sources: TA540,[2] Collins et al. (2014)[23], Northern Cancer Alliance,[108] Lymphoma Group.[109]			

The comparator acquisition costs for each SoC component are summarised in Table 52. The primary source for the acquisition costs were from the electronic market information (eMit) as these include the average price paid for medicines in the NHS.[65] Where the average prices for therapies were not available via eMit, prices were obtained from the current BNF. Given the differences in strengths and pack sizes among for each component, the model contained up to a maximum of four vial/pack size for each component. The lowest cost combination of vials to make up the required dosage for the average patient, inclusive of drug wastage, was calculated within the model. The cost per unit is calculated using the pack cost divided by the units per pack.

Table 52. Unit cost from sourced prices of SoC regimen components in the UK

Component	Strength	Units per pack	Pack cost (£)	Source	Cost per pack
Ifosfamide	1000mg	1	£151.49	BNF (2023)[66]	£151.49
	2000mg	1	£273.77		£273.77
Mesna	400mg	15	£203.24	eMIT (2023)[65]	£13.55
	1000mg	15	£442.75		£29.52
Carboplatin	50mg	1	£3.89	eMIT (2023)[65]	£3.89
	150mg	1	£6.29		£6.29
	450mg	1	£15.16		£15.16
	600mg	1	£21.32		£21.32
Etoposide	100mg	1	£3.94	eMIT (2023)[65]	£3.94
	500mg	1	£14.79		£14.79
Bleomycin	15mg (15,000 unit)	10	£190.60	BNF (2023)[66]	£19.60
Cyclophosphamide	500mg	1	£8.43	eMIT (2023)[65]	£8.43
	1000mg	1	£13.23		£13.23
	2000mg	1	£27.50		£27.50
Vincristine	1mg	1	£7.10	eMIT (2023)[65]	£7.10
	2mg	1	£16.76		£16.76
Chlorambucil	2mg	25	£11.15	BNF (2023)[66]	£0.45

Lomustine	40mg	20	£780.82	BNF (2023)[66]	£39.04
Ondanestron	4mg	30	£0.93	eMIT (2023)[65]	£0.03
	8mg	10	£0.76		£0.08
Carmustine	7.7mg	8	£5,203	BNF (2023)[66] as cited in TA462	£650.38
Melphalan	50mg	1	£35.88	eMIT (2023)[65]	£35.88
Mitoxantrone	20mg	1	£67.24	eMIT (2023)[65]	£67.24
Gemcitabine	1000mg	1	£9.36	eMIT (2023)[65]	£9.36
	2000mg	1	£19.64		£19.64
Vinorelbine	10mg	10	£74.72	eMIT (2023)[65]	£7.47
	50mg	10	£159.59		£15.96
Prednisolone	1mg	28	£0.20	eMIT (2023)[65]	£0.01
	5mg	28	£1.23		£0.04
	25mg	56	£12.41		£0.22
Cisplatin	10mg	1	£2.71	eMIT (2023)[65]	£2.71
	50mg	1	£9.10		£9.10
	100mg	1	£10.97		£10.97
Methyl-prednisolone	40mg	1	£1.58	eMIT (2023)[65]	£1.58
	125mg	1	£4.77		£4.77
	500mg	1	£5.66		£5.66
	1000mg	1	£11.46		£11.46
Dexamethasone	0.4mg	30	£2.56	eMIT (2023)[65]	£0.09
	2mg	50	£2.46		£0.05
	4mg	50	£30.73		£0.59
Pegylated liposomal doxorubicin	20mg	1	£317.85	eMIT (2023)[65]	£317.85
	50mg	1	£610.74		£610.74
Cytarabine	100mg	5	£12.84		£2.57

	500mg	5	£15.13	eMIT (2023)[65]	£3.03
	1000mg	1	£8.28		£8.28
	2000mg	1	£12.66		£12.66
Doxorubicin	10mg	1	£4.52	eMIT (2023)[65]	£4.52
	50mg	1	£7.29		£7.29
Bendamustine	100mg	5	£77.70	eMIT (2023)[65]	£15.54
	25mg	5	£28.75		£5.75
Abbreviations: BNF, British National Formulary; eMIT, electronic Market Information Tool.					

The cost for SoC administration (per cycle) is summarised below in Table 53. The cost of administration for the SoC chemotherapy regimens were sourced from NHS reference costs. The administration cost of SoC per cycle is calculated by multiplying the cost per administration for each component of the various regimens by the respective frequency in each cycle. HRG codes SB14Z[64] and SB15Z[64] were applied to delivering complex chemotherapy at first attendance and delivering subsequent elements of a chemotherapy cycle respectively in line with the TA462[27] appraisal and as applied in TA540.[2]

Table 53. Administration cost of SoC per cycle[64]

Regimen	Administration	Description
Bendamustine	£573.42	Delivering complex chemotherapy at first attendance and delivering a subsequent complex chemotherapy element within the same cycle
ICE	£1,211.82	Delivering complex chemotherapy at first attendance and delivering two subsequent complex chemotherapy elements within the same cycle
Weekly chemotherapy (PMitCEBO)	£843.38	Delivering complex chemotherapy at first attendance and delivering a subsequent complex chemotherapy element within the same cycle
IGEV	£1,580.26	Delivering complex chemotherapy at first attendance and delivering three subsequent complex chemotherapy elements within the same cycle

GEM-P	£1,211.82	Delivering complex chemotherapy at first attendance and delivering a subsequent complex chemotherapy element within the same cycle
GDP	£843.38	Delivering complex chemotherapy at first attendance and delivering a subsequent complex chemotherapy element within the same cycle
GVD	£843.38	Delivering complex chemotherapy at first attendance and delivering a subsequent complex chemotherapy element within the same cycle
Oral chemotherapy (DECC)	N/A. Oral only therefore no administration cost.	
Radiotherapy	N/A. Applied as one-off cost.	
Mini-BEAM	£1,948.70	Delivering complex chemotherapy at first attendance and delivering four subsequent complex chemotherapy elements within the same cycle
Where delivering complex chemotherapy at first attendance is £474.94 (SB14Z), and delivering a s subsequent complex chemotherapy element within the same cycle is £368.44 (SB15Z). Source: NHS Reference Costs 2021/22.[64]		

Table 54 summarises the acquisition cost per cycle for the SoC regimens used along with the respective cycle length and maximum number of treatment cycles. 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 GEM-P was applied at the start of week 0, 4, 8 and 12). Not all patients are expected to get to the maximum treatment cycles in all SoC regimens but we were unable to find data on mean cycles for each. Instead we down-weighted the total cost of SoC using the Time on Treatment curve from the BV arm of KEYNOTE-204. This approach was conceptually similar to the approach in TA540, which used the SoC's calculated PFS curve to do the same. Overall the down-weighting meant that SoC treatment costs were ~90% of the maximum. This is because the SoC regimens are uniformly short in duration. We were not sure whether this was an overestimate so examined an arbitrary reduction to 70% in sensitivity analysis.

Table 54 shows the max and weighted costs for each SoC regimen. Although costs do not accrue in the model as simply as this, we included this table to give a sense of the relative contribution of the different regimens to SoC costs.

Table 54. Acquisition costs per cycle and maximum number of cycles

Regimen	Cost per cycle (£)	Cycle length (days)	Maximum number of cycles
Bendamustine	69.00	28	6
ICE	1,379.91	14	3
IGEV	2,351.67	21	4
GEM-P	13.77	28	3
GDP	58.60	21	2
GVD	736.17	21	2
PMitCEBO	1,928.36	14	8
DECC	195.20	42	6
Mini-BEAM	9,694.81	28	3
Radiotherapy	See Table 56	N/A	N/A

Source: TA540,[2] BNF,[66] eMIT,[65] Collins et al. (2014),[23] Northern Cancer Alliance,[108] Lymphoma Group.[109]

Abbreviations: BEAM: carmustine, etoposide, cytarabine, melphalan; DECC: dexamethasone, etoposide, chlorambucil; ICE: ifosfamide, carboplatin, etoposide; IGEV: ifosfamide, gemcitabine, vinorelbine; GDP: gemcitabine, dexamethasone, cisplatin; GEM-P: gemcitabine, cisplatin, methylprednisolone; GVD: gemcitabine, vinorelbine, doxorubicin; PMitCEBO: prednisolone, mitoxantrone, cyclophosphamide, etoposide, bleomycin, vincristine.

Table 55. Total drug acquisition and administration costs for SoC

SoC regimen	Weight	Max cycles	Total Cost per Cycle	Weighted max costs
Bendamustine	14.29%	6.00	£642.42	£550.81
ICE	14.29%	3.00	£2,591.72	£1,111.07
IGEV	3.57%	4.00	£3,931.93	£561.48
GDP	3.57%	2.00	£901.98	£64.40
GEM-P	3.57%	3.00	£1,335.59	£143.04

GVD	3.57%	2.00	£1,579.55	£112.78
DECC	14.29%	6.00	£195.19	£167.36
PMitCEBO	14.29%	8.00	£2,771.74	£3,168.65
miniBEAM	14.29%	3.00	£11,643.51	£4,991.57
Radiotherapy (see below)	14.29%	1.00	£5,340.59	£763.17
SoC total weighted average at max cycles				£11,634.33

Source: TA540,[2] BNF,[66] eMIT,[65] Collins et al. (2014),[23] Northern Cancer Alliance,[108] Lymphoma Group.[109]

Abbreviations: BEAM: carmustine, etoposide, cytarabine, melphalan; DECC: dexamethasone, etoposide, chlorambucil; ICE: ifosfamide, carboplatin, etoposide; IGEV: ifosfamide, gemcitabine, vinorelbine; GDP: gemcitabine, dexamethasone, cisplatin; GEM-P: gemcitabine, cisplatin, methylprednisolone; GVD: gemcitabine, vinorelbine, doxorubicin; PMitCEBO: prednisolone, mitoxantrone, cyclophosphamide, etoposide, bleomycin, vincristine.

B.3.6.4 Radiotherapy cost

As mentioned, radiotherapy is an added treatment to the SoC composition and applied as a one-time cost in the model. Input from clinical experts advised that the use of radiotherapy in the R/R cHL population would most commonly be palliative, and therefore the cost of radiotherapy was sourced accordingly. As the evidence around the cost and resource use of radiotherapy in cHL is limited, MSD based the general approach on published NICE Guidelines for Lung Cancer (NG122) in the non-small cell lung cancer (NSCLC).[110] Consistent with NICE guidance,[54] each radiotherapy component was sourced using NHS reference costs. Table 56 summarises the HRG codes, descriptions, and multipliers (number of resource units) for each radiotherapy component.[64] To ensure the radiotherapy costs are R/R cHL specific, the number of fractions and admissions were reduced to 5 based on the Royal College of Radiologists (RCR) guidelines for relapsed HL.[111] Same day radiotherapy attendance was calculated using the weighted average across elective, non-elective, and day case same day radiotherapy admission or attendance from NHS reference costs. [64] To ensure the radiotherapy costs are R/R cHL specific, the number of fractions were decreased from 35 to 20 fractions based on the Royal College of Radiologists (RCR) guidelines for R/R cHL.[111] Same day radiotherapy attendance was calculated using the weighted average of all same day radiotherapy

admission or attendance from NHS reference costs. The total radiotherapy cost is then calculated using a weighted average of the multipliers and cost per HRG radiotherapy component.

Table 56. Cost of radiotherapy

HRG code	Description	Multiplier	Cost
SC56Z[64]	Other External Beam Radiotherapy Preparation	1	£911.34
SC22Z[64]	Deliver a Fraction of Treatment on a Megavoltage Machine	5	£177.61
SC97Z[64]	Same day Radiotherapy Admission or Attendance (weighted average)	5	£708.24
Total			**£5,340.59**
Abbreviation: RT, Radiotherapy. For SC97Z, the weighted average was taken across Elective, Non-elective long stay, Non-elective short stay, and Day case).			

B.3.6.5 Health-state unit costs and resource use

At an MSD advisory board meeting clinical advisors were presented with the resource use estimates from TA540[2] (originally sourced from TA524[26]). The advisors agreed that, during treatment:

- Monthly blood tests are appropriate but may be more frequent if, for example, the patient has a low blood count
- PET scans should be conducted every 3–4 months rather than every 6 months
- Not many HCPs carry out computerised tomography (CT) scans in current clinical practice.

Based on this feedback, CT scans have been omitted and the frequency of PET scans adjusted.

Following the advisory board, we emailed the advisors to ask whether resource would reduce for patients post 4-years who hadn't had or relapsed after SCT. Practice was variable but the general consensus was that appointments would be 6-monthly, scans and blood tests would be yearly or not at all if the patient was not symptomatic. We therefore costed these yearly. The costs and resource use pre-

landmark and post-landmark for patients with no SCT or failed SCT are summarised below in Table 57.

The clinicians at the UK advisory board confirmed it was appropriate to assume no health state resource costs for patients who were cured by their SCT. It is assumed that all SCT related costs and follow up are captured either within the SCT costs themselves, which are very high (see Section B.3.6.7), or via the general pre-landmark costs.

Table 57. Costs and resource use pre-landmark for all patients and post-landmark for patients with no SCT or relapsed after SCT

Resource	Unit Cost	Source	Weekly usage		Source
			Pre-landmark: all patients	Post-landmark: no/relapsed SCT	
Outpatient attendance	£209.41	NHS Reference costs 2021-22, 303: Clinical Haematology Service, Consultant let follow-up attendance, non-admitted face to face.[64]	1/month	2/year	NICE TA446 (replaced by TA524) as cited in TA540 and TA772 and validated by MSD Advisory Board, July 2023.
Biochemistry	£1.55	NHS Reference costs 2021-22, DAP204: Clinical Biochemistry.[64]	1/month	1/year	NICE TA446 (replaced by TA524) as cited in TA540 and TA772 and validated by MSD Advisory Board, July 2023.
Cell blood count	£2.96	NHS Reference costs 2021-22, DAPS05: Haematology. [64]	1/month	1/year	NICE TA446 (replaced by TA524) as cited in TA540 and TA772 and validated by

					MSD Advisory Board, July 2023.
PET scan	£927.78	NHS Reference costs 2021-22, RN03A: Positron Emission Tomography with computed Tomography (PETCT) of more than three areas[64]	4/year	1/year	NICE TA446 (replaced by TA524) as cited in TA540 and TA772 (0.03). Value adjusted and validated based on MSD Advisory Board, July 2023.
Abbreviations: PET; Positron Emission Tomography; SCT: Stem Cell Transplant.					

B.3.6.6 Miscellaneous unit costs and resource use

B.3.6.6.1 Subsequent therapy cost

From our consultations with experts at the UK advisory board, it was unclear the extent to which pembrolizumab would displace treatments in the SoC pathway. While pembrolizumab effectively provides an additional line of treatment, there are reasons to think that subsequent treatment use might be lower in the pembrolizumab arm. While it is clear from the SACT data that some patients will receive chemotherapy after pembrolizumab, it is also logical that patients who are cured by SCT will not require subsequent treatments. Based on expert feedback at the UK advisory board, some patients receive good disease control for many years after pembrolizumab and may not need a subsequent treatment. It is also likely that a further line of treatment would reduce patient fitness and that some patients would have died before subsequent treatment is considered. In the base case we made the following assumptions:-

- The proportion of patients receiving any subsequent treatments would be taken from KEYNOTE-204, with the BV arm representing SoC.
- In the absence of good evidence, subsequent treatment costs elicited from the UK advisory board would be applied to all those undergoing subsequent treatment in the SoC and pembrolizumab arms.

- Given the considerations above about the unknown level to which pembrolizumab displaces treatments in the current pathway, we decided to handle this via a series of arbitrary sensitivity analyses rather than make explicit assumptions. This included an extreme assumption where the only treatment costs in the model were pembrolizumab costs i.e. it was treated as entirely cost-additive.

In the sub-population of KEYNOTE-204 who hadn't received an SCT prior to the trial, the proportion of patients receiving subsequent therapy in the pembrolizumab and BV arms was 50.8% and 69.1% respectively. These data were used to represent subsequent treatment use in the pembrolizumab and SoC arms in the model. The proportion receiving SCT was very similar between the two arms and has likely not influenced this differential.

The subsequent treatment composition is summarised in Table 58. As previously outlined in TA540[2], there remains considerable uncertainty over what constitutes standard subsequent treatment for R/R cHL in the UK given the paucity of data. This limitation in scoping subsequent treatment regimens continues at the time of this CDF exit. As such, the same approach was adopted in TA540 of initially basing the subsequent treatment composition on TA306,[99] which was the approach taken in TA462.[27] We presented these data, along with those collected in SACT to clinical experts at the UK Clinical Advisory Board who provided their feedback. Of note, for pembrolizumab and SoC, bendamustine and radiotherapy was added to the composition and RVIG was removed. In the SoC arm only, nivolumab was included as a subsequent treatment based on NICE TA462.[27]

For patients who would receive subsequent therapy, the clinical experts were not able to estimate the proportions receiving specific treatments with confidence with the exception that they agreed 70% of treated patients would receive bendamustine, following either pembrolizumab or SoC arm. For pembrolizumab, the distribution of patients across the other therapies was estimated by taking the remaining proportion of 0.30 and dividing this equally among the number of treatments. As nivolumab was included for the SoC arm, only a small percentage of patients in the SoC arm would get an autoSCT and relapse (~8% [SCT] * 35% [autos] * 63% [relapses]) and not all

would likely be fit enough for nivolumab at this time, we assigned 1% of SoC patients to receive this regimen. The remaining treatments were reweighted to account for nivolumab.

Table 58. Composition of subsequent therapy among patients who receive active treatment

Therapy	Distribution of patients across therapies (%)	
	Pembrolizumab	SoC
Gemcitabine monotherapy (administered over 4 weeks)	5%	4.83%
DHAP	5%	4.83%
CHOP	5%	4.83%
IVAC	5%	4.83%
Weekly therapy (PMitCEBO)	5%	4.83%
Bendamustine	70%	70%
Radiotherapy	5%	4.83%
Nivolumab	N/A	1%
Abbreviations: CHOP: cyclophosphamide, doxorubicin, prednisolone, vincristine; DHAP: dexamethasone, cytarabine, cisplatin; IVAC: cytrabine, etoposide, ifosfamide, mesna; PMitCEBO: bleomycin, cyclophosphamide, etoposide, mitoxantrone, prednisolone, vincristine		

The dosages, cycles, acquisition unit costs, administration cost and expected treatment duration for each regimen included for subsequent therapy are summarised in Table 59, Table 60, Table 61 and Table 62 respectively. Radiotherapy is costed separately as outlined in section B.3.5.4.

Table 59. Dosing and descriptions for subsequent therapy

Regimen	Dosage (mg)	Dosage unit	Admins per cycle	Cycle length (days)	Max. cycles
Gemcitabine (monotherapy)	1000	mg/m^2	3	28	6
Bendamustine (monotherapy)	120	mg/m^2	2	28	6

DHAP	Dexamethasone	40	mg	4	21	2
	Cytarabine	2000	mg/m^2	1		
	Cisplatin	100	mg/m^2	1		
CHOP	Cyclophosphamide	750	mg/m^2	1	21	8
	Doxorubicin	50	mg/m^2	1		
	Vincristine	2	mg	1		
	Prednisolone	100	mg	5		
IVAC	Etoposide	60	mg/m^2	5	21	6
	Cytarabine	2000	mg/m^2	4 (twice on 2 days)		
	Ifosfamide	1500	mg/m^2	5		
	Mesna	300	mg/m^2	10 (twice on 5 days)		
PMitCEBO	Bleomycin	10	mg/m^2	1 (day 8)	14	Although PMitCEBO is otherwise known as 'weekly therapy', the cycle length is 14 days with a maximum of 8 cycles (16 weeks)
	Cyclophosphamide	300	mg/m^2	1		
	Etoposide	150	mg/m^2	1		
	Mitoxantrone	300mg/m²	mg/m^2	1		
	Prednisolone	50	mg	14 (each day)		
	Vincristine	1.4	mg/m^2	1		
Nivolumab		240	mg/m^2	1	14	

Table 60. Drug acquisition costs for subsequent therapy

Component	Strength	Units per pack	Pack cost	Source	Cost per unit
Bendamustine	25mg	5	£27.75	BNF (2023)[66]	£5.55

	100mg	5	£77.70	BNF (2023)[66]	£23.55
Gemcitabine	1000mg	1	£9.36	eMIT (2023)[65]	£9.36
	2000mg	1	£19.64		£19.64
Dexamethasone	0.5mg	30	£2.56	eMIT (2023)[65]	£0.09
	2mg	50	£2.46	eMIT (2023)[65]	£0.05
	4mg	50	£30.73		£0.61
Cytarabine	100mg	5	£12.84	eMIT (2023)[65]	£2.57
	500mg	5	£15.13		£3.03
	1000mg	1	£8.28		£8.28
	2000mg	1	£12.66		£12.66
Cisplatin	10mg	1	£2.71	eMIT (2023)[65]	£2.71
	50mg	1	£9.10		£9.10
	100mg	1	£10.97		£10.97
Cyclophosphamide	500mg	1	£8.43	eMIT (2023)[65]	£8.43
	1000mg	1	£12.23		£12.23
	2000mg	1	£27.50		£27.50
Doxorubicin	10mg	1	£4.52	eMIT (2023)[65]	£4.52
	50mg	1	£7.29		£7.29
Prednisolone	1mg	28	£0.20	eMIT (2023)[65]	£0.01
	5mg	28	£1.23		£0.01
	25mg	56	£12.41		£0.22
Vincristine	1mg	1	£7.10	eMIT (2023)[65]	£7.10
	2mg	1	£16.76		£16.76
Etoposide	100mg	1	£3.94	eMIT (2023)[65]	£3.94
	500mg	1	£14.79		£14.79
Mesna	400mg	15	£203.24	BNF (2023)[66]	£13.41
	1000mg	15	£442.75		£29.51
Ifosfamide	1000mg	1	£115.79	BNF (2023)[66]	£115.79
	2000mg	1	£273.77		£273.77

Bleomycin	15mg (15,000 unit)	10	£190.60	BNF (2023)[66]	£19.06
Mitoxantrone	20mg	1	£67.24	eMIT (2023)[65]	£67.24
Nivolumab	240mg	1	£2,633.00	BNF (2023)	£2,633.00
Abbreviations: BNF, British National Formulary; eMIT, electronic market information tool.					

Table 61. Administration cost of subsequent therapy (per cycle)

Regimen	Cost	Description
Bendamustine	£573.42	Delivering two elements of simple parenteral chemotherapy.
Gemcitabine	£860.13	Delivering three elements of simple parenteral chemotherapy.
DHAP	£474.94	Delivering complex chemotherapy at first attendance.
CHOP	£474.94	Delivering complex chemotherapy at first attendance.
IVAC	£1,948.70	Delivering complex chemotherapy at first attendance and delivering four subsequent elements of a chemotherapy cycle.
PMitCEBO	£843.38	Delivering complex chemotherapy at first attendance and delivering a subsequent element of a chemotherapy cycle.
Nivolumab	£286.71	Delivering simple parenteral chemotherapy at first attendance.
Where delivering a simple parenteral chemotherapy at the first attendance is £286.71 (SB12Z), delivering complex chemotherapy at first attendance is £474.94 (SB14Z), and delivering a subsequent element of a chemotherapy cycle is £368.44 (SB15Z). **Abbreviations:** CHOP: cyclophosphamide, doxorubicin, prednisolone, vincristine; DHAP: dexamethasone, cytarabine, cisplatin; IVAC: cytrabine, etoposide, ifosfamide, mesna; PMitCEBO: bleomycin, cyclophosphamide, etoposide, mitoxantrone, prednisolone, vincristine.		

The acquisition and administration cost per cycles for each component of the subsequent treatment regimens were multiplied by the expected duration and usage to give a one-off weighted average cost of £1,624.51 for the pembrolizumab arm and £2,230.43 for the SoC arm, as presented in Table 62. Subsequent treatment costs for both arms are calculated in the model by multiplying the weighted average subsequent treatment cost by the number if newly discontinued patients at each cycle.

Table 62. Weighted total subsequent treatment costs for pembrolizumab and SoC

Therapy	Treatment duration (cycles)*	Cost per cycle	Cost for treatment duration	Treatment distribution (%)	
				Pembro	SoC
Bendamustine	2.0	£642.42	£1,284.84	35.56	48.37
Gemcitabine monotherapy	4.0	£912.08	£3,648.31	2.64	3.34
DHAP	2.0	£521.56	£1,043.13	2.54	3.34
CHOP	6.0	£524.55	£3,147.33	2.54	3.34
IVAC	3.5	£3,825.09	£13,387.82	2.54	3.34
PMitCEBO	7.0	£2,771.74	£19,402.16	2.54	3.34
Radiotherapy	1.0	£5,340.59	£5,340.59	2.54	3.34
Nivolumab	36.5	£2,919.51	£88,972.07	0.00	0.07
No active treatment	N/A	£0.00	£0.00	49.20	31.54
Weighted total subsequent treatment cost				**£1,624.51**	**£2,230.43**

Abbreviations: CHOP: cyclophosphamide, doxorubicin, prednisolone, vincristine; DHAP: dexamethasone, cytarabine, cisplatin; IVAC: cytrabine, etoposide, ifosfamide, mesna; PMitCEBO: bleomycin, cyclophosphamide, etoposide, mitoxantrone, prednisolone, vincristine; SoC: Standard of Care.

***Source:** TA540,[2] Royal College of Radiologists guidelines for HL[111], Ansell et al. 2021[112]

B.3.6.6.2 Terminal care cost

To account for intensive disease management in the months leading to a death, a one-off terminal care cost was applied to all patients the pembrolizumab or SoC arm

when entering the death state. A one-off terminal care cost of £8,752.32 was applied to all patients who died prior to the landmark and to all patients in the No/Failed SCT state after the landmark, with a lower cost of £7,224.46 to patients in the curative SCT state after the landmark. Costs were sourced from a research report by the Nuffield Trust[113] exploring costs at the end of life, and relate to the terminal care costs for "all cancer patients" and "all patients irrespective of diagnosis" respectively. All costs were inflated into the 2021/22 cost year using the PSSRU index.[63] The breakdown of the total cost is summarised in Table 63.

Table 63. Terminal care costs

Resource use	Unit cost		Source
	No/Failed SCT	Curative SCT	
Secondary (acute) hospital care	£7,074.33	£5,500.92	Georgiou & Bardsley (2014)[113], inflated to 2021/22 prices
Local authority-funded social care	£533.28	£1,213.08	
District nursing	£706.23	£333.90	
GP contacts	£438.39	£176.56	
Total	**£8,752.23**	**£7,224.46**	

B.3.6.6.3 Adverse reaction unit costs and resource use

The cost per adverse event is summarised in Table 64. These adverse event unit costs were obtained from NHS reference costs using HRG codes from various NICE appraisals consistent with the TA540 appraisal. For fatigue and pyrexia, the HRG codes from TA540 were no longer in use and therefore alternative appraisals (TA772,[1] TA813[114]) were used to inform the AE HRG codes. The weighted unit cost was estimated for each adverse event by calculating the total cost per HRG code and dividing by the activity in that respective code.

Table 64. Adverse reaction unit costs

Adverse Event	Unit Cost	Source	Rationale

Anaemia	£941	NHS Reference Costs 2021/22. Weighted average of HRG Codes: SA03G-H: Haemolytic Anaemia with CC Score 0-3+, total; SA04G-L: Iron Deficiency Anaemia with CC Score 0-14+, total; SA05G-J: Megaloblastic Anaemia with CC Score 0-8+, total; SA08G-J: Other Haematological or Splenic Disorders, with CC Score 0-6+, total.	TA411,[103] TA399,[115] and TA391[116] cited in TA540[2]
Diarrhoea	£1,847	NHS Reference Costs 2021/22. Weighted average of HRG Codes: FD04A-B: Nutritional Disorders with Interventions, with CC Score 0-2+, total; FD04C-E: Nutritional Disorders without Interventions, with CC Score 0-6+, total; FD10A-D: Non-Malignant Gastrointestinal Tract Disorders with Multiple Interventions, with CC Score 0-8+, total; FD10E-H: Non-Malignant Gastrointestinal Tract Disorders with Single Intervention, with CC Score 0-9+, total; FD10J-M: Non-Malignant Gastrointestinal Tract Disorders without Interventions, with CC Score 0-11+, total.	TA399[115] and TA440[102] cited in TA540[2]
Dyspnea	£863	NHS Reference Costs 2021/22. Weighted average of HRG Codes: DZ19H: Other Respiratory Disorders with Multiple Interventions, total; DZ19J-K: Other Respiratory Disorders with Single Intervention, with CC Score 0-5+, total; DZ19L-N: Other Respiratory Disorders without Interventions, with CC Score 0-11+, total.	TA420[98] cited in TA540[2]
Fatigue	£2,015	NHS Reference Costs 2021/22. Weighted average of HRF codes SA01G, H, J, K: Acquired pure red cell aplasia or other aplastic anaemia, non-elective short stay. Assumed equal to fatigue (Brown et al., 2013)	TA772[1]

Leukopenia	£1,366	NHS Reference Costs 2021/22. Weighted average of HRG codes SA08G-J: Other Haematological or Splenic Disorders, with CC Score 0-6+, total.	TA391[116] cited in TA540[2]
Nausea	£1,030	NHS Reference Costs 2021/22. HRG code FF53A: Minor Therapeutic or Diagnostic, General Abdominal Procedures, 19 years and over, total.	TA411[103] cited in TA540[2]
Neutropenia	£1,366	NHS Reference Costs 2021/22. Weighted average of HRG codes SA08G-J: Other Haematological or Splenic Disorders, with CC Score 0-6+, total. Assumed equal to leukopenia.	TA411[103] and TA399[115] cited in TA540[2]
Pyrexia	£1,322	NHS Reference Costs 2021/22. Weighted average of HRG codes WJ07A-D: Fever of unknown origin, with and without interventions, total.	TA813[114]
Thrombocytopenia	£993	NHS Reference Costs 2021/22. Weighted average of HRG codes SA12G-K: Thrombocytopenia with CC Score 0-8+, total.	TA399[115] and TA440[102] cited in TA540[2]
Vomiting	£1,847	Assumed equal to Diarrhoea. NHS Reference Costs 2021/22. Weighted average of HRG Codes: FD04A-B: Nutritional Disorders with Interventions, with CC Score 0-2+, total; FD04C-E: Nutritional Disorders without Interventions, with CC Score 0-6+, total; FD10A-D: Non-Malignant Gastrointestinal Tract Disorders with Multiple Interventions, with CC Score 0-8+, total; FD10E-H: Non-Malignant Gastrointestinal Tract Disorders with Single Intervention, with CC Score 0-9+, total; FD10J-M: Non-Malignant Gastrointestinal Tract Disorders without Interventions, with CC Score 0-11+, total.	TA476[100] and TA440[102] cited in TA540[2]

B.3.6.7 Autologous and allogeneic stem cell transplantation

The cost of alloSCT in TA540 was taken from Radford et al 2015[21] as this was the preferred source by the committee in TA462.[27] Radford et al was a retrospective analysis that studied the cost and resource use in 40 cHL patients who had relapsed after autoSCT, of which 15 patients subsequently received alloSCT and were followed up to date of death or most recent follow-up. As mentioned for this CDF review, the population of interest is Cohort 2 from KEYNOTE-087, patients who have received BV and are ineligible for autoSCT and therefore the Radford study is not directly applicable to the population of interest.

Because of this, we followed approaches taken in recent haematology appraisals (e.g. TA567,[117] TA813[114]) and adopted a micro-costing approach to establish the cost of both autoSCT and alloSCT by splitting the process into three key components: stem cell harvesting; transplant procedure; and follow-up. The cost of stem cell harvest and the transplant procedure were taken from the NHS Reference Costs (2021/2022),[64] presented in Table 65. The procedure cost of alloSCT was calculated using a weighted average of the total cost for sibling, volunteer unrelated, and donor type not specified reference costs. The 24-month follow-up cost for alloSCT was taken from the UK Stem Cell Strategy Oversight Committee report (2014)[118] and was inflated to the 2021/2022 cost year using the PSSRU index,[63] HCHS index up to 2015 and NHSCII index for 2015 onwards. This was the chosen and accepted source of alloSCT costs in recent haematology appraisals (TA567,[117] TA813[114]). The follow-up cost for autoSCT was not presented in this report, and therefore was calculated as a proportion of the follow-up costs for alloSCT using the relative costs from Blommestein et al. 2021.[119] The final cost for autoSCT and alloSCT are presented in Table 66. These costs were validated by clinicians at an advisory board to be reflective of clinical practice, and 24 months follow-up was considered sufficient to capture all transplant-related follow-up costs.

Allogeneic SCT is notably more expensive than autologous SCT. The experts at the UK advisory board commented that this is likely because of significant issues with graft vs. host disease, lengthy hospital stays and intensive follow-up after the procedure. AlloSCT is the primary goal of treatment in this setting but it is not certain that it would have a cost-effective ICER if assessed in isolation.

Table 65. NHS Reference Costs (2021/22) for stem cell harvesting and transplant.

HRG tariff	Description	Activity	Unit cost
SA34Z	Peripheral Blood Stem Cell Harvest	3,103	£1,651.11
SA26A	Peripheral Blood Stem Cell Transplant, Autologous, 19 years and over	2,069	£17,898.16
SA38A	Peripheral Blood Stem Cell Transplant, Allogeneic (Sibling), 19 years and over	289	£32,148.73
SA39A	Peripheral Blood Stem Cell Transplant, Allogeneic (Volunteer Unrelated Donor), 19 years and over	478	£36,235.96
SA40Z	Peripheral Blood Stem Cell Transplant, Allogeneic (Donor Type Not Specified)	268	£49,132.68
SA38A, SA39A, SA40Z	Peripheral Blood Stem Cell Transplant, Allogeneic (weighted average by activity)	1,035	£38,434

Table 66. Total cost of autoSCT and alloSCT

	Stem cell harvest	Transplant procedure	Follow-up		Total Cost
			Year 1	Year 2	
AutoSCT	£1,651[a]	£21,150[b]	£8,152[d]	£1,019[d]	£31,972[f]
AlloSCT	£1,651[a]	£38,434[c]	£40,383[e]	£5,047[e]	£85,515[f]

[a] NHS National Reference Costs 2021/22. SA34Z.(64)

[b] NHS National Reference Costs 2021/22. SA26A.(64)

[c] NHS National Reference Costs 2021/22. Weighted average of SA38A, SA39A and SA40Z.(64)

[d] Calculated as a proportion alloSCT follow-up cost based on the relative costs from Blommestein et al. (2012).(119)

[e] NHS Blood and Transplant, Stem Cell Transplant Oversight Committee Report on Unrelated donor stem cell transplantation in the UK (2014),(118) inflated to 2021/22 using PSSRU index.

[f] Aggregated cost from stem cell harvest, transplant procedure, and follow-up costs.

The proportion of SCTs that were auto or allo (35%/65%) was taken from the SACT report cost of Allo SCT was applied outside the model engine to all patients who received an SCT in both arms. We inspected the Time to SCT curves from SACT and noted that ~90% of events occurred in the first year, with the remainder occurring soon afterwards. As such, we didn't implement discounting for these costs, reasoning that it wouldn't make any material difference to the model's results as total SCT costs would reduce by only one tenth of the discount rate, or 0.35%.

B.3.7. Severity

As stated in Section B.1.3, although cHL is rare, the disease at relapsed or refractory stage is a severe disease associated with significant morbidity and mortality. It often occurs in young patients who would otherwise have long lives ahead of them. In addition, older patients (>60 years) with advanced stage R/R cHL are more likely to die from non-lymphoma causes such as bleomycin-related lung toxicity or receiving chemotherapy regimens such as ABVD and BEACOPP.[24] If left untreated, R/R cHL can spread to multiple lymph nodes. At advanced stage, cHL can spread to extra nodal sites such as the bone marrow and spleen, and to organs outside the lymphatic system, such as the liver or lungs.[24]

There continues to be significant unmet need for treatment options for patients with R/R cHL, who receive BV but are ineligible for autoSCT. Before pembrolizumab was introduced in the clinical pathway, treatment options were limited and outcomes were typically poor. Clinicians from the MSD UK Advisory Board stated among patients who received prior BV and are autoSCT ineligible, pembrolizumab would be the preferred treatment option to help achieve a better or durable response to bridge them to SCT.[5]

Under NICE's Severity Modifier, the Appraisal Committee may apply a greater weight to QALY gains if the technology is indicated for a condition with a high degree of severity (as determined based on proportional and/or absolute QALY shortfall). To understand the extent to which R/R cHL deprives patients of their remaining QALYs, the total lifetime accrued QALY of patients receiving SoC (as estimated in the cost-effectiveness model) is measured as a proportion of the total lifetime QALY gain of the general population with the same age and sex distribution.[54] The total QALYs

associated with SoC were obtained from the results of the base-case analysis, and we estimated total QALYs for the general population using the starting age, the proportion female taken from the SACT report and the ScHARR QALY Shortfall calculator tool.[120]

The results of the QALY shortfall analysis (Table 67) indicate that the general population with this mean age and sex would expect to accrue 15.6 QALYs. This means that as long as expected QALYs on SoC are less than 3.6, this indication has a QALY shortfall >12 and qualifies for the 1.2 QALY modifier. If the patients on the SoC have an expected QALYs of <0.78, this indication would qualify for the 1.7 modifier.

Table 67. Summary features of QALY shortfall analysis

Factor	Value (reference to appropriate table or figure in submission)	Reference to section in submission
Sex distribution	40% female	Patient population
Starting age (mean)	51 years	Patient population
Abbreviation: QALY, quality adjusted life year.		

Table 68. Summary of QALY shortfall analysis

Expected total QALYs for the general population	Total QALYs that people living with a condition would be expected to have with current treatment	QALY shortfall
15.6	1.31	14.29
15.6	2.46 (highest sensitivity analysis)	13.14

In all scenario analyses, QALY shortfall was between 12 and 18. There were no scenarios where QALY shortfall was greater than 95%. The 1.2 severity modifier is therefore indicated and will be applied to all incremental QALY gains when calculating ICERs.

B.3.8. Uncertainty

A variety of factors impact the ability to generate high quality evidence in the R/R cHL population. These include the small patient population, lack of randomised

controlled trials and limited data on patient outcomes on standard of care. This uncertainty is substantially offset by the magnitude of the difference between what has been observed in patients treated with pembrolizumab and what is expected on SoC; most importantly 4-year OS of ~55% vs. 10-20% and proportion achieving SCT of 30% vs. <10% as well as significant improvements in HRQoL. In addition, real world evidence has been collected which directly addresses the NICE committee's uncertainty following TA540.

B.3.9. Summary of base-case analysis assumptions

B.3.9.1 Assumptions

Table 69. Base-case assumptions

Variable	Assumption for base case analysis	Justification
Model structure	2-states (alive/dead) prior to the landmark and 3 states (cured-SCT, no/failed SCT and dead) beyond it	No PFS as collected in the SACT. Landmark set at 4 years as adequately captures the point at which SCTs and SCT-associated relapses take place.
Model time horizon (years)	40 years	Sufficient to capture all relevant and important differences in the future costs or outcomes among the treatments.
Perspective	NHS and PSS	NICE reference case.
Discount rate: Costs and outcomes	3.5%	NICE reference case.
Pre-landmark OS - pembrolizumab	SACT: total population (one-piece parametric curves)	SACT dataset is much larger (N=215 vs N=81 in KEYNOTE-087) and represents how pembrolizumab is used in the UK NHS.

Pre-landmark OS – SOC	HR derived from Bucher ITC using KEYNOTE-204 and TA524	A formal ITC could not be conducted between SACT vs Cheah and SACT vs Eyre et al. as no access to IPD or baseline characteristics from these studies. KEYNOTE-204 contained subgroup who did not have an SCT. Markov trace from TA524 appraisal was used as BV was the common comparator where the treatment effect of pembrolizumab vs SoC were assumed to be the sum of the treatment effects for pembrolizumab vs BV and BV vs SoC.
SCT outcomes - Probability of SCT Pembrolizumab	SACT: total population	SACT dataset represents proportion of patients on pembrolizumab that get a SCT in the NHS.
SCT outcomes - Probability of SCT SoC	Structured expert elicitation	As probability of receiving an SCT whilst on SoC was not collected in the SACT dataset, an SEE was run to obtain robust probability estimate for SoC.
Ratio of auto and alloSCT – pembrolizumab	SACT: total population	SACT dataset contains the proportion of patients receiving auto and alloSCT in the UK NHS whilst on pembrolizumab.
Ratio of auto and alloSCT – SoC	Assumed equal to pembrolizumab	Absence of evidence on proportion receiving autoSCT and alloSCT whilst on SoC.
Probability of curative SCT – pembrolizumab	Structured expert elicitation	As no studies on treatment effect for pembrolizumab versus SoC was identified in the SLR, SEE was run to obtain robust cure estimates for pembrolizumab.
Probability of curative SCT – SoC	Structured expert elicitation	As no studies on treatment effect for pembrolizumab versus SoC was identified in the SLR, SEE was run to obtain robust cure estimates for SoC.

Post-landmark OS: Cured SCT patients pembrolizumab and SoC	No SMR applied	Clinicians from the Advisory Board were not able to suggest any studies on magnitude of any mortality decrement.
Post-landmark OS: Cured SCT patients – pembrolizumab and SoC	General population mortality	SCT cured patients assigned a general population QoL in line with previous R/R cHL appraisal (TA772).
ToT: pembrolizumab	KM ToT data from SACT	SACT is the most relevant source as pembrolizumab is a bridge to SCT and the patient mix is more reflective of UK practice.
ToT: SoC	From various studies	To align with the recommended treatment duration
Treatment stopping rules – pembrolizumab	No additional rule applied in the economic model. ToT informed by SACT. A maximum of 35 treatment cycles (approximately 2 years) applied in the model.	To align with the SACT data and NHS clinical practice.
Treatment stopping rules – SoC	Maximum number of treatment durations	As per studies/trial.
Treatment waning	Not applied	No evidence to indicate a treatment waning.
Utilities – pembrolizumab and SoC	KEYNOTE-204 (EQ-5D-3L)	KEYNOTE-204 captures utilities for both pembrolizumab and BV Trial uses EQ-5D-3L and provides the only comparative source of evidence on pembrolizumab versus a chemotherapy-based regimen. Mean EQ-5Ds between KEYNOTE-087 and KEYNOTE-204 were also generalisable.
AE type and incidence – pembrolizumab	AEs were based on grade 3-5 AEs from KEYNOTE-087. Relevant AEs were based on TA462 and validated by clinical experts.	As per TA540 methodology.

AE type and incidence - SoC	Drug-related AEs based on various literature	As KEYNOTE-087 is a single arm trial, AEs had to be sourced from the literature.
Disutilities – AEs	Various literature	As no disutilities were collected in KEYNOTE-087 or in any R/R HL study, disutilities were sourced from the literature (leukaemia, lung, breast, soft tissue carcinoma and pancreatic cancer).
AE durations	NICE appraisals (TA306 and TA476).	As no disutility durations were collected in KEYNOTE-087 or in any R/R HL study, disutility durations were sourced from NHL and pancreatic cancer appraisals.
SCT complications (QALY decrement)	Absolute QALY decrement calculated by the decrements for each time point by the time point (in years) and a disbenefit was capped at 2 years. This absolute QALY decrement was applied at cycle 0 to all patients that undergo SCT by the landmark timepoint.	Auto/alloSCT split and acute and chronic GvHD could not be explicitly accounted for as followed R/R cHL appraisal (TA524). A 2-year disbenefit was the point at which all SCT-related events of interest have taken place as validated from the Clinical Advisory board.
Pembrolizumab administration	Q3W, as per the SACT and KEYNOTE-087 trial.	To align with the SACT and clinical trial.
SoC comparator composition	SoC regimens assumed equal proportion.	As clinicians in the Advisory Board could not provide estimates for proportion receiving each SoC regimen.
SoC administration	Sourced from TA540, Collins et al. (2014), Northern Cancer Alliance and Lymphoma Group.	As per study.
Drug wastage costs	Vial sharing is applied with 0% drug wastage assumed	

Drug acquisition costs	BNF, eMIT. Dosages for pembrolizumab in line with KEYNOTE-087 and for SoC based on published literature/studies.	NICE's preferred source for costs. To align with sources.
Drug administration costs	Costs sourced from the National Schedule of NHS Costs.	NICE's preferred source for costs. To align with TA540 methodology and previous appraisals.
Radiotherapy cost	Line items based on NG122 in NSCLC. Costs sourced from the National Schedule of NHS Costs. Number of fractions and admission reduced to 5 in line with RCR for R/R cHL. Weighted average calculated and applied as a one time-cost in the model.	Clinical experts from the Advisory Board stated radiotherapy should be added to SoC. NHS PSSRU is NICE's preferred source for costs.
Costs and resource use pre-landmark no SCT or relapsed SCT	Outpatient attendance, biochemistry, cell blood count and PET scan costs sourced from the National Schedule of NHS Costs. Weekly usage pre landmark for all patients and post landmark no/relapsed SCT based on TA540 (originally TA524) and subsequently validated by clinical experts. Cured SCT patients assumed no health state resource use based on clinical experts.	Health state cost and resource use varies between patient groups i.e. cured SCT versus no/relapsed SCT.
Proportion receiving any subsequent treatment	Pembrolizumab from KEYNOTE-204, with the BV arm representing SoC.	As per KEYNOTE-204 trial.

Subsequent treatment Costs.	Those who are undergoing subsequent treatment in the SoC and pembrolizumab arm will receive subsequent treatments elicited from the Clinical Advisory Board. Subsequent therapies for SoC adjusted for nivolumab.	Proportions based on the clinical advisory board based on paucity of good evidence.
Subsequent treatment Costs – nivolumab SoC only	Nivolumab included in the subsequent treatments for SoC.	Nivolumab was included as a subsequent treatment based on NICE TA462.
Subsequent therapy dosing and cycle	Dosing and cycles consistent with the methodology in TA540.	Dosing and cycles are aligned with the literature.
Subsequent therapy drug acquisition costs	Cost sources based on BNF, eMIT. Dosages SoC based on published literature/studies.	NICE's preferred source for costs.
Subsequent therapy administration cost per cycle	Costs sourced from the National Schedule of NHS Costs.	NICE's preferred source for costs. To align with TA540 methodology and previous appraisals.
Terminal care costs	One-time terminal care cost applied to all patients who died before the landmark and to the no/failed SCT state after the landmark. A lower cost applied to the curative SCT group after the landmark.	We wanted to account for intensive disease management costs in the months leading to a death.
AE costs	List of AEs from various NICE appraisals. AE unit costs sourced from the National Schedule of NHS Costs.	To align with TA540 methodology and previous appraisals. NICE's preferred source for costs.

Cost of auto and alloSCT	Cost of autoSCT and alloSCT was microcosted as population in Radford (2015) concerned relapsed patients after autoSCT of which a proportion received alloSCT. Components sourced from previous NICE haematology appraisals (TA567, TA813). Harvesting and procedure costs were sourced from National Schedule of NHS Costs. Follow up costs were sourced from Blommestein et al. and NHS Blood and Transplant, report and inflated to current cost year.	To accurately capture cost of SCT for patients in cohort 2 i.e. , patients who have received BV and are ineligible for autoSCT. NHS reference costs are NICE preferred source for costs. To align with TA540 and previous haematology appraisals.
Abbreviations: AE: adverse event; alloSCT: allogenic stem cell transplant; autoSCT: autologous stem cell transplant, BNF: British National Formulary; BV: brentuximab vedotin; eMIT; electronic market information tool; NHS: National Health Service; OS; overall survival; Q3W: every 3 weeks; SACT: Systemic anti-cancer therapy; SoC: standard of care; ToT: time on treatment		

B.3.10. Base-case results

All results in this section are presented with the Commercial Access Arrangement price applied. All incremental QALY gains are multiplied by the 1.2 severity modifier. Life years are not discounted. Disaggregated results are available in Appendix O.

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

Table 70. Base-case results

Technologies	**Total costs (£)**	**Total LYG**	**Total QALYs**	**Inc costs (£)**	**Inc LYG**	**Inc QALYs** (inc sev mod)	**ICER (£/QALY)**
SoC	[redacted]	2.24	[redacted]				
Pembrolizumab	[redacted]	10.96	[redacted]	[redacted]	8.72	5.39	[redacted]

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

B.3.11. Exploring uncertainty

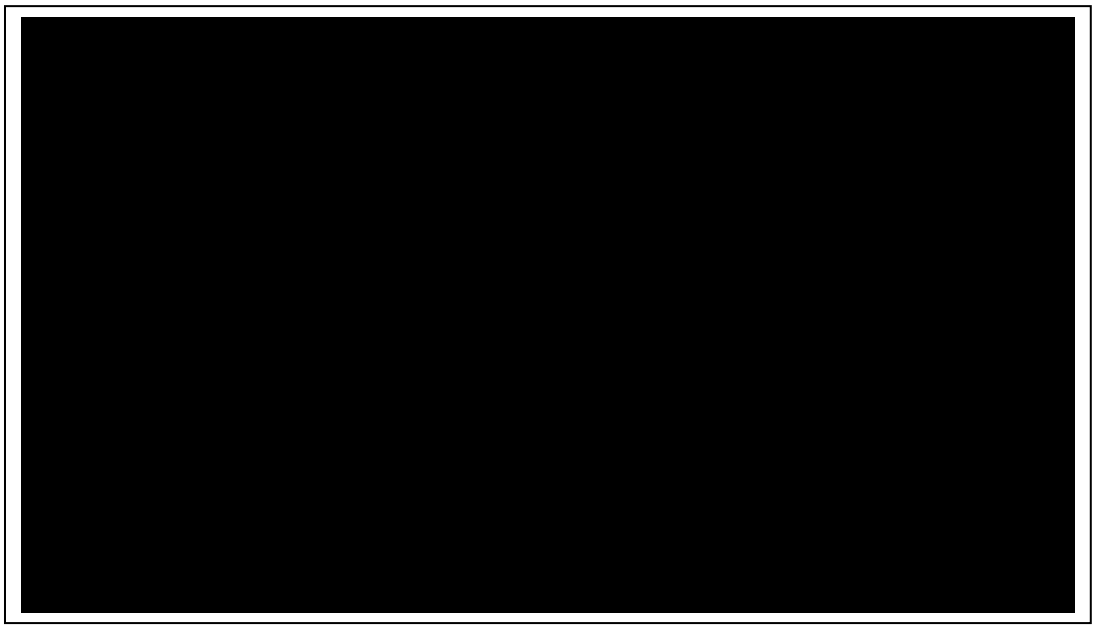
B.3.11.1 Probabilistic sensitivity analysis

We ran the model 1,000 times, varying uncertain parameters within their appropriate probability distributions. The ICER and all disaggregated results were found to be stable. Pembrolizumab had a greater than [redacted] probability of being cost-effective vs. a threshold of £30,000/QALY gained (including a 1.2 severity modifier to incremental QALYs).

Table 71. PSA Results

Technologies	Total costs (£)	Total LYG	Total QALYs	Inc costs (£)	Inc LYG	Inc QALYs (inc sev mod)	ICER (£/QALY)
SoC	[redacted]	2.23	[redacted]				
Pembrolizumab	[redacted]	10.93	[redacted]	[redacted]	8.62	5.32	[redacted]
Abbreviations: ICER, incremental cost-effectiveness ratio; LYG, life years gained; QALYs, quality-adjusted life years.							

Figure 21. PSA Scatterplot

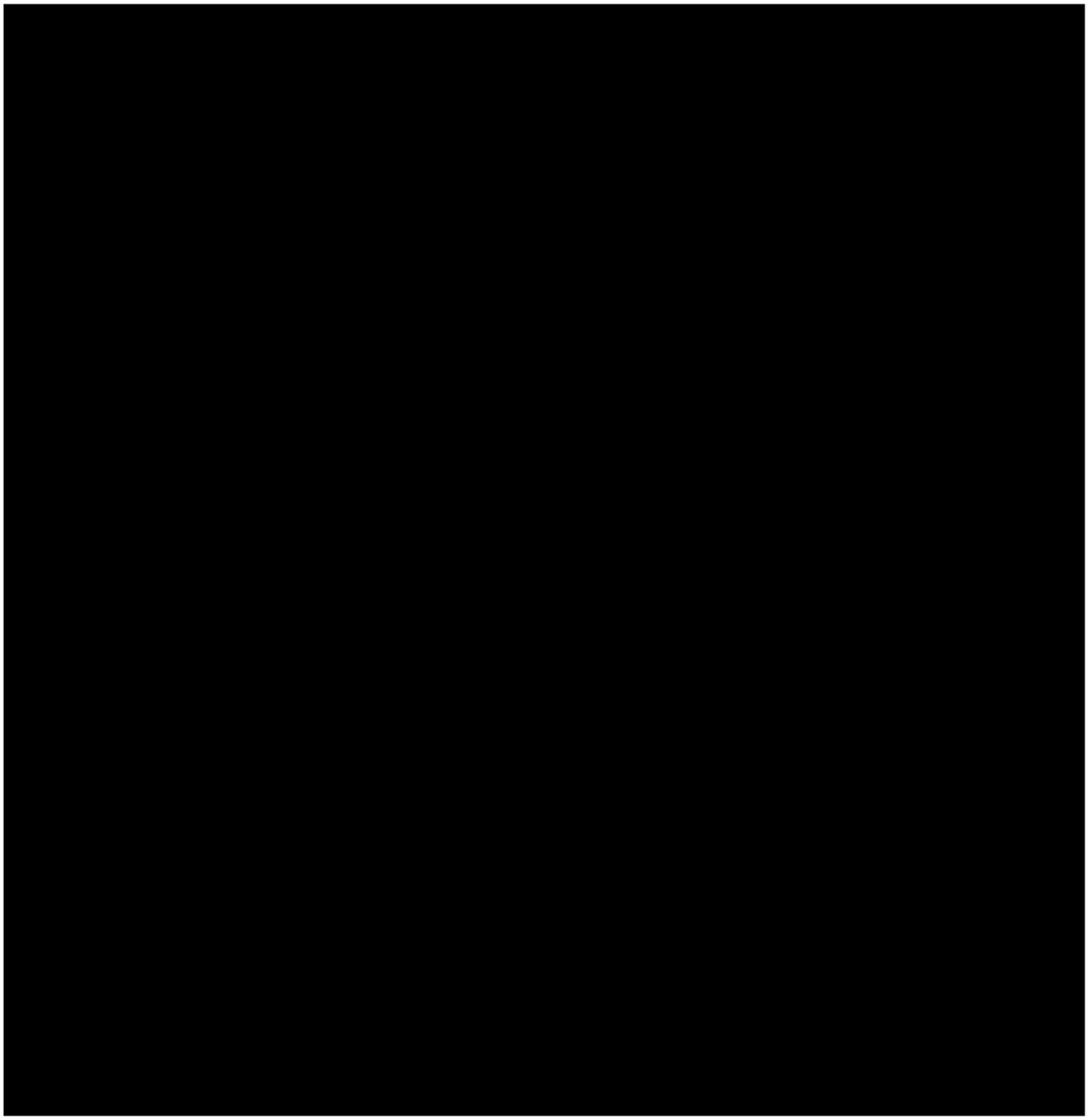


Key: PSA, probabilistic sensitivity analysis; QALY, quality-adjusted life-year; WTP, willingness-to-pay threshold

B.3.11.2 Deterministic sensitivity analysis

We undertook deterministic sensitivity analyses, varying key inputs to the extremes of their confidence intervals or, in the case of certain parameters with no associated probability distribution, to arbitrarily decided values e.g. the location of the landmark was varied to 3 and 5 years. The conclusions of the model were robust to all these changes, with all ICERs below [redacted].

Figure 22. Tornado diagram from deterministic sensitivity analyses



B.3.11.3 Scenario analysis

Table 72. Summary of scenario analyses

Base case	Scenario value	Justification
Model set up		
Baseline age of 51 years	KEYNOTE-087 starting age	To reflect the baseline age of the cohort from the KEYNOTE-087 trial.

Baseline age of 51 years	34 years	Among the mean age of the cohort in the SACT was 51 years, among these patients, the cohort who received SCT had a mean age of 34 years.
Time horizon 40 years	Lifetime (49 years)	Up to 100 years in line with NICE reference case.
3.5% discount rate	1.5% discount rate	The NICE manual recommends alternative analyses using rates of 1.5% for both costs and health effects.
Health-related quality-of-life		
Utility treatment effect for patients who never received SCT after landmark based on progressed disease from KN-204	Equal utility after the landmark	Pessimistic scenario regarding length of utility benefit for pembrolizumab.
Include age disutility	Exclude age disutility	To assess the impact of applying no age disutilities on the ICER
Include adverse event disutility	Exclude adverse event disutility	Adverse event disutilities may already be captured in the health state utility values.
Clinical effectiveness (survival)		
Landmark at 4 years	Landmark at 3 years	All SCTs and SCT-associated relapses may occur earlier.
Landmark at 4 years	Landmark at 5 years	Extreme scenario where all SCTs and SCT-associated relapses may occur later than the evidence suggested in the SACT dataset and AETHERA trial.
No treatment waning effect for mortality	3-5 years post pembrolizumab cessation for mortality treatment effects	3-5 years post cessation is in line with a NICE committee's most recent stated preferences (TA885).[72]

Survival after the landmark for the No/Failed SCT group - Log-logistic	Exponential	Exponential model had a similar visual fit to the extrapolated curves and had the lowest AIC.
No SMR applied	Applying a SMR by multiplying all cycle-specific OS event rates by 1.2	To account for survival of cured SCT patients not being equal to the general population.
Bucher ITC (KEYNOTE-204 and TA524)	KEYNOTE-204 (BV) 3L+ patients who did not receive prior SCT HR	As KEYNOTE-087 is a single arm trial, an alternative from the KEYNOTE-204 trial was used to inform the clinical outcomes of patients who did not receive prior SCT.
Bucher ITC (KEYNOTE-204 and TA524)	Bucher ITC vs Eyre et al. and TA524	As KEYNOTE-087 is a single arm trial, alternative ITC was conducted to inform the clinical outcomes of patients in the SoC arm of the economic model.
Bucher ITC (KEYNOTE-204 and TA524)	ITC vs Eyre et al.	As KEYNOTE-087 is a single arm trial, alternative ITC was conducted to inform the clinical outcomes of patients in the SoC arm of the economic model.
Bucher ITC (KEYNOTE-204 and TA524)	ITC vs Cheah et al.	As KEYNOTE-087 is a single arm trial, alternative ITC was conducted to inform the clinical outcomes of patients in the SoC arm of the economic model.
Bucher ITC (KEYNOTE-204 and TA524)	MAIC KEYNOTE-087 vs Eyre et al.	As KEYNOTE-087 is a single arm trial, a MAIC was incorporated to reweight the IPD from KEYNOTE-087 versus the SoC arm from Eyre et al.
Bucher ITC (KEYNOTE-204 and TA524)	MAIC KEYNOTE-087 vs Cheah et al.	As KEYNOTE-087 is a single arm trial, a MAIC was incorporated to reweight the IPD from KEYNOTE-087 versus the SoC arm from Cheah et al.
Resource use and costs		

Pembrolizumab 200mg Q3W dosing schedule as per the trial	Pembrolizumab 400mg Q6W dosing schedule	Q6W schedule is more commonly used in NHS practice to reduce the burden for patients and clinic capacity.
SoC distribution across all comparators in the SoC composition	SoC 100% bendamustine	Alternative SoC composition as no estimates were provided by the clinicians from the Advisory Board
Maximum treatment cycles for SoC - 90% (based on time on treatment curve from TA524)	SoC costs halved	Scenario explored as the BV ToT data from KN204 might have been an overestimate.
Include SoC and subsequent therapy costs	Only pembrolizumab costs applied to the treatment costs in the model	Extreme scenario where pembrolizumab is entirely additive to the treatment pathway and displaces nothing.
Differential SCT probabilities	Pembro SCT probability set equal to SoC	Examine influence of this treatment effect
Differential SCT cure probabilities	Pembro SCT cure probability set equal to SoC	Examine influence of this treatment effect
Allo SCT 65%/Auto 35%	Auto 65%/Allo 35%	Examine influence of auto/allo probabilities
Base case	OS Bucher Eyre+TA524 + 100% bendamustine SoC + treatment waning 3-5 + exp curve for No/Failed + equal utility after landmark + SMR 1.2	Conservative scenario
Abbreviations: AIC: Akaike information criterion; BIC: Bayesian information criterion; ICER: incremental cost effectiveness ratio; SCT: stem cell transplant; SMR: standardized mortality ratio; SoC: standard of care		

Table 73. Results of Scenario Analyses

Scenario	LYs SoC	LYs Pem	Inc Costs	Inc QALYs	ICER
Base Case	2.24	10.96	[redacted]	5.39	[redacted]
KEYNOTE-087 Starting Age	2.26	11.07	[redacted]	5.44	[redacted]
Starting Age = SCT patients	2.43	12.06	[redacted]	5.90	[redacted]

Time horizon 49 years	2.27	11.22	[redacted]	5.43	[redacted]
1.5% Discount Rate	2.24	10.96	[redacted]	6.75	[redacted]
Equal utility after the landmark	2.24	10.96	[redacted]	5.02	[redacted]
Exclude age related disutility	2.24	10.96	[redacted]	5.72	[redacted]
Exclude AE disutility	2.24	10.96	[redacted]	5.39	[redacted]
Landmark 3 Years	2.25	10.87	[redacted]	5.32	[redacted]
Landmark 5 years	2.21	11.00	[redacted]	5.44	[redacted]
Treatment waning effect on No/Failed SCT OS	2.24	8.74	[redacted]	4.38	[redacted]
Exponential survival curve after landmark	2.16	8.41	[redacted]	4.22	[redacted]
SMR 1.2 applied to Cured SCT	2.20	10.78	[redacted]	5.34	[redacted]
OS HR from KN204	3.17	10.96	[redacted]	4.71	[redacted]
OS HR Bucher Eyre+TA524	3.11	10.96	[redacted]	4.75	[redacted]
OR HR from Eyre ITC	4.14	10.96	[redacted]	4.01	[redacted]
OS HR from Cheah ITC	3.90	10.96	[redacted]	4.18	[redacted]
OS HR from MAIC 087 vs Eyre	2.04	10.96	[redacted]	5.54	[redacted]
OS HR from MAIC 087 vs Cheah	2.10	10.96	[redacted]	5.49	[redacted]
Pembrolizumab Q6W	2.24	10.96	[redacted]	5.39	[redacted]
SoC = 100% bendamustine	2.24	10.96	[redacted]	5.39	[redacted]
SoC treatment costs halved	2.24	10.96	[redacted]	5.39	[redacted]
All SoC and subs trt costs removed	2.24	10.96	[redacted]	5.39	[redacted]
SCT probabilities set equal	2.24	8.96	[redacted]	4.44	[redacted]
Probability SCT curative set equal	2.24	10.15	[redacted]	4.99	[redacted]
Auto/allo % reversed	2.24	10.96	[redacted]	5.39	[redacted]
OS Bucher Eyre+TA524 + 100% bendamustine SoC + treatment waning 3-5 + exp curve for No/Failed + equal utility after landmark + SMR 1.2	2.81	7.78	[redacted]	3.24	[redacted]

B.3.12. Subgroup analysis

None.

B.3.13. Benefits not captured in the QALY calculation

None.

B.3.14. Validation

B.3.14.1 Validation of cost-effectiveness analysis

The figures below illustrate health state membership over time in the health economic model.

Figure 23. Proportion surviving observed and modelled

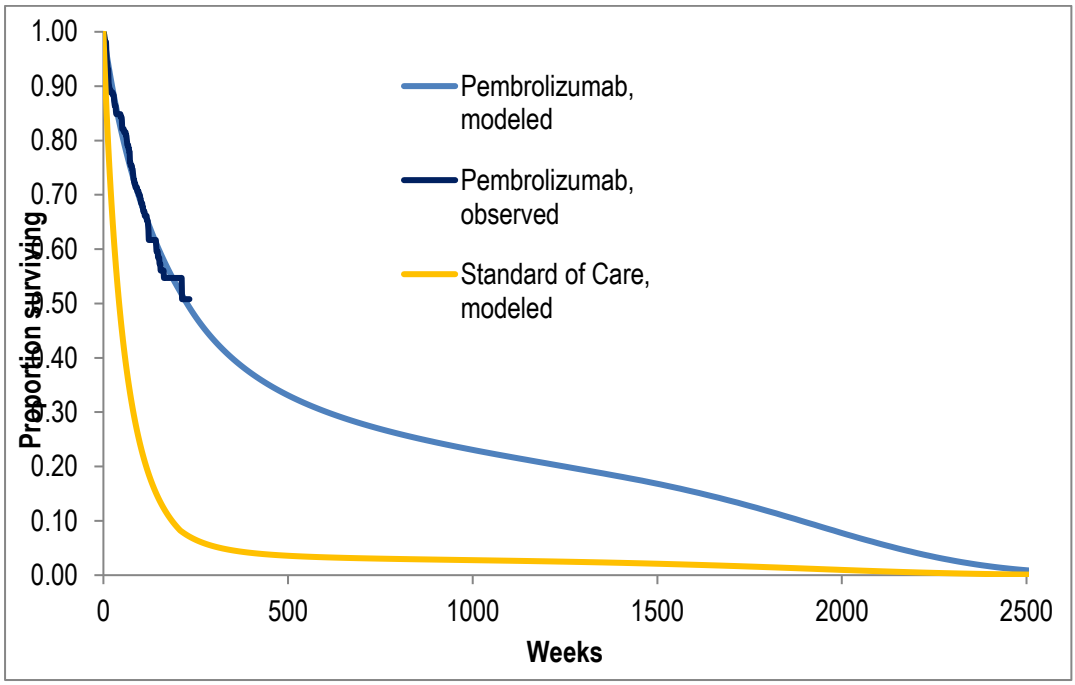


Figure 24. Proportion of patients in each health state over time in the pembrolizumab arm under base case assumptions

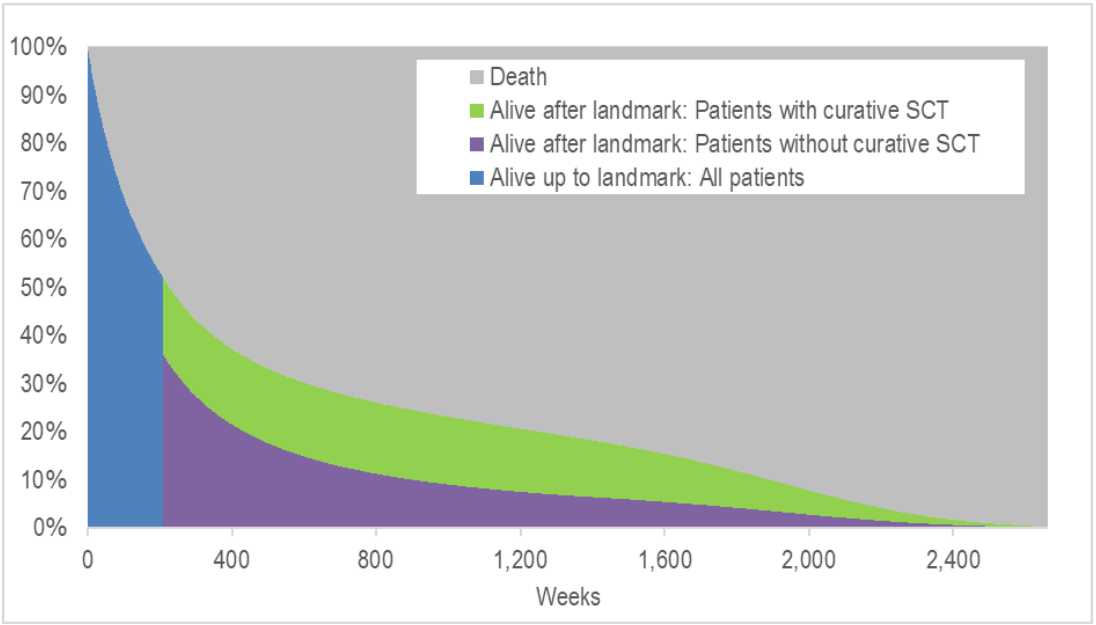
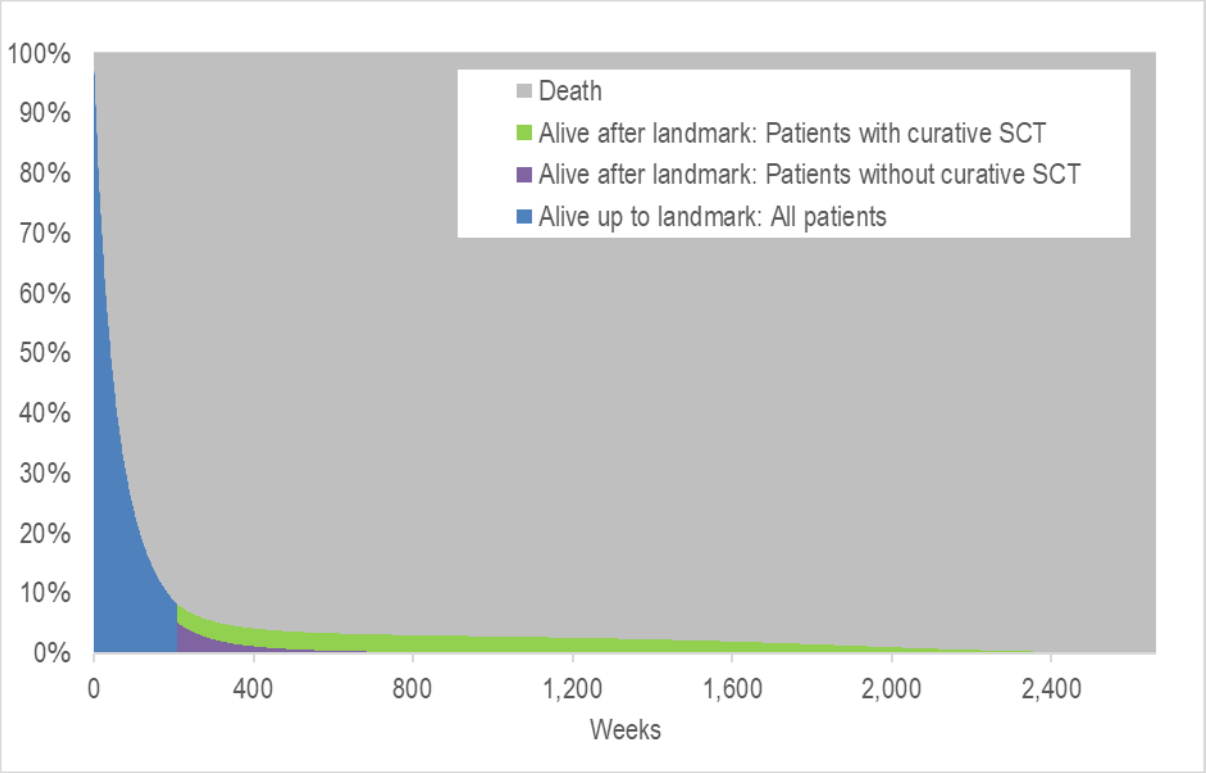
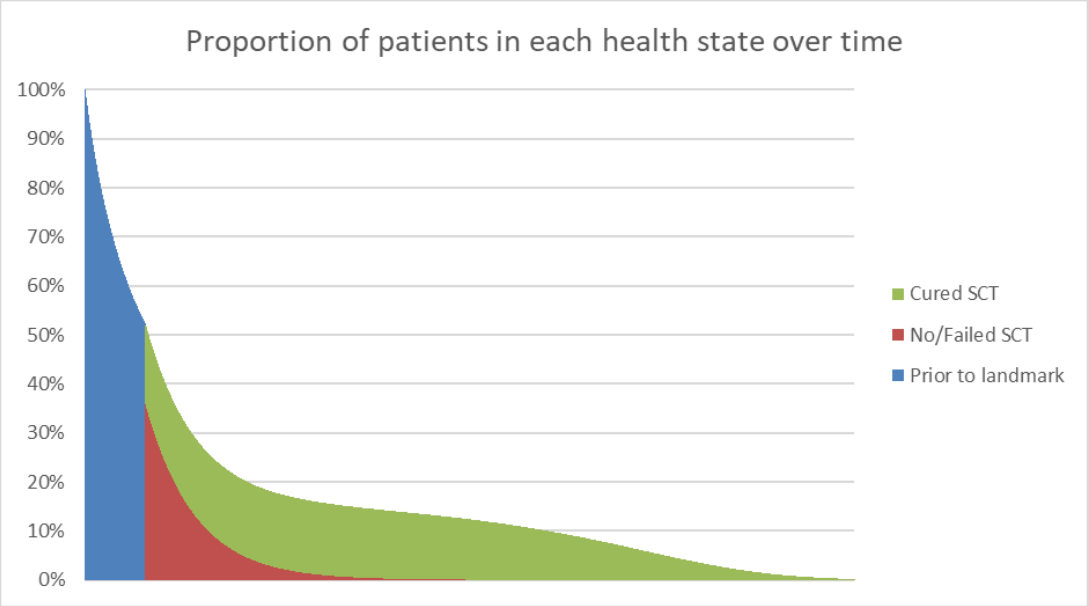


Figure 25. Proportion alive in each health state over time in the SoC arm under base case assumptions



We were not able to clinically validate the economic model's predictions for long term OS on pembrolizumab in time for the submission but note that the survival among the No/Failed SCT group is optimistic. Fitting an exponential curve provides a more conservative scenario as illustrated in Figure 26 below.

Figure 26. Proportion in each health state over time in the pembrolizumab arm when an exponential curve is used for No/Failed SCT post landmark



We examined the 4-year OS that is estimated for SoC using our preferred ITC and found that it was significantly lower than had been estimated by advisors at the UK advisory board (~20%) although it was similar to the SoC survival that had been estimated and accepted as plausible for 3L SoC during TA524 (~10%). OS among patients who failed BV and didn't get an SCT in the Eyre study was 25% at just 2 years with no emergent plateau. Were follow up longer in this study, a 4-year OS of <10% is certainly plausible. Estimate 3 (a Bucher ITC using Eyre and TA524) results in an HR of ~0.4, which estimates approximately 20% 4-year OS on SoC.

Overall we concluded that OS HRs between 0.2 and 0.4 generate 4-year OS in line with the range of data sources available. A treatment effect of this magnitude is plausible given very few patients get an SCT on SoC and survival expectations are low.

We note that the 'plunge-plateau' morphology of OS produced by the health economic model is consistent with that observed and accepted in NICE TA524 and appears to better reflect the natural history and care pathway than would direct parametric extrapolation of the SACT OS data. Nevertheless, we also extrapolated the SACT OS data fully to compare long term OS predicted by the economic model with simple survival extrapolations as shown in Figure 27.

Figure 27. SACT OS extrapolations

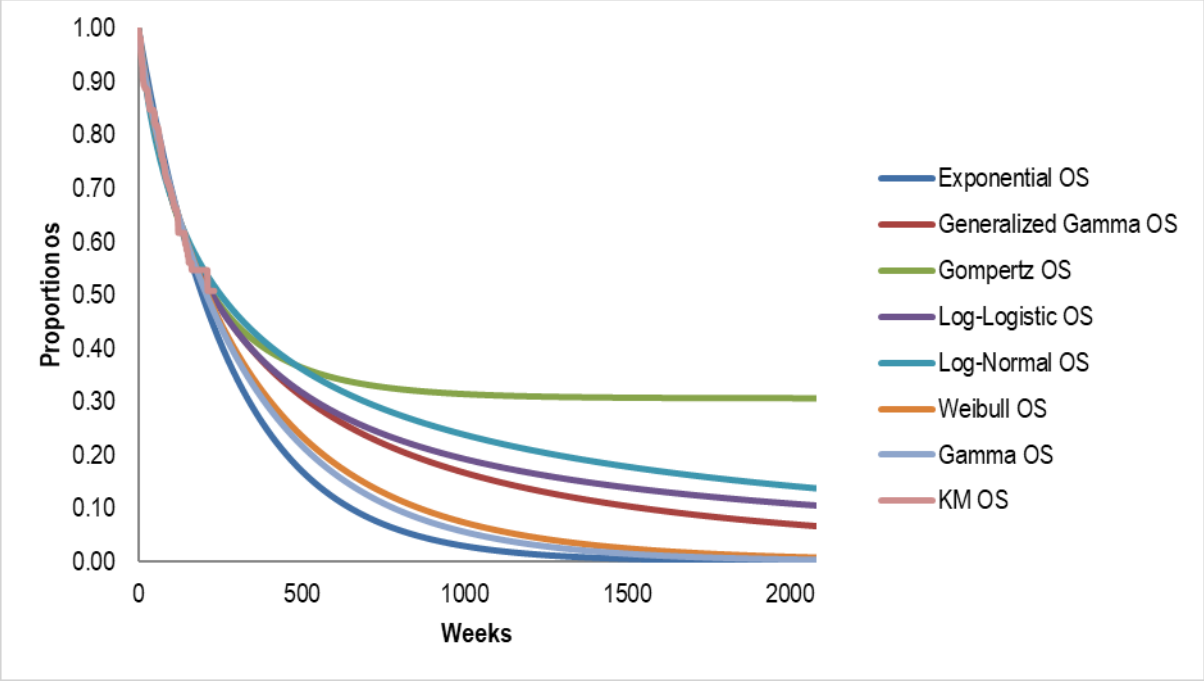


Table 74: Model fit statistics for extrapolating SACT OS

Functional Form	AIC	BIC	Average of AIC and BIC
Exponential	972.398	975.769	974.084
Weibull	972.039	978.780	975.409
Log-logistic	971.615	978.357	974.986
Log-normal	972.077	978.818	975.448
Gompertz	971.727	978.469	975.098
Generalised Gamma	973.198	983.310	978.254
Gamma	972.278	979.020	975.649

With OS being immature, AIC/BIC statistics being very similar and no curve able to explicitly capture the effect of curative SCTs, the figure above is unimportant except to demonstrate that the model's predictions about long term survival are within the (admittedly wide) range of predictions produced by standard extrapolation methods. The base case economic model produces OS of 32%, 22% and 16% at 10, 20 and 30 years respectively, which is roughly equivalent to the log-normal curve above (35%, 23%, 17% at 10, 20 and 30 years respectively).

B.3.15. Interpretation and conclusions of economic evidence

Like many haematology-oncology indications submitted to NICE, this model is based on a single arm trial and has made used of a linked evidence approach to estimate the longer term outcomes associated with potentially curative SCTs. The inherent uncertainty is fortunately tempered by the magnitude of treatment benefit for pembrolizumab in this setting. Expectations for patients on SoC are low, with very few receiving SCT and surviving to four years. By contrast, in a real-world UK setting, 30% of pembrolizumab patients received an SCT and median OS was not reached at 4 years. Although point estimates for the various treatment effects are inherently uncertain because they have not been estimated via a parallel RCT, the model is not sensitive to even conservative sets of adjustments to input parameters and assumptions. Most of the treatment effects used in the model can even be removed entirely without rendering pembrolizumab cost-ineffective. The base case ICER and all plausible sensitivity analyses are far below the £20-£30k/QALY threshold and the PSA demonstrated an [redacted] probability of cost-effectiveness vs a threshold of £30,000/QALY. The most important treatment effect in the model appears to be OS HR to the landmark. Although SCT is the primary generator of life years in the model, the great expense of the intervention means that it doesn't contribute a large amount of net monetary benefit to the overall results. This is why the ICER is not very sensitive to altering the SCT related assumptions. The selection of an exponential curve that implements constant transition probabilities to death for the No/Failed SCT group in both arms and implementation of a treatment waning assumption on these probabilities also both appear to impact the ICER somewhat.

A combinatorial analysis where a more conservative pre-landmark OS HR was chosen along with exponential post-landmark transitions, treatment waning, an SMR on the cured patients, SoC being comprised 100% of bendamustine (an inexpensive SoC option) and equal utility after the landmark was undertaken. This analysis could reasonably be characterized as a conservative alternative to the base case and produced an ICER of [redacted].

The approach taken to economic modelling has several strengths. Base case inputs for pembrolizumab are based on SACT data, which reflects clinical practice in the

NHS. Utility estimates are drawn from a randomised clinical trial (albeit at an earlier line of therapy). Estimates of parameters for which there were no data available from trials or the literature were elicited from experts using a SEE approach to fully capture the uncertainty in these estimates and minimise the risk of bias. Thanks to its time in the CDF, clinical experts have significant experience using pembrolizumab in cHL and were able to estimate model parameters more confidently than would have been the case had this treatment been entirely novel. KEYNOTE-204 has also established pembrolizumab as more effective than BV, which had been the standard of care in R/R cHL for a number of years and itself represented a step change in management over normal chemotherapy options when first approved. This gives more confidence to the primary OS HR treatment effect estimates. The long term portion of the model is dominated by patients who are cured by SCT. This is a strength because the long term outcomes are less reliant on extrapolating observed hazards into the future but instead relate to known mortality rates.

Overall, the economic model gives confidence that pembrolizumab is highly cost-effective in this indication.

B.4. References

1. NICE. 2022. Pembrolizumab for treating relapsed or refractory classical Hodgkin lymphoma after stem cell transplant or at least 2 previous therapies. Last accessed: 17 May 2023. Available from: https://www.nice.org.uk/guidance/TA772/chapter/1-Recommendations.
2. NICE. 2018. Pembrolizumab for treating relapsed or refractory classical Hodgkin lymphoma. Last accessed: 17 May 2023. Available from: https://www.nice.org.uk/guidance/ta540/chapter/1-Recommendations.
3. NICE. 2023. Pembrolizumab for treating relapsed or refractory classical Hodgkin lymphoma [Review of TA540] [ID5084]. Last accessed: 23 August 2023. Available from: https://www.nice.org.uk/guidance/topic-selection/gid-ta11317/documents.
4. 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.
5. MSD UK Ltd. Classical Hodgkin Lymphoma Advisory Board Meeting, London, Monday 10 July 2023. MSD data on file: 2023. 2023.
6. Robert C, Ribas A, 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;384(9948):1109-17.
7. European Medicines Agency. 2021. Keytruda: Assessment report. Last accessed: 19 July 2023. Available from: https://www.ema.europa.eu/en/documents/variation-report/keytruda-h-c-3820-ii-0090-epar-assessment-report-variation_en.pdf.
8. electronic Medicines Compendium. 2022. KEYTRUDA. Last accessed: 06 April 2023. Available from: https://www.medicines.org.uk/emc/product/2498.
9. Cancer Research UK. 2020. Hodgkin lymphoma. Last accessed: 21 July 2023. Available from: https://www.cancerresearchuk.org/about-cancer/hodgkin-lymphoma.
10. Cancer Research UK. 2023. Hodgkin lymphoma: statistics. Last accessed: 25 July 2023. Available from: https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/hodgkin-lymphoma.
11. Brice P, de Kerviler E, Friedberg JW. Classical Hodgkin lymphoma. *Lancet* 2021;398(10310):1518-27.
12. American Cancer Society. 2018. What Is Hodgkin Lymphoma? Last accessed: 23 July 2023. Available from: https://www.cancer.org/cancer/types/hodgkin-lymphoma/about/what-is-hodgkin-disease.html.
13. Connors JM, Cozen W, Steidl C, Carbone A, Hoppe RT, Flechtner HH, et al. Hodgkin lymphoma. *Nat Rev Dis Primers* 2020;6(1):61.

14. Carbone PP, Kaplan HS, Musshoff K, Smithers DW, Tubiana M. Report of the Committee on Hodgkin's Disease Staging Classification. *Cancer Res* 1971;31(11):1860-1.

15. Follows GA, Ardeshna KM, Barrington SF, Culligan DJ, Hoskin PJ, Linch D, et al. Guidelines for the first line management of classical Hodgkin lymphoma. *Br J Haematol* 2014;166(1):34-49.

16. Cancer Data NHS. 2020. Haematological malignancies. Last accessed: 26 July 2023. Available from: https://www.cancerdata.nhs.uk/getdataout/haem.

17. Cancer Intelligence Team. 2023. Performance measures across the cancer pathway: key stats. Last accessed: 26 July 2023. Available from: https://www.cancerresearchuk.org/sites/default/files/covid_and_cancer_key_stats_2023.07_stats_for_may_2021_released_july_2023.pdf.

18. Cancer Data NHS. 2023. Cancer incidence and mortality. Last accessed: 26 July 2023. Available from: https://www.cancerdata.nhs.uk/incidence_and_mortality.

19. Engelhardt BG, Holland DW, Brandt SJ, Chinratanalab W, Goodman SA, Greer JP, et al. High-dose chemotherapy followed by autologous stem cell transplantation for relapsed or refractory Hodgkin lymphoma: prognostic features and outcomes. *Leuk Lymphoma* 2007;48(9):1728-35.

20. 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.

21. 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.

22. Ortega-Ortega M, Hanly P, Pearce A, Soerjomataram I, Sharp L. Paid and unpaid productivity losses due to premature mortality from cancer in Europe in 2018. *Int J Cancer* 2022;150(4):580-93.

23. Collins GP, Parker AN, Pocock C, Kayani I, Sureda A, Illidge T, et al. Guideline on the management of primary resistant and relapsed classical Hodgkin lymphoma. *Br J Haematol* 2014;164(1):39-52.

24. Follows GA, Barrington SF, Bhuller KS, Culligan DJ, Cutter DJ, Gallop-Evans E, et al. Guideline for the first-line management of Classical Hodgkin Lymphoma - A British Society for Haematology guideline. *Br J Haematol* 2022;197(5):558-72.

25. Mohty R, Dulery R, Bazarbachi AH, Savani M, Hamed RA, Bazarbachi A, et al. Latest advances in the management of classical Hodgkin lymphoma: the era of novel therapies. *Blood Cancer J* 2021;11(7):126.

26. NICE. 2018. Brentuximab vedotin for treating CD30-positive Hodgkin lymphoma. Last accessed: 17 May 2023. Available from: https://www.nice.org.uk/guidance/ta524/chapter/1-Recommendations.

27. NICE. 2017. Nivolumab for treating relapsed or refractory classical Hodgkin lymphoma. Last accessed: 28 July 2023. Available from: https://www.nice.org.uk/guidance/ta462/chapter/1-Recommendations.

28. Merryman RW, Redd RA, Nishihori T, Chavez J, Nieto Y, Darrah JM, et al. Autologous stem cell transplantation after anti-PD-1 therapy for multiply relapsed or refractory Hodgkin lymphoma. *Blood Adv* 2021;5(6):1648-59.

29. Kuruvilla J, Ramchandren R, Santoro A, Paszkiewicz-Kozik E, Gasiorowski R, Johnson NA, et al. Pembrolizumab versus brentuximab vedotin in relapsed or refractory classical Hodgkin lymphoma (KEYNOTE-204): an interim analysis of a multicentre, randomised, open-label, phase 3 study. *Lancet Oncol* 2021;22(4):512-24.

30. Chen R, Zinzani PL, Fanale MA, Armand P, Johnson NA, Brice P, et al. Phase II Study of the Efficacy and Safety of Pembrolizumab for Relapsed/Refractory Classic Hodgkin Lymphoma. *J Clin Oncol* 2017;35(19):2125-32.

31. European Medicines Agency. 2012. Adcetris: brentuximab vedotin. Last accessed: 02 August 2023. Available from: https://www.ema.europa.eu/en/documents/smop-initial/chmp-summary-positive-opinion-adcetris_en.pdf.

32. Armand P, Zinzani PL, Lee HJ, Johnson NA, Brice P, Radford J, et al. Five-year follow-up of KEYNOTE-087: pembrolizumab monotherapy in relapsed/refractory classical Hodgkin lymphoma. *Blood* 2023.

33. Merck Sharp & Dohme. KEYNOTE-087 clinical study report. 2021.

34. NHS England. Pembrolizumab for treating relapsed or refractory classical Hodgkin lymphoma – data review. 2023.

35. NHS England. 2023. National Cancer Drugs Fund list. Last accessed: 04 August 2023. Available from: https://www.england.nhs.uk/wp-content/uploads/2017/04/National-Cancer-Drugs-Fund-list-version-1.270.pdf.

36. ClinicalTrials.gov. 2023. Study of pembrolizumab (MK-3475) in participants with relapsed or refractory classical Hodgkin Lymphoma (MK-3475-087/KEYNOTE-087). Last accessed: 07 August 2023. Available from: https://clinicaltrials.gov/study/NCT02453594.

37. 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.

38. 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.

39. Forero-Torres A, Fanale M, Advani R, Bartlett NL, Rosenblatt JD, Kennedy DA, et al. Brentuximab vedotin in transplant-naive patients with relapsed or refractory hodgkin lymphoma: analysis of two phase I studies. *Oncologist* 2012;17(8):1073-80.

40. Wells GA, O'Connell D, Peterson J, Welch V, Losos M, Tugwell P. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. Last accessed: 09 August 2023. Available from: https://www.ohri.ca/programs/clinical_epidemiology/oxford.asp.

41. Merck Sharp & Dohme. Pembrolizumab (MK-3475) in subjects with relapsed or refractory (R/R) classical Hodgkin Lymphoma (cHL). Data on file. 2023.

42. BMJ Best Practice. 2023. Graft versus host disease. Last accessed: 15 August 2023. Available from: https://bestpractice.bmj.com/topics/en-us/946.

43. Phillippo D, Ades A, Dias S, S. P, Abrams K, NJ. W. 2016. NICE DSU Technical support document 18: Methods for population-adjusted indirect comparisons in submissions to NICE. Last accessed: 29 August 2023. Available from: https://www.sheffield.ac.uk/media/34216/download?attachment.

44. Keeping S, Wu E, Chan K, Mojebi A, Ferrante SA, Balakumaran A. Pembrolizumab versus the standard of care for relapsed and refractory classical Hodgkin's lymphoma progressing after brentuximab vedotin: an indirect treatment comparison. *Expert Rev Hematol* 2018;11(6):503-11.

45. Rossi C, Gilhodes J, Maerevoet M, Herbaux C, Morschhauser F, Brice P, et al. Efficacy of chemotherapy or chemo-anti-PD-1 combination after failed anti-PD-1 therapy for relapsed and refractory Hodgkin lymphoma: A series from Lysa centers. *Am J Hematol* 2018;93:1042-9.

46. Carreau NA, Pail O, Armand P, Merryman R, Advani RH, Spinner MA, et al. Checkpoint Blockade Treatment May Sensitize Hodgkin Lymphoma to Subsequent Therapy. *Oncologist* 2020;25(10):878-85.

47. Carreau NA, Armand P, Merryman RW, Advani RH, Spinner MA, Herrera AF, et al. Checkpoint blockade treatment sensitises relapsed/refractory non-Hodgkin lymphoma to subsequent therapy. *Br J Haematol* 2020;191(1):44-51.

48. Schvartsman G, Peng SA, Bis G, Lee JJ, Benveniste MFK, Zhang J, et al. Response rates to single-agent chemotherapy after exposure to immune checkpoint inhibitors in advanced non-small cell lung cancer. *Lung Cancer* 2017;112:90-5.

49. Costantini A, Cadranel J. Increased Response Rates to Salvage Chemotherapy Administered after PD-1/PD-L1 Inhibitors in Patients with Non-Small Cell Lung Cancer. *J Thorac Oncol* 2018;13(4):e55-e6.

50. Harada D, Takata K, Mori S, Kozuki T, Takechi Y, Moriki S, et al. Previous Immune Checkpoint Inhibitor Treatment to Increase the Efficacy of Docetaxel and Ramucirumab Combination Chemotherapy. *Anticancer Res* 2019;39(9):4987-93.

51. Shiono A, Kaira K, Mouri A, Yamaguchi O, Hashimoto K, Uchida T, et al. Improved efficacy of ramucirumab plus docetaxel after nivolumab failure in previously treated non-small cell lung cancer patients. *Thorac Cancer* 2019;10(4):775-81.

52. Casadei B, Argnani L, Morigi A, Lolli G, Broccoli A, Pellegrini C, et al. Potential survival benefit for patients receiving autologous hematopoietic stem cell

transplantation after checkpoint inhibitors for relapsed/refractory Hodgkin lymphoma: A real-life experience. *Hematol Oncol* 2020;38(5):737-41.

53. Merryman RW, Castagna L, Giordano L, Ho VT, Corradini P, Guidetti A, et al. Allogeneic transplantation after PD-1 blockade for classic Hodgkin lymphoma. *Leukemia* 2021;35(9):2672-83.
54. NICE. 2022. NICE health technology evaluations: the manual. Last accessed: 23 August 2023. Available from: https://www.nice.org.uk/process/pmg36/resources/nice-health-technology-evaluations-the-manual-pdf-72286779244741.
55. Jones B, Ward T, Harrison JP, Hurst M, Tyas D, McEwan P, et al. The cost-effectiveness of nivolumab for the treatment of people with relapsed or refractory classical Hodgkin lymphoma following autologous stem cell transplant and brentuximab vedotin. *Value in Health* 2017;20:A433.
56. NICE. 2018. Inotuzumab ozogamicin for treating relapsed or refractory B-cell acute lymphoblastic leukaemia. Last accessed: 10 September 2023. Available from: https://www.nice.org.uk/guidance/TA541/chapter/1-Recommendations.
57. NICE. 2023. Brexucabtagene autoleucel for treating relapsed or refractory B-cell acute lymphoblastic leukaemia in people 26 years and over. Last accessed: 10 September 2023. Available from: https://www.nice.org.uk/guidance/TA893/chapter/1-Recommendations.
58. NICE. 2018. Tisagenlecleucel for treating relapsed or refractory B-cell acute lymphoblastic leukaemia in people aged up to 25 years. Last accessed: 10 September 2023. Available from: https://www.nice.org.uk/guidance/ta554/chapter/1-Recommendations.
59. NICE. 2021. Brexucabtagene autoleucel for treating relapsed or refractory mantle cell lymphoma. Last accessed: 10 September 2023. Available from: https://www.nice.org.uk/guidance/TA677/chapter/1-Recommendations.
60. Ara R, Brazier J, Zouraq IA. The Use of Health State Utility Values in Decision Models. *Pharmacoeconomics* 2017;35(Suppl 1):77-88.
61. 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.
62. van Agthoven M, Vellenga E, Fibbe WE, Kingma T, Uyl-de Groot CA. Cost analysis and quality of life assessment comparing patients undergoing autologous peripheral blood stem cell transplantation or autologous bone marrow transplantation for refractory or relapsed non-Hodgkin's lymphoma or Hodgkin's disease. a prospective randomised trial. *Eur J Cancer* 2001;37(14):1781-9.
63. Personal Social Services Research Unit. 2022. Unit Costs of Health and Social Care programme (2022 – 2027). Last accessed: 30 August 2023. Available from: https://www.pssru.ac.uk/unitcostsreport/.

64. NHS England. 2023. National Schedule of NHS Costs - Year 2020-21 - NHS trusts and NHS foundation trusts. Last accessed: 25 August 2023. Available from: https://www.england.nhs.uk/costing-in-the-nhs/national-cost-collection/.

65. Department of Health and Social Care. 2023. Drugs and pharmaceutical electronic market information tool (eMIT). Last accessed: 29 August 2023. Available from: https://www.gov.uk/government/publications/drugs-and-pharmaceutical-electronic-market-information-emit.

66. BNF. 2023. British National Formulary. Last accessed: 29 August 2023. Available from: https://bnf.nice.org.uk/.

67. Eyre TA, Phillips EH, Linton KM, Arumainathan A, Kassam S, Gibb A, et al. Results of a multicentre UK-wide retrospective study evaluating the efficacy of brentuximab vedotin in relapsed, refractory classical Hodgkin lymphoma in the transplant naive setting. *Br J Haematol* 2017;179(3):471-9.

68. MSD UK Ltd. Classical Hodgkin Lymphoma Advisory Board Meeting The King's Fund, London Monday 13 March 2017. MSD data on file: . 2017.

69. Bucher HC, Guyatt GH, Griffith LE, Walter SD. The results of direct and indirect treatment comparisons in meta-analysis of randomized controlled trials. *J Clin Epidemiol* 1997;50(6):683-91.

70. WebPlotDigitizer. 2022. WebPlotDigitizer. Last accessed: 23 August 2023. Available from: https://automeris.io/WebPlotDigitizer/.

71. Office for National Statistics. 2021. National life tables: England. Last accessed: 25 August 2023. Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/lifeexpectancies/datasets/nationallifetablesenglandreferencetables/current.

72. NICE. 2023. Pembrolizumab plus chemotherapy with or without bevacizumab for persistent, recurrent or metastatic cervical cancer. Last accessed: 23 August 2023. Available from: https://www.nice.org.uk/guidance/TA885/chapter/1-Recommendations.

73. Bojke L, Soares M, Claxton K, Colson A, Fox A, Jackson C, et al. Developing a reference protocol for structured expert elicitation in health-care decision-making: a mixed-methods study. *Health Technol Assess* 2021;25(37):1-124.

74. Jankovic D, Soares M, Bojke L, Horscroft J, Lee D. 2022. R code for building bespoke Shiny apps for conducting SEE. Last accessed: 25 August 2023. Available from: https://www.york.ac.uk/che/research/teehta/elicitation/steer/r-tool/.

75. SHELF. 2023. Tools to Support the Sheffield Elicitation Framework. Last accessed: 25 August 2023. Available from: https://rdrr.io/cran/SHELF/.

76. Horscroft J, Lee D, Jankovic D, Soares M, Bojke L. 2023. Structured expert elicitation resources (STEER). Last accessed: 25 August 2023. Available from: https://www.york.ac.uk/che/research/teehta/elicitation/steer/.

77. Moskowitz CH, Walewski J, Nademanee A, Masszi T, Agura E, Holowiecki J, et al. Five-year PFS from the AETHERA trial of brentuximab vedotin for Hodgkin lymphoma at high risk of progression or relapse. *Blood* 2018;132(25):2639-42.

78. MSD UK Ltd. Clinicians Survey Results, MSD UK and MedeConnect. MSD data on file. 2017.

79. MSD UK Ltd. Clinician Survey Questionnaire (MSD UK, MedeConnect). MSD data on file. 2017.

80. Moskowitz AJ, Hamlin PA, Jr., Perales MA, Gerecitano J, Horwitz SM, Matasar MJ, et al. Phase II study of bendamustine in relapsed and refractory Hodgkin lymphoma. *J Clin Oncol* 2013;31(4):456-60.

81. 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. *J Clin Oncol* 1999;17(12):3776-85.

82. Mainwaring PN, Cunningham D, Gregory W, Hoskin P, Hancock B, Norton AJ, et al. Mitoxantrone is superior to doxorubicin in a multiagent weekly regimen for patients older than 60 with high-grade lymphoma: results of a BNLI randomized trial of PAdriaCEBO versus PMitCEBO. *Blood* 2001;97(10):2991-7.

83. 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.

84. 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. *Br J Haematol* 2003;120(6):970-7.

85. 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. *Ann Oncol* 2003;14(12):1762-7.

86. Bartlett NL, Niedzwiecki D, Johnson JL, Friedberg JW, Johnson KB, van Besien K, et al. Gemcitabine, vinorelbine, and pegylated liposomal doxorubicin (GVD), a salvage regimen in relapsed Hodgkin's lymphoma: CALGB 59804. *Ann Oncol* 2007;18(6):1071-9.

87. Maybury B, Kimpton G, Otton S. A retrospective multicentre study of COCKLE, an oral chemotherapy regimen, as palliative treatment for high grade lymphoma. *Br J Haematol* 2019;185(4):803-6.

88. Martin A, Fernandez-Jimenez MC, Caballero MD, Canales MA, Perez-Simon JA, Garcia de Bustos J, et al. Long-term follow-up in patients treated with Mini-BEAM as salvage therapy for relapsed or refractory Hodgkin's disease. *Br J Haematol* 2001;113(1):161-71.

89. Kumar A, Casulo C, Yahalom J, Schoder H, Barr PM, Caron P, et al. Brentuximab vedotin and AVD followed by involved-site radiotherapy in early stage, unfavorable risk Hodgkin lymphoma. *Blood* 2016;128(11):1458-64.

90. Merck Sharp & Dohme. Keytruda (MK-3475). KN087A Phase II Clinical Trial of MK-3475 (Pembrolizumab) in Subjects with Relapsed or Refractory (R/R) Classical Hodgkin Lymphoma (cHL). Data on file. 2023.

91. 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):860-7.

92. 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.

93. Doyle S, Lloyd A, Walker M. Health state utility scores in advanced non-small cell lung cancer. *Lung Cancer* 2008;62(3):374-80.

94. 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.

95. 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.

96. 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.

97. 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.

98. NICE. 2016. Ticagrelor for preventing atherothrombotic events after myocardial infarction. Last accessed: 23 August 2023. Available from: https://www.nice.org.uk/guidance/TA420/chapter/1-Recommendations.

99. NICE. 2014. Pixantrone monotherapy for treating multiply relapsed or refractory aggressive non-Hodgkin's B-cell lymphoma. Last accessed: 23 August 2023. Available from: https://www.nice.org.uk/guidance/ta306/chapter/1-Guidance.

100. NICE. 2017. Paclitaxel as albumin-bound nanoparticles with gemcitabine for untreated metastatic pancreatic cancer. Last accessed: 29 August 2023. Available from: https://www.nice.org.uk/guidance/TA476/chapter/1-Recommendations.

101. NICE. 2016. Ceritinib for previously treated anaplastic lymphoma kinase positive non-small-cell lung cancer. Last accessed: 29 August 2023. Available from: https://www.nice.org.uk/guidance/TA395/chapter/1-Recommendations.

102. NICE. 2017. Pegylated liposomal irinotecan for treating pancreatic cancer after gemcitabine. Last accessed: 23 August 2023. Available from: https://www.nice.org.uk/guidance/TA440/chapter/1-Recommendations.

103. NICE. 2016. Necitumumab for untreated advanced or metastatic squamous non-small-cell lung cancer. Last accessed: 23 August 2023. Available from: https://www.nice.org.uk/guidance/TA411/chapter/1-Recommendations.

104. NICE. 2015. Idelalisib for treating chronic lymphocytic leukaemia. Last accessed: 29 August 2023. Available from: https://www.nice.org.uk/guidance/ta359/chapter/1-Guidance.

105. NICE. 2015. Pembrolizumab for advanced melanoma not previously treated with ipilimumab. Last accessed: 29 August 2023. Available from: https://www.nice.org.uk/guidance/ta366/chapter/1-Guidance.

106. NICE. 2017. Pembrolizumab for treating PD-L1-positive non-small-cell lung cancer after chemotherapy. Last accessed: 29 August 2023. Available from: https://www.nice.org.uk/guidance/TA428/chapter/1-Recommendations.

107. BNF. 2023. Keytruda 100mg/4ml concentrate for solution for infusion vials Merck Sharp & Dohme (UK) Ltd. Last accessed: 23 August 2023. Available from: https://bnf.nice.org.uk/drugs/pembrolizumab-specialist-drug/medicinal-forms/.

108. NHS England. 2016. DECC: Relapsed and progressive Hodgkin and Non Hodgkin Lymphoma. Last accessed: 29 August 2023. Available from: https://www.northerncanceralliance.nhs.uk/wp-content/uploads/2018/11/DECC-NECN-protocol-CRP08-H0062.pdf.

109. Lymphoma Group. 2022. Mini-BEAM (modified) Last accessed: 29 August 2023. Available from: https://nssg.oxford-haematology.org.uk/lymphoma/documents/lymphoma-chemo-protocols/L-30-mini-beam-modified.pdf.

110. NICE. 2019. B. Evidence reviews for the clinical and costeffectiveness of routine MRI or CT of the brain in the management of people with lung cancer prior to therapy with curative intent Last accessed: 25 August 2023. Available from: https://www.nice.org.uk/guidance/ng122/evidence/evidence-review-b-clinical-and-costeffectiveness-of-routine-mri-or-ct-of-the-brain-in-the-management-of-people-with-lung-cancer-prior-to-therapy%20-with-curative-intent-pdf-6722112207.

111. Radiologists RCo. 2019. Radiotherapy dose fractionation, third edition. Last accessed: 25 August 2023. Available from: https://www.rcr.ac.uk/publication/radiotherapy-dose-fractionation-third-edition.

112. Ansell S, Brockelmann PJ, von Keudell G, Lee HJ, Santoro A, Zinzani PL, et al. HL-398: Five-Year Overall Survival from the CheckMate 205 Study of Nivolumab for Relapsed or Refractory (R/R) Classical Hodgkin Lymphoma (cHL). *Clinical Lymphoma, Myeloma and Leukemia* 2021;21:S373-S4.

113. Georghiou T, Bardsley M. Exploring the cost of care at the end of life. *Research report* 2014.

114. NICE. 2022. Asciminib for treating chronic myeloid leukaemia after 2 or more tyrosine kinase inhibitors. Last accessed: 25 August 2023. Available from: https://www.nice.org.uk/guidance/TA813/chapter/1-Recommendations.

115. NICE. 2016. Azacitidine for treating acute myeloid leukaemia with more than 30% bone marrow blasts. Last accessed: 23 August 2023. Available from: https://www.nice.org.uk/guidance/TA399/chapter/1-Recommendations.

116. NICE. 2016. Cabazitaxel for hormone-relapsed metastatic prostate cancer treated with docetaxel. Last accessed: 23 August 2023. Available from: https://www.nice.org.uk/guidance/TA391/chapter/1-Recommendations.

117. NICE. 2019. Tisagenlecleucel for treating relapsed or refractory diffuse large B-cell lymphoma after 2 or more systemic therapies. Last accessed: 25 August 2023. Available from: https://www.nice.org.uk/guidance/ta567/chapter/1-Recommendations.

118. NHS Blood and Transport. Unrelated donor stem cell transplantation in the UK. 2014.

119. Blommestein HM, Verelst SG, Huijgens PC, Blijlevens NM, Cornelissen JJ, Uyl-de Groot CA. Real-world costs of autologous and allogeneic stem cell transplantations for haematological diseases: a multicentre study. *Ann Hematol* 2012;91(12):1945-52.

120. Schneider P, McNamara S, Love-Koh J, Doran T, Gutacker N. 2021. QALY shortfall calculator. Last accessed: 30 August 2023. Available from: https://r4scharr.shinyapps.io/shortfall/.

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single technology appraisal

Pembrolizumab for treating relapsed or refractory classical Hodgkin lymphoma [review of TA540]

Summary of Information for Patients (SIP)

September 2023

File name	Version	Contains confidential information	Date
ID5084_KN087 MSD_cHL_NICE STA_SIP_v1.0	Final	No	20/09/2023

Summary of Information for Patients (SIP):

The pharmaceutical company perspective

What is the SIP?

The Summary of Information for Patients (SIP) is written by the company who is seeking approval from NICE for their treatment to be sold to the NHS for use in England. It is a plain English summary of their submission written for patients participating in the evaluation. It is not independently checked, although members of the public involvement team at NICE will have read it to double-check for marketing and promotional content before it is sent to you.

The **Summary of Information for Patients** template has been adapted for use at NICE from the Health Technology Assessment International – Patient & Citizens Involvement Group (HTAi PCIG). Information about the development is available in an open-access IJTAHC journal article

SECTION 1: Submission summary

1a) Name of the medicine (generic and brand name):

Pembrolizumab (KEYTRUDA®)

1b) Population this treatment will be used by. Please outline the main patient population that is being appraised by NICE:

The patient population being looked at by NICE is adults with relapsed or refractory classical Hodgkin lymphoma (R/R cHL) who have had brentuximab vedotin (BV) and cannot have autologous stem cell transplant (autoSCT). This patient population can have pembrolizumab through the Cancer Drugs Fund (CDF). The current NICE technology appraisal is a review of evidence collected in clinical trials and from the CDF after pembrolizumab was made available in September 2018. In the original appraisal (NICE Technology Appraisal 540)[1], NICE had some questions on i) the time to stem cell transplant, ii) how many patients treated with pembrolizumab would go on to receive treatments that could cure cHL, such as autoSCT, and iii) how much longer patients treated with pembrolizumab would live compared with treatments already available in the NHS. As part of the CDF review, NHS England collected data to address these uncertainties which will inform the evidence in this current review.

1c) Authorisation: Please provide marketing authorisation information, date of approval and link to the regulatory agency approval. If the marketing authorisation is pending, please state this, and reference the section of the company submission with the anticipated dates for approval.

Under its licence, which states how pembrolizumab can be used, pembrolizumab (also known as KEYTRUDA) as “monotherapy is indicated for the treatment of adult and paediatric patients aged 3 years and older with relapsed or refractory classical Hodgkin lymphoma who have failed

autologous stem cell transplant (ASCT) or following at least two prior therapies when ASCT is not a treatment option."[2]

1d) Disclosures. Please be transparent about any existing collaborations (or broader conflicts of interest) between the pharmaceutical company and patient groups relevant to the medicine. Please outline the reason and purpose for the engagement/activity and any financial support provided:

Response: The company has an ongoing working relationship with the Clinical Trials Support Service (CTSS) at Blood Cancer UK, involving the provision of information related to the company's haematology clinical trials recruiting in the UK

SECTION 2: Current landscape

2a) The condition – clinical presentation and impact

Please provide a few sentences to describe the condition that is being assessed by NICE and the number of people who are currently living with this condition in England.

Please outline in general terms how the condition affects the quality of life of patients and their families/caregivers. Please highlight any mortality/morbidity data relating to the condition if available. If the company is making a case for the impact of the treatment on carers this should be clearly stated and explained.

Hodgkin lymphoma is an uncommon cancer that develops in the lymphatic system, which is a network of vessels and glands spread throughout the body.[3]

The lymphatic system is part of the immune system. Clear fluid called lymph flows through the lymphatic vessels and contains infection-fighting white blood cells, known as lymphocytes. In Hodgkin lymphoma, the body makes too many lymphocytes which means the immune system does not function as it should and increases vulnerability to infection and prevents other blood cells from working properly. The most common symptom of Hodgkin lymphoma is a painless swelling in a lymph node, usually in the neck, armpit or groin.

cHL is the most common form of Hodgkin lymphoma and accounts for approximately 93% of all cases.[4] This appraisal focuses on patients who relapse (disease returns after treatment) or become refractory (disease does not respond to treatment). Patients who relapse often have the same symptoms as when they were first diagnosed. A lymphoma diagnosis or relapse can often trigger a range of feelings and concerns for both patients and their caregivers. Additionally, cancer treatment can cause physical discomfort.

In the UK, around 2,100 patients are diagnosed with cHL each year most commonly occurring in people aged 20-40 and over 75.[5] Overall, around 8 out of 10 people with HL live at least 5 years and most patients have the potential to be cured.[3] There is evidence that people with R/R cHL who cannot have SCT do not survive as long as this. Clinical experts consulted for this appraisal estimated that 1 in 5 people in the patient population for this evaluation would live for 4 years. Therefore, there is an urgent need for effective treatments for these patients.

The population of interest for this appraisal is people who are not suitable for SCT. Having an SCT is an intense and challenging experience, and some people may be too unwell to undergo this

procedure or may decide that they are unwilling to accept the risks associated with it. Therefore, people with relapsed or refractory cHL who cannot have an SCT would benefit from a treatment that either improves their health enough that they can receive a potentially curative SCT or improves their quality of life if they are not able to have an SCT.

2b) Diagnosis of the condition (in relation to the medicine being evaluated)

Please briefly explain how the condition is currently diagnosed and how this impacts patients. Are there any additional diagnostic tests required with the new treatment?

The only way to confirm a diagnosis of Hodgkin lymphoma is by carrying out a biopsy.[3]

This is a minor surgical procedure where a sample of affected lymph node tissue is removed and studied in a laboratory.

After confirmed diagnosis, further testing is needed to check how far the lymphoma has spread, or to see whether it has returned after treatment. These tests may include (but are not limited to):[3]

- blood tests – samples of blood will be taken throughout diagnosis and treatment to check general health, the levels of red and white cells and platelets in the blood, and how well organs such as the liver and kidneys are working
- positron emission tomography (PET) scan – this scan measures the activity of cells in different parts of the body and can check the spread of the cancer and the impact of treatment

After a diagnosis of lymphoma, patients will need to have tests, either blood tests or scans, every few weeks to check for signs of whether the lymphoma has improved on treatment, or gotten worse, which can cause anxiety for the patient and family members before.[6] If their cancer goes away completely, doctors will want to regularly check that the cancer has not come back.

2c) Current treatment options:

The purpose of this section is to set the scene on how the condition is currently managed:

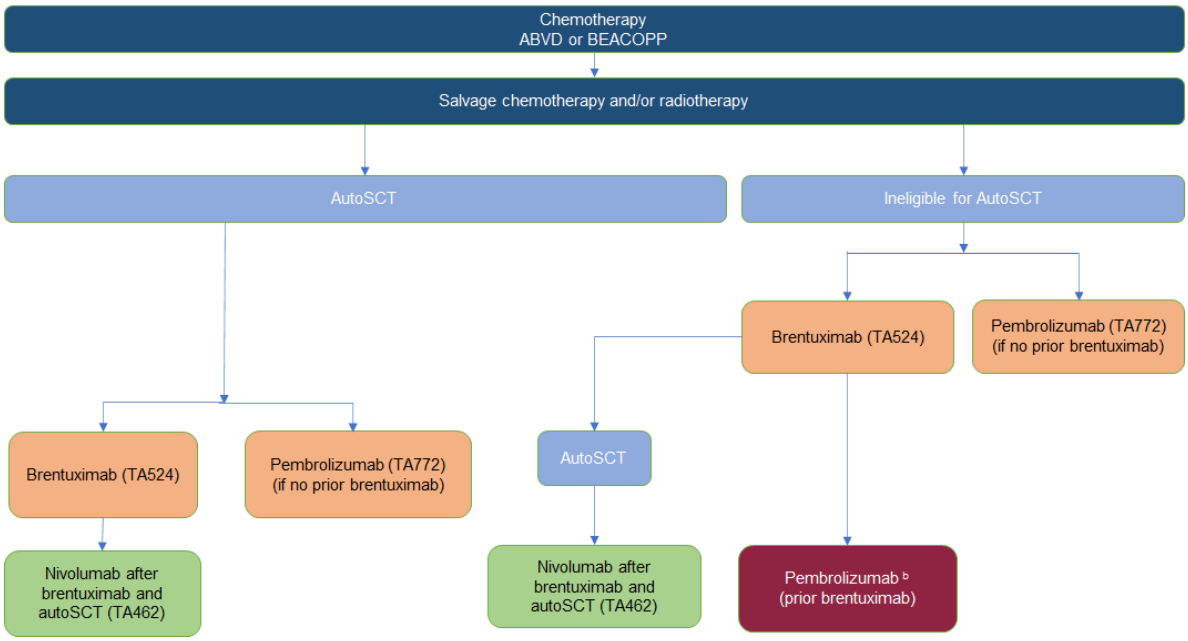
- What is the treatment pathway for this condition and where in this pathway the medicine is likely to be used? Please use diagrams to accompany text where possible. Please give emphasis to the specific setting and condition being considered by NICE in this review. For example, by referencing current treatment guidelines. It may be relevant to show the treatments people may have before and after the treatment under consideration in this SIP.
- Please also consider:
    - if there are multiple treatment options, and data suggest that some are more commonly used than others in the setting and condition being considered in this SIP, please report these data.
    - are there any drug–drug interactions and/or contraindications that commonly cause challenges for patient populations? If so, please explain what these are.

The current treatment pathway for R/R cHL is shown in figure 1. After chemotherapy some people may be able to have a SCT. The most common type of SCT is autoSCT, where the patient's own healthy blood stem cells are used to replace blood cells destroyed by high doses of chemotherapy and other treatments. In line with the treatment pathway, those who cannot have autoSCT can

receive BV (a form of targeted chemotherapy) or pembrolizumab (a type of immunotherapy). Pembrolizumab can be a bridging therapy to allogeneic SCT (alloSCT), a donor stem cell transplant which aims to replace bone marrow that is no longer working properly with healthy stem cells from another person.[7] In this scenario the aim of treatment with pembrolizumab is to delay rapid disease progression and get patients fit enough to receiving alloSCT.

The focus of this appraisal is patients who have disease that relapses after, or does not respond to, BV. The aim of treatment with pembrolizumab at this stage is either to achieve sufficient disease response to enable SCT to be done (which may cure the disease) or to achieve sufficient disease response to improve quality of life.

Figure 1. Treatment pathway for R/R cHL



2d) Patient-based evidence (PBE) about living with the condition

Context:

- **Patient-based evidence (PBE)** is when patients input into scientific research, specifically to provide experiences of their symptoms, needs, perceptions, quality of life issues or experiences of the medicine they are currently taking. PBE might also include carer burden and outputs from patient preference studies, when conducted in order to show what matters most to patients and carers and where their greatest needs are. Such research can inform the selection of patient-relevant endpoints in clinical trials.

In this section, please provide a summary of any PBE that has been collected or published to demonstrate what is understood about **patient needs and disease experiences**. Please include the methods used for collecting this evidence. Any such evidence included in the SIP should be formally referenced wherever possible and references included.

Patients with relapsed and refractory blood cancers are faced with many challenges, including the difficulties with taking chemotherapy, and the mental and emotional impacts associated with the diagnosis of a potentially fatal illness.

By targeting the rapidly dividing cancer cells, chemotherapy aims to ease some of these symptoms. However further issues can be caused by the side effects of chemotherapy. Each person experiences

side effects from chemotherapy differently, and different chemotherapy drugs cause different side effects.[8] Many people feel fine for the first few hours following chemotherapy. Usually, some reaction occurs about four to six hours later. However, some people do not react until 12 or even 24 to 48 hours after treatment. Some people experience many of the side effects described, while others experience almost none. Some of the most common side effects are summarised below:

- Infection and fever – due to chemotherapy reducing a patient's white blood cell count (the cells that help fight infection), chemotherapy patients are more susceptible to infection. This can result in a fever.
- Flu-like symptoms - Around the third day following a chemotherapy treatment, some people may experience flu-like symptoms such as muscle aches and pains.
- Nausea (though not all chemotherapy drugs cause nausea).
- Fatigue, which can range from mild (usually cured by additional rest) to severe which may routinely impact a patient's ability to carry out everyday tasks such as cooking or bathing.
- Hair loss - begins about two to three weeks after starting chemotherapy. Some people will lose relatively little hair, while others may lose the hair on their head, eyelashes and eyebrows, as well as other body hair. Many people feel that hair loss is one of the most difficult aspects of chemotherapy treatment.

People with blood cancer may have the option of receiving a SCT This is an intensive form of treatment. It can take 3 to 6 months to recover physically, with the immune system taking 6 to 12 months to recover.[9] Usually patients will stay in protective isolation in a transplant unit for a few months during the stem cell transplant and recovery process (though some patients who are well enough may be able to stay at home or at arranged accommodation overnight and return to hospital for treatment, blood tests and check ups). Among alloSCT patients, the transplanted stem cells may fail to settle or make new blood cells (known as graft failure) often requiring a second SCT. In addition, alloSCT patients have a greater risk of developing a complication known as graft versus host disease (GvHD) where the new immune system from the donor cells attacks the healthy tissues, causing serious side effects. The reaction to the transplant of cells can happen straight away or much later.[10] If the reaction happens straight away, this is called acute and occurs in about half of alloSCTs. If the reaction happens a few months after the SCT, then this is chronic and happens in about 1 in 4 cases. Graft failure and GvHD are rare in autoSCT.

People who have had successful treatment for blood cancer still face challenges.[11] The mental after-effects of blood cancer treatment are often described by patients as being like an unexpected side effect of treatment. Patients often state that instead of "returning to normal" they instead "create a new normal." While people who recover from blood cancer may be relieved or happy to have completed treatment, they may also feel pressure to return to their previous responsibilities, isolated due to a decrease in support, angry or sad about how much their life has changed and fearful of relapse. Some people may feel guilty for experiencing these mixed emotions. People may also have to deal with long-term side effects of intensive chemotherapy, for example, being unable to conceive.

SECTION 3: The treatment

3a) How does the new treatment work?

What are the important features of this treatment?

Please outline as clearly as possible important details that you consider relevant to patients relating to the mechanism of action and how the medicine interacts with the body

Where possible, please describe how you feel the medicine is innovative or novel, and how this might be important to patients and their communities. If there are relevant documents which have been produced to support your regulatory submission such as a summary of product characteristics or patient information leaflet, please provide a link to these.
An important role of the immune system is the ability to be able to tell the difference between healthy and unhealthy cells. The level of activity of immune cells, such as T cells, is crucial to maintaining a balanced immune response. Under normal conditions, a protein called programmed death-ligand 1 (PD-L1) which naturally occurs on cells, plays an important role in maintaining this balanced immune response. PD-L1 binds to its PD-1 receptor on immune T cells, which lessens the ability of immune T cells to attack. This ensures that normal cells are protected from excessive damage. However, PD-L1 is produced in larger amounts on cancerous cells than normal cells. As a result, when binding to PD-1 on immune T cells, this interaction tricks the immune system thereby protecting the tumour from being attacked by the body's immune system. PD-1 inhibitors, such as pembrolizumab, act to block the interaction between PD-1 and PD-L1 and by doing so, boost the immune response which helps the person's own immune cells to attack the cancer cells. The summary of product characteristics (SmPC) and the patient information leaflet (PIL) for pembrolizumab can be found by following this link: KEYTRUDA

3b) Combinations with other medicines

Is the medicine intended to be used in combination with any other medicines? • Yes / No If yes, please explain why and how the medicines work together. Please outline the mechanism of action of those other medicines so it is clear to patients why they are used together. If yes, please also provide information on the availability of the other medicine(s) as well as the main side effects. **If this submission is for a combination treatment, please ensure the sections on efficacy (3e), quality of life (3f) and safety/side effects (3g) focus on data that relate to the combination, rather than the individual treatments.**
No.

3c) Administration and dosing

How and where is the treatment given or taken? Please include the dose, how often the treatment should be given/taken, and how long the treatment should be given/taken for. How will this administration method or dosing potentially affect patients and caregivers? How does this differ to existing treatments?
Pembrolizumab comes in a 25mg/mL concentrate solution for infusion. One 4mL vial of concentrate contains 100 mg of pembrolizumab.

The recommended dose of KEYTRUDA in adults is either 200 mg every 3 weeks or 400 mg every 6 weeks administered as an infusion into the vein (intravenous infusion) over 30 minutes.[12, 13].

In line with its licence, pembrolizumab may be given for up to 35 cycles (approximately two years) as long as it is working (i.e. as long as the cancer does not progress) and side effects are tolerable. Patients may also stop treatment prior to receiving an SCT or if they choose to do so for any other reason.

3d) Current clinical trials

Please provide a list of completed or ongoing clinical trials for the treatment. Please provide a brief top-level summary for each trial, such as title/name, location, population, patient group size, comparators, key inclusion and exclusion criteria and completion dates etc. Please provide references to further information about the trials or publications from the trials.

A search on clinicaltrials.gov for recruited, enrolling by invitation, active but not recruiting, or completed studies on pembrolizumab returns 1,637 (search conducted 6th September 2023). Of these, 17 are in relapsed or refractory cHL and listed below. Further details of these studies can be found by following the included links or searching for the study identifiers (NCT number or study name) on clinicaltrials.gov.

Table 1. Current clinical trials of pembrolizumab in cHL

Study Title	NCT Number	Status	Phase
Study of Magrolimab and Pembrolizumab in Relapsed or Refractory Classic Hodgkin Lymphoma	NCT04788043	Recruiting	Phase 2
Study of PD-1 Inhibitors After CD30.CAR T Cell Therapy in Relapsed/Refractory Hodgkin Lymphoma	NCT04134325	Recruiting	Early Phase 1
Study of Pembrolizumab With Bendamustine in Hodgkin Lymphoma	NCT04510636	Recruiting	Phase 2
Study of Pembrolizumab (MK-3475) in Participants With Relapsed or Refractory Classical Hodgkin Lymphoma (MK-3475-087/KEYNOTE-087)	NCT02453594	Active, not recruiting	Phase 2
Study of the Combination of AFM13 and Pembrolizumab in Patients With Relapsed or Refractory Classical Hodgkin Lymphoma	NCT02665650	Completed	Phase 1
Pembrolizumab and Combination Chemotherapy in Treating Patients With Relapsed or Refractory Hodgkin Lymphoma	NCT03077828	Active, not recruiting WITH RESULTS	Phase 2
Study of Pembrolizumab (MK-3475) vs. Brentuximab Vedotin in Participants With Relapsed or Refractory Classical Hodgkin Lymphoma (MK-3475-204/KEYNOTE-204)	NCT02684292	Active, not recruiting	Phase 3
A Study of Coformulated Favezelimab/Pembrolizumab (MK-4280A) Versus Physician's Choice Chemotherapy in PD-(L)1-refractory, Relapsed or Refractory Classical Hodgkin Lymphoma (MK-4280A-008)	NCT05508867	Recruiting	Phase 3
Pembrolizumab and Brentuximab Vedotin vs GDP and Stem Cell Transplant for Relapsed/Refractory Hodgkin Lymphoma	NCT05180097	Recruiting	Phase 2

A Study of Pembrolizumab (MK-3475) in Relapsed or Refractory Classical Hodgkin's Lymphoma (rrcHL) or Relapsed or Refractory Primary Mediastinal Large B-cell Lymphoma (rrPMBCL) (MK-3475-B68)	NCT04875195	Active, not recruiting	Phase 2
A Study of SAR444245 With or Without Other Anticancer Therapies for the Treatment of Adults and Adolescents With Relapsed or Refractory B Cell Lymphoma (Master Protocol) [Pegathor Lymphoma 205]	NCT05179603	Active, not recruiting	Phase 2
Pembrolizumab After ASCT for Hodgkin Lymphoma, DLBCL and T-NHL	NCT02362997	Completed WITH RESULTS	Phase 2
A Study of Pembrolizumab (MK-3475) in Pediatric Participants With an Advanced Solid Tumor or Lymphoma (MK-3475-051/KEYNOTE-051)	NCT02332668	Recruiting	Phase 1Phase 2
Pembrolizumab and Vorinostat in Treating Patients With Relapsed or Refractory Diffuse Large B-Cell Lymphoma, Follicular Lymphoma, or Hodgkin Lymphoma	NCT03150329	Active, not recruiting	Phase 1
Pembrolizumab and Ibrutinib in Treating Patients With Relapsed or Refractory Non-Hodgkin Lymphoma	NCT02950220	Completed	Phase 1
Testing the Addition of an Experimental Medication MK-3475 (Pembrolizumab) to Usual Anti-Retroviral Medications in Patients With HIV and Cancer	NCT02595866	Active, not recruiting	Phase 1
TTI-622 and TTI-621 in Combination With Pembrolizumab for the Treatment of Relapsed or Refractory Diffuse Large B-Cell Lymphoma	NCT05507541	Recruiting	Phase 2

3e) Efficacy

Efficacy is the measure of how well a treatment works in treating a specific condition.

In this section, please summarise all data that demonstrate how effective the treatment is compared with current treatments at treating the condition outlined in section 2a. Are any of the outcomes more important to patients than others and why? Are there any limitations to the data which may affect how to interpret the results? Please do not include academic or commercial in confidence information but where necessary reference the section of the company submission where this can be found.

The key results on how effective pembrolizumab is at treating R/R cHL comes from the KEYNOTE-087 trial. This trial included patients who are typically younger and more able to carry on with their daily activities than those in the NHS (the mean age of patient in the trial was 42 years versus 51 years in the CDF). These differences meant there was uncertainty around how many patients treated with pembrolizumab in NHS clinical practice would be able to receive SCTs and how long they would survive compared with those in the clinical trial. To address these uncertainties, information on survival and SCT rates were collected in the CDF. The proportion of people undergoing a SCT was similar in both KEYNOTE-087 and the SACT dataset, with 29.6% from KEYNOTE-087 and 30.2% of people in the SACT dataset receiving SCT. This contrasts with clinical expert opinion on the likelihood of SCT rate for standard of care (typically chemotherapy), which was estimated to be around 8%.

A summary of survival data from both sources is shown in the table below:

Table 2. Overall survival outcomes for treatment with pembrolizumab

Outcome	Cohort 2 of KEYNOTE-087 (N=81)	CDF dataset (N=215)
Number of events, n (%)	24 (29.6)	73 (34.0%)
Median OS (95% CI), months	Not reached	Not reached
OS rate at various time points		
• 6 months (%)	100.0[a]	88% (95% CI: 83% to 92%)
• 12 months (%)	96.3[a]	82% (95% CI: 76% to 87%)
• 18 months (%)	93.7[a]	75% (95% CI: 68% to 80%)
• 24 months (%)	91.1[a]	68% (95% CI: 61% to 75%)
• 36 months (%)	85.9[a]	56% (95% CI: 47% to 64%)
• 48 months (%)	76.5[a]	55% (95% CI: 46% to 63%)
• 60 months (%)	69.2[a]	N/A

* From product-limit (Kaplan–Meier) method for censored data.

Abbreviations: BICR, blinded independent central review; CDF, Cancer Drugs Fund; CI: confidence interval; IWG, international working group; N/A, not available; OS, overall survival; SACT, Systemic Anti-Cancer Therapy; SD, standard deviation.

As expected, survival rates were lower in clinical practice than in the KEYNOTE-087 trial. However, they are much higher than those expected for standard of care in the NHS. Clinical experts estimated that the survival rate at 4 years with standard of care would be about 20%, which is less than half of the 55% survival rate reported for pembrolizumab in the SACT data.

3f) Quality of life impact of the medicine and patient preference information

What is the clinical evidence for a potential impact of this medicine on the quality of life of patients and their families/caregivers? What quality of life instrument was used? If the EuroQol-5D (EQ-5D) was used does it sufficiently capture quality of life for this condition? Are there other disease specific quality of life measures that should also be considered as supplementary information?

Please outline in plain language any quality of life related data such as **patient reported outcomes (PROs).**

Please include any **patient preference information (PPI)** relating to the drug profile, for instance research to understand willingness to accept the risk of side effects given the added benefit of treatment. Please include all references as required.

The KEYNOTE-087 trial used two types of questionnaire to measure the quality of life (QoL) of patients — the European Organization for the Research and Treatment of Cancer Quality of Life Questionnaire (EORTC QoL QLQ-C30), which looks specifically at the quality of life of cancer patients, and the European Quality of Life 5 Dimensions 3 Level Version (EQ-5D-3L), which is a generic tool that looks at the general health of a patient.

The EQ-5D-3L is of most relevance to a NICE appraisal and consists of 2 pages: the EQ-5D descriptive system and the EQ visual analogue scale (EQ VAS). The EQ-5D descriptive system has

five questions on mobility, self-care, pain, usual activities, and psychological status with three possible answers for each item (1=no problem, 2=moderate problem, 3=severe problem). Results from these questions can then be combined and scaled to produce a single score with a maximum score of 1. Scores can vary from 0, which represents death, to 1 which represents the best possible health state. The EORTC uses different questions, however it also produces a score that is meant to represent a patient's QoL. The EQ VAS records the patient's self-rated health on a vertical visual analogue scale, where the endpoints are labelled 'The best health you can imagine' and 'The worst health you can imagine'. From this we can gather three scores (from the EQ-5D questionnaire, the EQ-5D VAS and the EORTC questionnaires) that can assess how a patient feels throughout their treatment.

There is also comparative quality of life data available from the KEYNOTE-204 trial comparing pembrolizumab with BV. These data are based on patients who were treated at an earlier line of therapy and BV treatment is more effective than standard of care. In the absence of a comparator arm in the KEYNOTE-087 trial, data from KEYNOTE-204 may be a suitable alternative for assessing the impact of pembrolizumab on patient's quality of life.

Results

For both questionnaires, on average patients in KEYNOTE-087 reported a clinically meaningful improvement in quality of life after 12 weeks of treatment. However, the scores were different depending on whether the patients achieved a response on pembrolizumab. Patients who had a complete or partial response reported the largest improvement. Patients whose cancer neither became better or worse (stable disease) reported a smaller improvement. Patients whose cancer got worse (progressive disease) reported a smaller improvement still. Full details are available in the submission documents.

Quality of life data from the KEYNOTE-204 trial, comparing pembrolizumab to BV at an earlier line of therapy, was also used in the economic model (see section 3i). In this trial patients treated with pembrolizumab had a clinically meaningful improvement in quality of life compared with those treated with BV.

3g) Safety of the medicine and side effects

When NICE appraises a treatment, it will pay close attention to the balance of the benefits of the treatment in relation to its potential risks and any side effects. Therefore, please outline the main side effects (as opposed to a complete list) of this treatment and include details of a benefit/risk assessment where possible. This will support patient reviewers to consider the potential overall benefits and side effects that the medicine can offer.

Based on available data, please outline the most common side effects, how frequently they happen compared with standard treatment, how they could potentially be managed and how many people had treatment adjustments or stopped treatment. Where it will add value or context for patient readers, please include references to the Summary of Product Characteristics from regulatory agencies etc.

Pembrolizumab has been used in hospitals in England since 2015 [14]. Section 1b describes the different cancers that pembrolizumab is licensed to treat. The safety and side effects data from all the trials that have led to these licences are included in the pembrolizumab SmPC.[15] A summary of relevant safety information from the pembrolizumab SmPC has been provided below, giving doctors and other hospital staff clear guidance on what to do if a patient experiences an immune-related side effect.

The safety of pembrolizumab as monotherapy has been evaluated in 7,631 patients across tumour types. In this patient population, the median observation time was 8.5 months (range: 1 day to 39 months) and the most frequent adverse reactions with pembrolizumab were fatigue (31%), diarrhoea (22%), and nausea (20%). The majority of adverse reactions reported for monotherapy were of mild or moderate severity. The most serious adverse reactions were immune-related adverse reactions and severe infusion-related reactions. The incidences of immune-related adverse reactions were and 24.2% all Grades and 6.4% for Grades 3-5 in the metastatic setting.

Immune-related adverse reactions, including severe and fatal cases, have occurred in patients receiving pembrolizumab. Most immune-related adverse reactions occurring during treatment with pembrolizumab were reversible and managed with interruptions of pembrolizumab, administration of corticosteroids and/or supportive care. Immune-related adverse reactions have also occurred after the last dose of pembrolizumab. Immune-related adverse reactions affecting more than one body system can occur simultaneously.

For suspected immune-related adverse reactions, adequate evaluation to confirm aetiology or exclude other causes should be ensured. Based on the severity of the adverse reaction, pembrolizumab should be withheld and corticosteroids administered. Upon improvement to Grade ≤ 1, corticosteroid taper should be initiated and continued over at least 1 month. Based on limited data from clinical studies in patients whose immune-related adverse reactions could not be controlled with corticosteroid use, administration of other systemic immunosuppressants can be considered.

Pembrolizumab may be restarted within 12 weeks after last dose of pembrolizumab if the adverse reaction recovers to Grade ≤ 1 and corticosteroid dose has been reduced to ≤ 10 mg prednisone or equivalent per day.

Pembrolizumab must be permanently discontinued for any Grade 3 immune-related adverse reaction that recurs and for any Grade 4 immune-related adverse reaction toxicity, except for endocrinopathies that are controlled with replacement hormones.

The grading system for adverse reactions, or side effects, referred to above is explained in section 4a.

The side effects that were reported in the KEYNOTE-087 clinical trial are consistent with the common side effects listed in the pembrolizumab SmPC.

The most frequently reported adverse events (>15%) in the population in KEYNOTE-087 aligned to this evaluation are reported in the table below. As KEYNOTE-087 was a single arm trial, there are no data available to directly compare the frequency of adverse events with standard of care.

Table 3: Adverse events on treatment with pembrolizumab in the KEYNOTE-087 trial

Adverse effect	**KEYNOTE-087 Cohort 2 (N=81) n (%)**
Pyrexia	19 (23.5)

Cough	22 (27.2)	
Fatigue	17 (21.0)	
Diarrhoea	12 (14.8)	
Upper respiratory tract infection	7 (8.6)	
Nausea	11 (13.6)	
Vomiting	9 (11.1)	
Nasopharyngitis	16 (19.8)	
Arthralgia	12 (14.8)	
Hypothyroidism	13 (16.0)	

3h) Summary of key benefits of treatment for patients

Issues to consider in your response:

- Please outline what you feel are the key benefits of the treatment for patients, caregivers and their communities when compared with current treatments.
- Please include benefits related to the mode of action, effectiveness, safety and mode of administration
-

The key benefits to patients, caregivers and communities may include:

- Based on data from the CDF approximately 30% of patients treated with pembrolizumab in the NHS are able to have a potentially curative SCT. This contrasts with independent clinical expert estimates of approximately 8% for current standard of care, such as chemotherapy.
- Independent clinical experts consulted by the company estimated that the likelihood that a SCT will result in a cure is higher for patients taking immunotherapies such as pembrolizumab than it is for patients who have standard of care such as chemotherapy.
- Data from clinical trials shows that patients treated with pembrolizumab are likely to have improved QoL compared to those treated with standard of care (whether or not they receive an SCT).
- Providing a path to a potential cure provides hope to both patients and caregivers, whereas previously the prognosis for patients with relapsed/refractory cHL was exceptionally poor.
- Data from the clinical trial shows that side effects of treatment with pembrolizumab occur less frequently than reported for standard of care in the literature
- Many of the chemotherapy regimens available as standard of care consist of multiple components. As no other drugs are administered along with pembrolizumab, infusion times may be shorter.

3i) Summary of key disadvantages of treatment for patients

Issues to consider in your response:

• Please outline what you feel are the key disadvantages of the treatment for patients, caregivers and their communities when compared with current treatments. Which disadvantages are most important to patients and carers? • Please include disadvantages related to the mode of action, effectiveness, side effects and mode of administration • What is the impact of any disadvantages highlighted compared with current treatments
• Patients are at an increased risk of developing immune-related side effects, some of which may last beyond the patient stopping pembrolizumab. Please note there is clear guidance provided in the SmPC that instructs healthcare providers on how to manage these side effects. • Pembrolizumab, like any other medicine, does not work the same in every patient. Not all patients' cancer will respond to treatment and it may not result in an extended life expectancy.

3i) Value and economic considerations

Introduction for patients: Health services want to get the most value from their budget and therefore need to decide whether a new treatment provides good value compared with other treatments. To do this they consider the costs of treating patients and how patients' health will improve, from feeling better and/or living longer, compared with the treatments already in use. The drug manufacturer provides this information, often presented using a health economic model. In completing your input to the NICE appraisal process for the medicine, you may wish to reflect on: • The extent to which you agree/disagree with the value arguments presented below (e.g., whether you feel these are the relevant health outcomes, addressing the unmet needs and issues faced by patients; were any improvements that would be important to you missed out, not tested or not proven?) • If you feel the benefits or side effects of the medicine, including how and when it is given or taken, would have positive or negative financial implications for patients or their families (e.g., travel costs, time-off work)? • How the condition, taking the new treatment compared with current treatments affects your quality of life.
Cost-effectiveness relates to how much new health (or quality-adjusted life years, QALYs) the new medicine produces compared to its additional cost (vs. current care), for a typical/average patient and whether the new health is worth the extra cost required to pay for it. The cost-effectiveness of pembrolizumab is evaluated for the typical/average patient via modelling that uses short-term trial data to predict clinical effectiveness (efficacy) and costs over 40 years. The challenges of modelling average lifetime outcomes (overall survival, chance of receiving a stem cell transplant that is curative and quality of life) from trial data arise from the short-term nature of data available for pembrolizumab (either from clinical trials or collected in the CDF). Additional challenges also arise because there are limited data available (from either clinical trials or NHS clinical practice) to estimate these average lifetime outcomes for patients treated with standard of care in the NHS without pembrolizumab. The survival data used in the model for pembrolizumab is based on data collected on outcomes with pembrolizumab during its time in the CDF. As there are no clinical trial data for timepoints

after this, survival estimates must be based on extrapolation (using statistical methods to predict how many patients will survive over the time beyond the available data). Different extrapolations were used for people in the model who had been cured and people who cannot have SCT or whose cancer returned after an SCT.

As KEYNOTE-087 is a single-arm trial i.e. only included pembrolizumab and no chemotherapy arm, we had to estimate by how much pembrolizumab improves length of life compared with standard of care by comparing results from the pembrolizumab trial to those for standard of care from similar trials that had been conducted previously. This is a process called "indirect comparison." A range of different sources were used to approximate survival on standard of care. Each of these sources had limitations, which created some uncertainty around cost-effectiveness results. We explored this uncertainty by looking at cost-effectiveness results across the full range of sources, many of which are likely to underestimate the effectiveness of pembrolizumab. We also asked clinical experts to estimate survival on standard of care after 4 years so we could see whether the survival estimates at 4 years predicted by the indirect comparison matched clinical expectations.

There were also no data available to estimate the proportion of people who would receive SCT on standard of care or the proportion of SCTs for patients treated with pembrolizumab or standard of care that would result in patients being cured. Therefore, we asked clinical experts to estimate these values and used them as inputs in our model.

Quality of life data are available from both the KEYNOTE-087 and KEYNOTE-204 trials of pembrolizumab. KEYNOTE-204 shows that pembrolizumab improves the quality of life of patients compared with BV treatment. As no data are available comparing pembrolizumab directly with the relevant standard of care for this evaluation, the QoL improvement versus BV from KEYNOTE-204 was used in the model. As having an SCT is an intensive experience (as outlined in section 2d) the model includes a decrease in modelled patients' quality of life related to undergoing an SCT procedure.

Differences in costs in the model are driven by the cost of pembrolizumab and the cost of stem-cell transplants. Differences in QALYs gained are largely driven by which indirect comparison is used to estimate the difference in survival on pembrolizumab and standard of care, the higher proportion of patients who can receive successful SCTs when treated with pembrolizumab compared with standard of care and whether it is assumed that the effectiveness of pembrolizumab decreases over time.

NICE's new health technology evaluation manual uses a severity modifier to see if the treatment could qualify for a higher willingness to pay threshold. The severity modifier determines a weight which can be assigned to the QALYs accrued by the treatments. The severity modifier depends on the current standard of care.[16]

Given that survival and quality of life and survival outcomes for patients on standard of care are much poorer when compared with the general population of a similar age, a severity modifier of 1.2 is likely to apply for this condition, which means NICE can consider a higher threshold for pembrolizumab to be cost-effective.

The company has calculated a base case cost per QALY gained that is substantially below the NICE's cost-effectiveness threshold of £20,000-£30,000 per QALY gained. But, the company's calculation does not account for potential confidential discounts of drugs that are included in the comparator group. The Committee will discuss how the assumptions made by the company to get their ICER match with what happens in the NHS. To address the uncertainties caused by the limited data available for standard of care a significant number of scenario analyses that use different methods and data sources are presented. Some increase or decrease the cost per QALY a small amount but pembrolizumab is cost-effective across all key scenarios. This demonstrates that despite the uncertainties created by the limited data available pembrolizumab is highly likely to be within the range that NICE normally considers an acceptable use of NHS resources.

3j) Innovation

NICE considers how innovative a new treatment is when making its recommendations.

If the company considers the new treatment to be innovative please explain how it represents a 'step change' in treatment and/ or effectiveness compared with current treatments. Are there any QALY benefits that have not been captured in the economic model that also need to be considered (see section 3f)

In the original evaluation a clinical expert explained that there is an unmet need for treatment to allow people with disease that has not responded or relapsed after BV, and who cannot have autoSCT, to have alloSCT, which is potentially curative. Subsequent data from the CDF have confirmed this view, with patients receiving either autoSCT or alloSCT following treatment with pembrolizumab. The data also show that some patients may receive an SCT after progressing on pembrolizumab and receiving a subsequent line of chemotherapy. Clinical experts have suggested that this may be due to pembrolizumab treatment allowing patients to become sensitive to chemotherapy treatments which may have not previously worked for them.

In the original evaluation the NICE Appraisal Committee noted that there were no additional benefits of treatment with pembrolizumab that had not already been captured by the economic model.[1] MSD note that this is also likely to be the case for the new economic model developed for this review.

3k) Equalities

Are there any potential equality issues that should be taken into account when considering this condition and this treatment? Please explain if you think any groups of people with this condition are particularly disadvantaged.

Equality legislation includes people of a particular age, disability, gender reassignment, marriage and civil partnership, pregnancy and maternity, race, religion or belief, sex, and sexual orientation or people with any other shared characteristics

More information on how NICE deals with equalities issues can be found in the NICE equality scheme

Find more general information about the Equality Act and equalities issues here

No equality issues are anticipated.

SECTION 4: Further information, glossary and references

4a) Further information

Feedback suggests that patients would appreciate links to other information sources and tools that can help them easily locate relevant background information and facilitate their effective contribution to the NICE assessment process. Therefore, please provide links to any relevant online information that would be useful, for example, published clinical trial data, factual web content, educational materials etc. Where possible, please provide open access materials or provide copies that patients can access.
In oncology clinical trials, the severity of adverse events are usually graded according to US National Cancer Institute's AE Severity Grading Scale - Common Terminology Criteria for Adverse Events (CTCAE).[17] CTCAE can also be used to grade the AE for non-oncology studies, but generally not appropriate for studies using healthy volunteers. • Grade 1 Mild; asymptomatic or mild symptoms; clinical or diagnostic observations only; no intervention indicated • Grade 2 Moderate; minimal, local or non-invasive intervention indicated; limiting age-appropriate instrumental ADL • Grade 3 Severe or medically significant but not immediately life-threatening; hospitalisation or prolongation of hospitalisation indicated; disabling; limiting self-care activities of daily living (e.g. bathing, dressing or feeding). • Grade 4 Life-threatening consequences; urgent intervention indicated. • Grade 5 Death related to AE. Further information on NICE and the role of patients**:** • Public Involvement at NICE Public involvement | NICE and the public | NICE Communities | About | NICE • NICE's guides and templates for patient involvement in HTAs Guides to developing our guidance | Help us develop guidance | Support for voluntary and community sector (VCS) organisations | Public involvement | NICE and the public | NICE Communities | About | NICE • EUPATI guidance on patient involvement in NICE: https://www.eupati.eu/guidance-patient-involvement/ • EFPIA – Working together with patient groups: https://www.efpia.eu/media/288492/working-together-with-patient-groups-23102017.pdf • National Health Council Value Initiative. https://nationalhealthcouncil.org/issue/value/ • INAHTA: http://www.inahta.org/ • European Observatory on Health Systems and Policies. Health technology assessment - an introduction to objectives, role of evidence, and structure in Europe: http://www.inahta.org/wp-content/themes/inahta/img/AboutHTA_Policy_brief_on_HTA_Introduction_to_Objectives_Role_of_Evidence_Structure_in_Europe.pdf

4b) Glossary of terms

Response: **Allogeneic stem cell transplant** - A donor stem cell transplant which aims to replace bone marrow that is no longer working properly with healthy stem cells from another person

Autologous stem cell transplant – A stem cell transplant where the patient's own healthy blood stem cells are used to replace blood cells destroyed by high doses of chemotherapy

Arthralgia - Pain in your joints.

Diarrhoea - Loose, watery stools three or more times a day.

Extrapolation - the action of estimating or concluding something by assuming that existing trends will continue or a current method will remain applicable

Fatigue - tired, weak feeling of the whole body, feeling tired all over.

Hypothyroidism - When your thyroid makes too much thyroid hormone.

Indirect comparison - A method of estimating the effectiveness of two treatments which have not been directly compared in a head to head clinical trial

Nausea - When you have an upset stomach or feel like throwing up.

Prognosis - the likely course of a medical condition

Pyrexia - A body temperature that is higher than normal. Also called fever.

QALY – A measure used to quantify the effectiveness of a health intervention that takes into account both length and quality of life

Upper respiratory tract infection - infections of parts of your body involved in breathing, such as the sinuses, and throat. Types of upper respiratory tract infection include the common cold, sinus infection, tonsillitis and laryngitis.

Vomiting - To throw up

4c) References

Please provide a list of all references in the Vancouver style, numbered and ordered strictly in accordance with their numbering in the text:

Response:

1. NICE. 2018. Pembrolizumab for treating relapsed or refractory classical Hodgkin lymphoma. Last accessed: 17 May 2023. Available from: https://www.nice.org.uk/guidance/ta540/chapter/1-Recommendations.

2. European Medicines Agency. 2021. Keytruda: Assessment report. Last accessed: 19 July 2023. Available from: https://www.ema.europa.eu/en/documents/variation-report/keytruda-h-c-3820-ii-0090-epar-assessment-report-variation_en.pdf.

3. NHS. 2021. Hodgkin lymphoma. Last accessed: 17 September 2023. Available from: https://www.nhs.uk/conditions/hodgkin-lymphoma/.

4. Cancer Research UK. 2020. Hodgkin lymphoma. Last accessed: 21 July 2023. Available from: https://www.cancerresearchuk.org/about-cancer/hodgkin-lymphoma.

5. Anthony Nolan. 2023. Hodgkin lymphoma. Last accessed: 17 September 2023. Available from: https://www.anthonynolan.org/patients-and-families/blood-cancers-and-blood-disorders/what-blood-cancer/hodgkin-lymphoma.

6. Cancer.Net. 2022. Lymphoma - Hodgkin: Follow-Up Care. Last accessed: 17 September 2023. Available from: https://www.cancer.net/cancer-types/lymphoma-hodgkin/follow-care.

7. Lymphoma Action. 2023. Donor (allogeneic) stem cell transplants. Last accessed: 17 September 2023. Available from: https://lymphoma-action.org.uk/about-lymphoma-treatment-lymphoma-stem-cell-transplants/donor-stem-cell-transplants.

8. Cancer Research UK. 2023. About side effects of chemotherapy. Last accessed: 17 September 2023. Available from: https://www.cancerresearchuk.org/about-cancer/treatment/chemotherapy/side-effects/about.

9. Lymphoma Action. 2023. Self (autologous) stem cell transplants. Last accessed: 17 September 2023. Available from: https://lymphoma-action.org.uk/about-lymphoma-treatment-lymphoma-stem-cell-transplants/self-stem-cell-transplants.

10. Ijaz A, Khan AY, Malik SU, Faridi W, Fraz MA, Usman M, et al. Significant Risk of Graft-versus-Host Disease with Exposure to Checkpoint Inhibitors before and after Allogeneic Transplantation. *Biol Blood Marrow Transplant* 2019;25(1):94-9.

11. Blood Cancer UK. 2023. After blood cancer treatment ends. Last accessed: 17 September 2023. Available from: https://bloodcancer.org.uk/understanding-blood-cancer/treatment/blood-cancer-after-treatment/.

12. Medicines and Healthcare products Regulatory Agency. KEYTRUDA® Summary of Product Characteristics [Available from: https://products.mhra.gov.uk/search/?search=keytruda&page=1&doc=Spc&rerouteType=0.

13. European Medicines Agency. European public assessment report (EPAR): Keytruda, pembrolizumab [Available from: https://www.ema.europa.eu/en/medicines/human/EPAR/keytruda.

14. National Institute for H, Care E. Pembrolizumab for advanced melanoma not previously treated with ipilimumab [TA366]. Available from: https://www.nice.org.uk/guidance/ta366. [Access Date: 11 January 2023]. 2015.

15. electronic Medicines Compendium. 2022. KEYTRUDA. Last accessed: 06 April 2023. Available from: https://www.medicines.org.uk/emc/product/2498.

16. NICE. 2022. NICE health technology evaluations: the manual. Last accessed: 23 August 2023. Available from: https://www.nice.org.uk/process/pmg36/resources/nice-health-technology-evaluations-the-manual-pdf-72286779244741.

17. National Cancer Institute. 2017. Common Terminology Criteria for Adverse Events (CTCAE) Version 5.0 Last accessed: 17 September 2023. Available from: https://ctep.cancer.gov/protocoldevelopment/electronic_applications/docs/CTCAE_v5_Quick_Reference_8.5x11.pdf.

Confidential

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single technology appraisal

Pembrolizumab for treating relapsed or refractory classical Hodgkin's lymphoma after brentuximab vedotin [review of TA540]

[ID5084]

Clarification questions

October 2023

File name	Version	Contains confidential information	Date
		Yes/no	

Confidential

Notes for company

Highlighting in the template

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

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

Section A: Clarification on effectiveness data

Literature searches

A 1. Priority question: The Evidence Assessment Group (EAG) noted that the previous search strategies utilised in the 2017 systematic literature review (SLR) were not rerun for this update. In the company submission (CS), the company explained that "*The target population was based on the approved indication from NICE Technology Appraisal (TA) 540,(1) that is, adults with R/R cHL who have had BV and cannot have autoSCT. Pembrolizumab is already recommended by NICE as a treatment option for patients with no prior exposure to BV and who have either failed autoSCT or cannot undergo autoSCT but have had at least two previous therapies (TA772).(2) Therefore, evidence outside the specific target population of TA540 was not required*". This updated search appears to carry a number of limitations:

- **Removal of conference proceedings from the Embase strategy despite additional named conference searches performed using Northern Light conference database and hand searching.**

- **Restriction of papers to English only, despite the company's justification the EAG feels that limiting the results to only studies published in English**

Confidential

may introduce 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"* (Morrison, 2012) and that *"research related to language bias supports the inclusion of non-English studies in systematic reviews"*. (Egger, 1997; Lefebvre, 2022)

- **Lack of synonyms for the term relapsed additional terms to consider: resist$ or persist$ or return$ or reocur$ or reoccur$ or recurren$ or recidiv$ or regenerat$**

- **The increased number of facets utilised in the update searches whilst justified by the change in scope raises issues of specificity versus sensitivity. The more specific the search strategy the greater the chance of missing relevant papers that don't mention your search terms directly in the title and abstract.**

- **Line #20 for terms related to Post chemo etc in title and abstract (later combined with terms for brentuximab OR stem cell transplant) feels particularly restrictive and contains some redundant terms. The EAG would recommend that this facet be removed.**

Given the lack of relevant papers found, the EAG would request that the above searches be rerun and expanded with the above points in mind and the resulting new papers screened for includes. Given that the previous searches for the original submission were conducted in 2017 and the searches already undertaken the EAG would suggest the inclusion of a date limit of 2017 onwards.

MSD understand the Evidence Assessment Group's (EAG's) concern about removing conference proceedings from the Embase strategy. However, Embase frequently generates a high number of irrelevant results from conference abstracts. To maintain the precision of our initial search and minimise the volume of non-relevant records, we opted to exclude conference proceedings in favour of additional searches in the Northern Light conference database, augmented by manual hand searches to capture any relevant data from conference proceedings. The searches were intentionally

Confidential

designed to ensure that no relevant conference data were omitted and, as such, we do not believe rescreening these records would identify any additional studies.

Regarding the restriction to publications in only English language, MSD recognise that current best practice suggests attempting to identify and include reports of studies irrespective of language. However, given the convention of international conferences and journals publishing in English, we consider missing a study published in non-English language has a low probability of relevance to the decision problem. Furthermore, the language to describe eligibility for autoSCT, a key population inclusion criterion, is not uniform, even within English-language studies; therefore, identifying studies in such a population published in other languages presents a high degree of difficulty. Although it is not possible to estimate the exact number of non-English records that were excluded due to how the restriction was applied in the searches, it is anticipated to be fewer than 80 records. Given the likely small number of studies excluded for non-English language, taken together with the above considerations, MSD deem that rescreening the records would not yield any additional relevant studies.

For the remaining critiques on the search strategies, based on confirmatory hand searches of relevant materials (e.g., published systematic reviews, narrative reviews, and treatment guidelines), applying the suggested changes would not be expected to yield additional studies and MSD consider it unnecessary to rerun the searches as proposed by the EAG.

A 2. Priority question: The PRISMA flow diagram (Appendix D, Figure 1) shows results for Trials registries Clinical Trials.gov and EUCTR. Please provide full details of the search strategies used.

The following search terms were used to identify relevant studies from trial registries:

- US National Institutes of Health Clinical Trial Registry (http://www.clinicaltrials.gov): “Classical hodgkin lymphoma”;
- European Union Clinical Trials Register (https://www.clinicaltrialsregister.eu): “Classical hodgkin lymphoma”.

A 3. Priority question: Appendix G, table 37 provides a copy of the combined updated search strategy for economic evaluation, health-related quality of life (HRQoL) and cost and resource use data. However, the EAG has concerns regarding the rationale behind the date limits used in lines #76 (#22 AND #74 AND [01-07-2017]/sd NOT [13-01-2023]/sd) and #147 (#22 AND #145 AND [01-07-2017]/sd NOT [13-01-2023]/sd. Given that these searches were run on 20th February 2023, both lines would appear to discard results added to the database since 13th January 2023. Please confirm is this is the case and if so, rerun these searches and screen the previously discarded records.

MSD apologise for the error in reporting of the searches. MSD confirm that the date limit applied in the searches for economic evaluations, health-related quality of life (HRQoL) studies, and cost and resource use data was 20-02-2023. MSD ran a test search on 12th January 2023 and the final search on 20th February 2023. The number of records reported in Appendix G are for the searches carried out on 20th February. Please see a screenshot below from Embase.com showing the history of the saved searches (Figure 1).

Figure 1. Record of search on Embase.com

	Query	Name	Last results	Last updated	Created
☐ #147	#22 AND #145 AND [01-07-2017]/sd NOT [20-03-2023]/sd	KN087 economic u pdate	949	2023-02-20	2023-01-12
☐ #146	#22 AND #145	KN087 economic u pdate	1,984	2023-02-20	2023-01-12
☐ #145	#77 OR #78 OR #79 OR #80 OR #81 OR #82 OR #83 OR #84 OR #85 OR #86 OR #87 OR #88 OR #89 OR #90 OR #91 OR #92 OR #93 OR #94 OR #95 OR #96 OR #97 OR #98 OR #99 OR Show full	KN087 economic u pdate	1,579,816	2023-02-20	2023-01-12

A 4. Section G1.3 of the CS states “Due to the lack of new emerging evidence in the small population of interest, hand-searching of additional publications from conferences and grey literature was unlikely to identify new evidence and therefore was not conducted.” However, this does not explain the decision not to include searches of Pubmed or the Cochrane databases in the update searches. Please explain the rationale behind this and the effect that this may have had on the recall of results.

Confidential

The search of Medline for published cost effectiveness studies was carried out via the Embase.com platform and the researcher selected “Medline in process” as a source of records. Thus, MSD considered it unnecessary to search “Medline in process” through Pubmed. MSD decided against searching the Cochrane database, because of the high probability that no unique study evaluating cost would be retrieved from the Cochrane database. Since the cessation of updating of the NHS Economic Evaluation Database (NHS EED) at the end of March 2015, which was hosted on the Cochrane platform, the Cochrane Database is primarily a repository for records of systematic reviews and randomised controlled trials.

Decision problem

A 5. Priority question: The scope, as defined by the National Institute for Health and Care Excellence (NICE), is “People with relapsed or refractory classical Hodgkin lymphoma who have had brentuximab vedotin [BV] and cannot have autologous stem cell transplant”. The company further clarifies the population as autoSCT naïve. Eligibility for autoSCT can change according to the overall status of the patient. Please confirm that the scope of the decision problem is patients who have not previously had autoSCT and remain ineligible.

MSD confirm that the population relevant to the decision problem is adults who have not had autologous stem cell transplant (autoSCT) and remain ineligible for autoSCT. As described in Table 1 of the company submission (CS), in TA540,[(1)] pembrolizumab was not recommended for treating relapsed or refractory classical Hodgkin lymphoma (R/R cHL) in adults who had received autoSCT and brentuximab vedotin (BV) and, thus, the population of interest to this STA are those who are transplant naïve. Additionally, the patient population is equivalent to those patients treated with pembrolizumab during its time in the Cancer Drugs Fund (CDF). All patients were deemed ineligible for autoSCT at the point of treatment initiation. However, some patients could become suitable candidates for SCT, should treatment induce a sufficient level of response. Because of older age and presence of comorbidities, some patients would be deemed not suitable for SCT at any time, be that either autologous or allogeneic SCT, and would continue treatment with pembrolizumab up to 35 cycles as long as clinical benefit remains.

Confidential

A 6. Priority question: The inclusion criteria for the systemic anti-cancer therapy (SACT) data states that the eligible population "Has not received SCT of any kind and is ineligible for SCT; Patient is either a candidate for future SCT if there is sufficient benefit from pembrolizumab, or not a candidate for SCT however good the response to pembrolizumab may be". Please clarify whether the population in the decision problem, the SLR and the SACT is transplant-naïve, people who are ineligible for autoSCT at any point including after responding to treatment, or if it includes those who might become eligible for SCT (stem cell transplant), whether auto or alloSCT, after responding to treatment and discuss the potential inconsistency with the NICE scope.

As above for A5.

MSD confirm that the population relevant to the decision problem is adults who have not had autologous stem cell transplant (autoSCT) and remain ineligible for autoSCT. As described in Table 1 of the company submission (CS), in TA540,[(1)] pembrolizumab was not recommended for treating relapsed or refractory classical Hodgkin lymphoma (R/R cHL) in adults who had received autoSCT and brentuximab vedotin (BV) and, thus, the population of interest to this STA are those who are transplant naïve. Additionally, the patient population is equivalent to those patients treated with pembrolizumab during its time in the Cancer Drugs Fund (CDF). All patients were deemed ineligible for autoSCT at the point of treatment initiation. However, some patients could become suitable candidates for SCT, should treatment induce a sufficient level of response. Because of older age and presence of comorbidities, some patients would be deemed not suitable for SCT at any time, be that either autologous or allogeneic SCT, and would continue treatment with pembrolizumab up to 35 cycles as long as clinical benefit remains.

A 7. Priority question: According to NICE TA guidance 524, "brentuximab vedotin is recommended for relapsed or refractory CD30+ Hodgkin lymphoma after autologous stem cell transplant, or after at least 2 prior therapies when autologous stem cell transplant or multi-agent chemotherapy is not a treatment option". Please confirm that the population addressed by the decision problem is a subset of the above

Confidential

recommendation: patients that have had brentuximab vedotin after at least 2 prior therapies when autologous stem cell transplant or multi-agent chemotherapy is not a treatment option.

That is correct. The population in the decision problem are those who have had treatment with BV after at least two prior therapies (per NICE TA524) and never had an SCT.

A 8. Priority question: The NICE scope defines the comparators to be single or combination chemotherapy including drugs such as gemcitabine, vinblastine and cisplatin and best supportive care (BSC). The list of comparators the company is reporting in Table 1 is not completely aligned e.g., the drugs vinblastine and cisplatin are not included in the list of comparators. The company justified its basket of comparators based on the study by Cheah et al. 2016. However, the TA540 final appraisal document (FAD) stated that this had limited application partly because it was conducted in the United States of America (USA) and partly because 70% of patients received SCT. The Eyre et al. 2017 study, which was conducted in the United Kingdom (UK), was mentioned in the FAD as an alternative, albeit also with some limitations.

a) Please provide further justification for the inconsistency in comparators with the NICE scope.

b) Please clarify if the basket of comparators was only from the 30% of patients who had not received a SCT. If not, then please adjust the comparators accordingly.

c) Please provide the list of comparators that would be standard of care (SoC) in the UK with percentage use based on objective evidence.

d) Please redo all analyses using the list of comparators based on Eyre et al. 2017.

The approach to determining the basket of comparators that represent standard care is explained in the CS, but briefly; we listed comparators from all the relevant sources (TA462, TA540, Eyre, Cheah) with no consideration of subgroups. The eight advisors

Confidential

at the UK advisory board then excluded or added options so that the final list reflected those that are available and in use in UK clinical practice. There are no guidelines, practice is heterogeneous and also depends on patient level factors. The advisors were not able to provide percentages given that pembrolizumab has been the SoC for a number of years and there are only ~50 patients per year nationally in this treatment line. Prior to the introduction of pembrolizumab many patients went into clinical trials. There is no objective evidence available so we felt that an even split was the starting point that made the fewest explicit assumptions. The blend of comparators is only relevant for determining the cost in the model and so is best covered using sensitivity analyses rather than explicit breakdowns that are based on indirectly applicable populations in the literature.

A 9. Priority question: The list of comparators provided by the company in Table 1: *"Investigational agent; Gemcitabine; Bendamustine; Other alkylatory; BV retreatment; Platinum based; AutoSCT; Other"*, which is taken from Cheah et al. 2016 is not the same list of drugs used in the blended comparator for the economic model reported in Table 50. Table 50 includes further drugs: ICE (ifosfamide, carboplatin, etoposide); PMitCEBO (prednisolone, mitoxantrone, cyclophosphamide, etoposide, bleomycin and oncovin); DECC (dexamethasone, etoposide, chlorambucil, lomustine); radiotherapy and Mini-BEAM (carmustine, etoposide, cytarabine, melphalan). These additional drugs are said to have been informed by Eyre et al. 2017 and expert opinion. Some of these drugs (ICE, IGEV, GEM-P) were reported in Eyre et al. 2017 to have been administered pre-BV treatment while others are not reported in Eyre et al. 2017 or the expert opinion report at all (GDP, GVD). Please elaborate.

As above for A8:

The approach to determining the basket of comparators that represent standard care is explained in the CS, but briefly; we listed comparators from all the relevant sources (TA462, TA540, Eyre, Cheah) with no consideration of subgroups. The eight advisors at the UK advisory board then excluded or added options so that the final list reflected those that are available and in use in UK clinical practice. There are no guidelines, practice is heterogeneous and also depends on patient level factors. The advisors

Confidential

were not able to provide percentages given that pembrolizumab has been the SoC for a number of years and there are only ~50 patients per year nationally in this treatment line. Prior to the introduction of pembrolizumab many patients went into clinical trials. There is no objective evidence available so we felt that an even split was the starting point that made the fewest explicit assumptions. The blend of comparators is only relevant for determining the cost in the model and so is best covered using sensitivity analyses rather than explicit breakdowns that are based on indirectly applicable populations in the literature.

A 10. Priority question: The company states that BSC (defined as *"no active treatment"*) was not included in the CS as according to feedback from clinicians it is not a comparator of interest. Please provide evidence to support this statement.

Clinical experts engaged by MSD stated during an advisory board that, in the absence of pembrolizumab as a treatment option, active chemotherapy would typically be given as a 4th line therapy.(2) Clinical experts went on to comment that 5th line treatment (after either pembrolizumab or standard of care) would commonly consist of oral palliative chemotherapy (e.g., DECC), with radiotherapy, single agent gemcitabine or bendamustine and participation in clinical trials also potential options. Based on this advice, MSD considered it appropriate to exclude BSC (defined as "no active treatment") as a comparator in the economic model. However, we contacted a clinician for information during responding to CQs who said that up to 10% of patients might actually get BSC. A sensitivity analysis could be done on the model that reduced SoC costs by 10%, which would very slightly increase the ICER.

Systematic review

A 11. Priority question: The eligibility criteria for the SLR included: *"patients…who have previously received BV and cannot undergo auto-SCT"*. However, the Cheah study included only a subgroup that might fulfil this criterion. Please clarify that no studies were excluded that might have had a subgroup that fulfilled this population criterion. If that is not the case, please redo the SLR and include all relevant data including those reported as a subgroup of the main population.

Confidential

MSD confirm no study with relevant subgroup data was excluded.

A 12. Priority question: The study by Hanel et al. 202) was excluded with population given as the reason for exclusion. This was a Phase 2 Trial of Ibrutinib and Nivolumab in Patients with Relapsed or Refractory cHL. 76.5% of the participants had prior brentuximab therapy, while 47.1% had prior autoSCT.

a) Why was this study not included in the SLR?

58.8% of patients had prior nivolumab. Patients previously treated with nivolumab are outside of the population in the decision problem. There are only 7 patients in the no prior nivolumab group and baseline characteristics are not reported for this group (so it is not clear how many of these patients received brentuximab). Therefore, the study was excluded as a subgroup that aligns with the decision problem cannot be identified.

b) Given the lack of alignment of the population in Cheah 2016 as well as the limitations of an observational study, the company is asked to consider running the ITC analysis using Hanel 2023 instead.

As noted in the response to part a, it is not possible to identify the relevant subgroup from the above data. Additionally, the study only reports PFS results and not OS. Therefore, it is not possible to conduct an ITC analysis using the data from Hanel 2023.

A 13. Priority question: In Table 63 of Appendix M, the study NCT02824029 of Ibrutinib is reported as an ongoing study of Ibrutinib.

a) How was this study retrieved? It is not reported in Table 7 of Appendix D.

b) This study includes patients with relapsed or refractory HL who have failed at least 2 lines of prior therapy and are not eligible for autologous stem cell transplant. Why is the company claiming that the population in this study does not align with the population relevant to the decision problem?

Confidential

c) Please include any data reported by this study in the indirect treatment comparison (ITC) analysis as well as in the adverse events section.

The study described in Appendix M is an ongoing study that could provide additional evidence in the next 12 months. A search for ongoing studies was carried out to fulfil the request detailed in Section B.2.11 of the NICE STA template. A simple search of ClinicalTrials.gov was carried out, using the population term of classical Hodgkin lymphoma, and applying the restriction of studies scheduled to report in the subsequent 12 months. Retrieved records were evaluated further to identify studies relevant to the decision problem. The identified study is a single arm trial of ibrutinib, and MSD are not aware of any results being published. At the time of writing, ibrutinib is not part of established clinical practice in the UK for the management of R/R cHL, therefore, had data been available, they would not have been used to address the decision problem. However, MSD considered that the study could be relevant in the future, should ibrutinib be evaluated through the NICE TA process.

A 14. Priority question: In the eligibility criteria reported in Table 5 of the Appendices:

a) Only the records in English language were to be included. This exclusion criterion goes against best practise guidelines (e.g., ROBIS). Please do not apply this exclusion criterion after running the requested updated searches (see Literature searches section).

b) The interventions: Pembrolizumab; Single or combination chemotherapy including drugs such as gemcitabine, vinblastine, and cisplatin are only included and any *"interventions not listed"* are to be excluded. Please remove this criterion for the updated searches since an exhaustive list of interventions in not provided in the NICE scope.

Regarding the English language restriction, please see MSD's response to A1.

Specific intervention terms were not added to the search strategies to maximise the number of records retrieved on the available evidence base. Rescreening of studies excluded for "interventions" did not yield any additional relevant studies.

Confidential

Clinical effectiveness evidence

A 15. Priority question: A SACT data set is presented for pembolizumab.

a) Does the company have access to SACT data for the comparators of interest? If so, this should be shared.

b) Are there available historical data for the comparators of interest?

c) Please conduct the indirect treatment comparison (ITC) analysis using the above data.

MSD do not have access to either contemporary or historical data from SACT for the comparators of interest, and is unaware of any such data being published. The data made available to the company from the SACT are exclusively for patients treated with pembrolizumab as detailed in section 7.1 of the managed access agreement for pembrolizumab in the CDF (data analysis plan). Individual patient level data were not provided to MSD. Digitisation was necessary for the secondary survival analysis used in the economic model. As the company does not have access to the data highlighted in A15, it is not possible to conduct an ITC.

A 16. Priority question: According to Figure 2 of the CS, pembrolizumab is placed in the 5th line of therapy. The company states that 78 of the patients (96.3%) in KEYNOTE-087 had ≥3 previous lines of therapy. Since the outcomes and specifically the survival values *"are likely to be substantially lower in the context of the later lines of therapy"*, please provide a breakdown for each previous line of therapy as well as a subgroup analysis by number of prior lines of therapy.

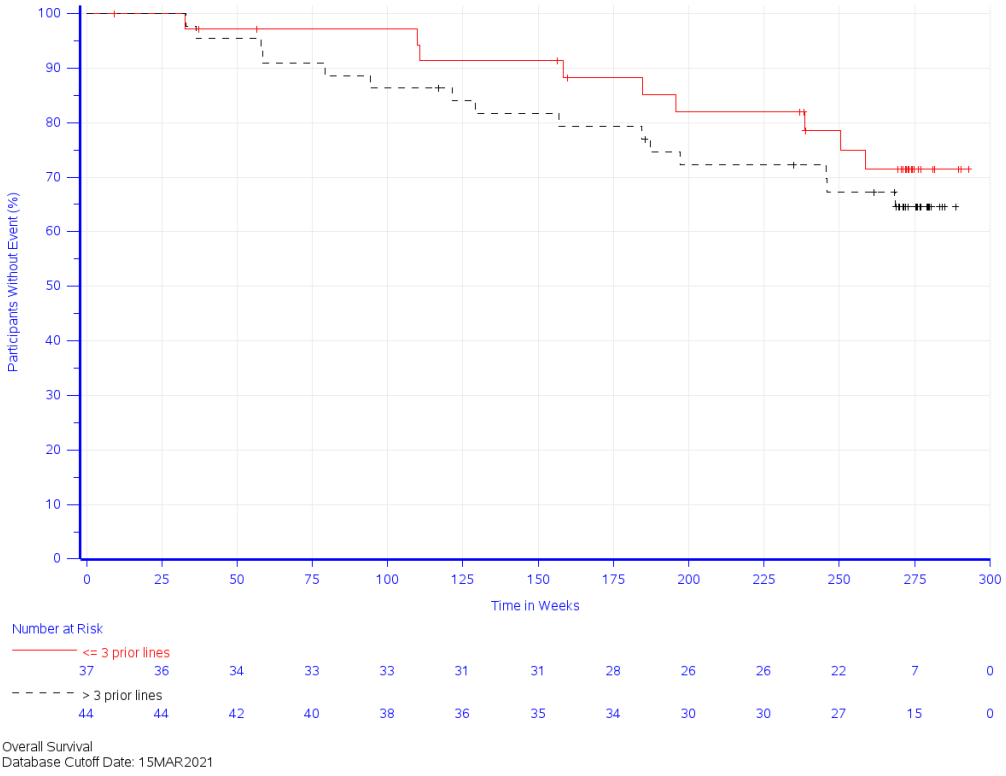
Figure 2 positions pembrolizumab in the 4th line of therapy for those who are autoSCT naive. As discussed in section B1.3.4.1 of the submission, the pathway to treatment with pembrolizumab of relevance to the decision problem is:

- 1st line: Chemotherapy (ABVD or BEACOPP)
- 2nd line: Salvage chemotherapy and/or radiotherapy
- 3rd line: Brentuximab vedotin

- 4th line: pembrolizumab

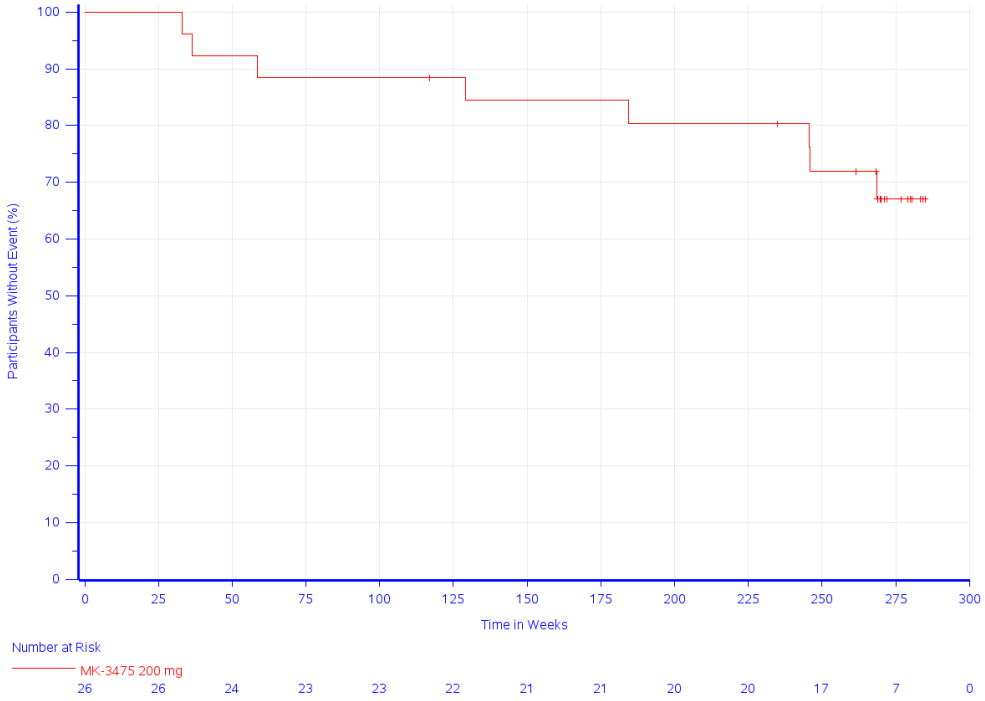
The positioning of pembrolizumab as 4th line of therapy was validated by clinicians at an advisory board.(2) Please see below Kaplan–Meier curves (Figure 2, Figure 3, and Figure 4) for overall survival (OS) for Cohort 2 from KEYNOTE-087 based on number of lines of prior therapy.

Figure 2. KM plot for OS for Cohort 2 by ≤3 vs >3 prior lines of therapy



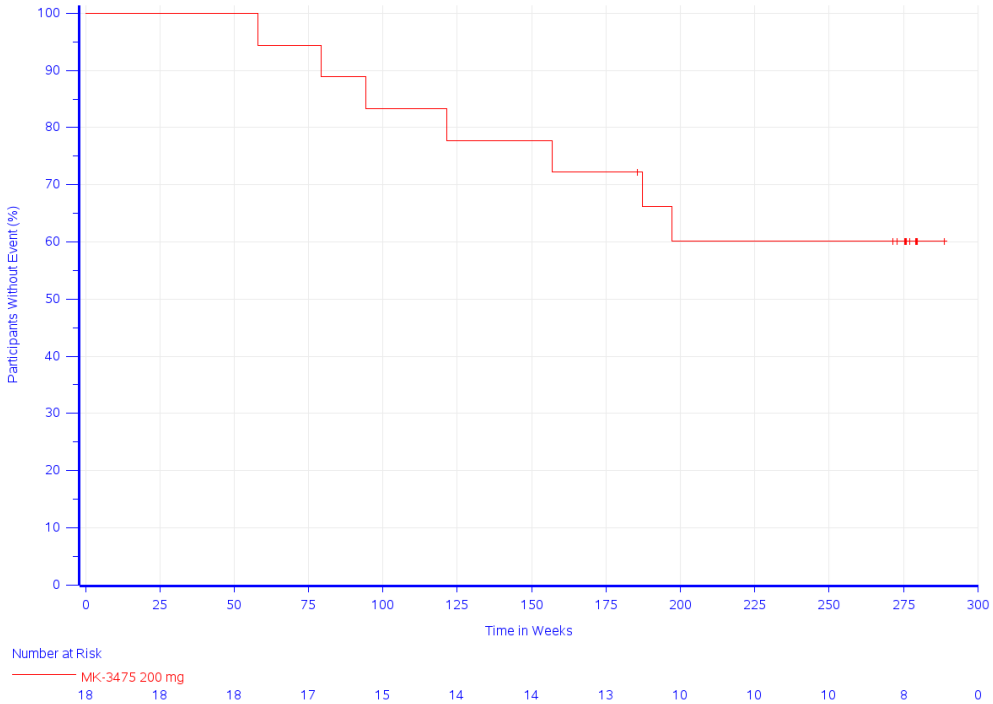
Overall Survival
Database Cutoff Date: 15MAR2021

Figure 3. KM plot for OS for Cohort 2 with 4 prior lines of therapy



Overall Survival
Database Cutoff Date: 15MAR2021

Figure 4. KM plot for OS for Cohort 2 with ≥5 prior lines of therapy



Overall Survival
Database Cutoff Date: 15MAR2021

Confidential

A 17. Priority question: Table 10 of the CS, which presents the baseline characteristics of the SACT cohort, does not present previous lines of therapy. Please provide this information and conduct a subgroup analysis to this effect.

Baseline characteristics provided by NHSE for the SACT cohort have been reported in full and do not include previous lines of therapy (see Section B2.3.1.6, table 10, page 42), therefore it is not possible to conduct this analysis. As noted in MSD's response to A16, clinical experts validated that the proposed positioning of pembrolizumab in the submission reflects clinical practice in the NHS in England. MSD note that no subgroup analysis was identified in the NICE scope.[3]

A 18. Priority question: According to the CS, "*Of the 215 people forming the SACT dataset, 132 (61%) patients were identified in Blueteq as being suitable candidates for SCT*". In fact, 65 (30.2%) patients went on to receive SCT; 23 [35.4%] receiving autoSCT vs. 42 [64.6%] receiving alloSCT. This seems to contradict the decision problem of the CS where only patients ineligible for autoSCT are included, as well as the inclusion criteria for patients in the SACT. Please comment on this and justify the inclusion of these participants.

The decision problem relates to patients who remain ineligible for autoSCT after treatment with BV. As noted in MSD's response to A5, in those deemed potential candidates for SCT, the goal of treatment subsequent to BV is, in some cases, to act as a bridge therapy, triggering sufficient response to treatment to facilitate SCT. However, not all those receiving pembrolizumab after BV will be considered a potential candidate for SCT, with age and comorbidities often rendering a patient unfit for SCT, and such patients can benefit from treatment with pembrolizumab. The population enrolled in the SACT cohort align with the decision problem – a proportion were judged suitable candidates for SCT and the remainder in the dataset were those who, prior to treatment initiation, it was thought would never be able to undergo SCT regardless of their response to pembrolizumab.

A 19. Priority question: KEYNOTE-087 had only three sites in the UK out of the 51 included in the trial. Is the position of BV in the treatment pathway comparable between the 48 non-UK and the three UK centres?

Confidential

MSD consider the position of BV in the treatment pathway across the participating centres not to be relevant to the interpretation of the results from KEYNOTE-087 and the SACT dataset. MSD note that, although few sites were located in the UK, patients had to meet strict eligibility criteria to be enrolled into Cohort 2 of KEYNOTE-087 and those entering into the study received treatment with pembrolizumab in line with the protocol. Thus, no patient enrolled in Cohort 2 had undergone SCT, and all patients received pembrolizumab after BV, and were deemed to be ineligible for SCT after BV. Due to the low number of sites in the UK, MSD acknowledged in the CS that we consider the results from the SACT cohort to be more generalisable to the management of R/R cHL at the line of therapy relevant to the decision problem than the results from KEYNOTE-087, which we consider to be strong supporting evidence of the results from the SACT dataset.

A 20. Priority question: In the FAD for TA540 recommendations for data collection the proposals for further data collection included:

- **proportion of people having pembrolizumab who have an alloSCT,**
- **time to alloSCT,**
- **duration of treatment with pembrolizumab before alloSCT,**
- **long-term follow-up of people having pembrolizumab with or without subsequent alloSCT (in particular, collection of data on overall survival).**

Please provide clarifications on how the above recommendations were met and provide further details, if needed.

The above recommendations were met through collection of data in the SACT database. MSD note that the data provided by NHSE included patients who received both alloSCT and autoSCT (rather than alloSCT alone) as both types of SCT were used in clinical practice. Data for time on treatment (ToT) was presented for all patients in the SACT, regardless of whether they received an SCT. The data were reported in the following sections of the company submission:

- proportion of people having pembrolizumab who have an SCT (auto, allo and any SCT) – Section B.2.6.1.3, table 17, page 52;
- time to SCT (any SCT) – Section B2.6.1.3, figure 4, page 54;

Confidential

- duration of treatment with pembrolizumab – Section B3.6.3.1, figure 20, page 137;
- overall survival data;
    - OS for all patients from the SACT dataset – Section B2.6.1.5, figure 7, page 57;
    - OS for patients from the SACT dataset who did not receive an SCT – Section B2.6.1.5, figure 8, page 59.

A 21. Priority question: A large proportion of patients in both KEYNOTE-087 (29.6%) and the SACT data set (30.2%), went on to receive SCT after being treated with pembrolizumab. Please discuss whether pembrolizumab should be considered a bridge therapy for SCT and how this relates to the outcomes presented in the SCT.

As in the Blueteq criteria, patients should be treated with pembrolizumab until they are either able to be bridged to SCT (if suitable for SCT) or until loss of clinical benefit or 35 cycles (if unsuitable for SCT). Assessment of suitability for SCT is a complex and subjective judgement based on a variety of patient-level factors, as is the decision to attempt the procedure.

A 22. Only 10 patients from the UK were included in cohort 2 of KEYNOTE-087. Please discuss the potential implications for patients in England and Wales.

As noted in MSD's response to A19, MSD acknowledge the low representation of patients from England and Wales in KEYNOTE-087 and consider the results from the SACT cohort to be more generalisable to the management of R/R cHL at the line of therapy relevant to the decision problem.

A 23. No data on the prespecified subgroup analysis are presented in Document B, the Appendices nor the clinical study report (CSR) for KEYNOTE-087. Please provide the results of the subgroup analysis for age category (≤65 vs >65 years), sex (female vs male), race (white vs non-white), region (US vs ex-US) and number of prior therapies (<4 vs ≥4).

MSD note that pre-specified subgroup analysis of ORR for Cohort 2 of KEYNOTE-087 are not available. As discussed in Appendix E of the CS, pre-specified subgroup analysis was planned for only ORR. The CSR stipulates that if the observed

Confidential

numbers for a particular subgroup are too small to make a meaningful clinical interpretation, then that subgroup analysis would not be conducted.

Indirect treatment comparison (ITC)

A 24. Priority question: Only three effect modifiers were selected for the matching-adjusted indirect comparison (MAIC) process: age, sex and Eastern Cooperative Oncology Group (ECOG) performance status. According to the CS, *"In patients with R/R cHL, time to initial relapse after high-dose chemotherapy and autoSCT was identified as a key prognostic factor for survival"*, as well as *"time to relapse"* and *"history of primary refractory disease"*.

a) Why have these factors not been included as potential effect modifiers?

b) Please conduct the ITC analysis using the above additional effects modifiers.

The factors highlighted by the EAG as omitted from the MAIC were not reported in all sources used to inform the MAIC and so adjustment was not possible.

A 25. Priority question: There are major differences between the baseline characteristics of the three cohorts used in the ITC (KEYNOTE-087, SACT data set, Cheah et al. 2016). A key difference is that in Cheah et al. 2016 *"before treatment with BV, 70 (72%) people had undergone SCT, predominantly autoSCT (n=66)"*, where in KEYNOTE-087 previous autoSCT was an exclusion criterion. How is this key difference addressed in the ITC?

MSD agree with the EAG, and have discussed the limitations of the analyses within the CS and Appendix D2. For the comparison of the SACT dataset versus Cheah (2016), the adjustment suggested by the EAG could not be carried out because MSD had access to only aggregated data for both studies. When comparing Cohort 2 of KEYNOTE-087 versus Cheah (2016), the aim was to provide a conservative ITC using 'fitter' patients in the control group (Cheah (2016)).

Confidential

A 26. Priority question: Cheah et al. 2016 reports baseline characteristics before treatment with BV and after. Please clarify which data were used in the ITC analysis.

The characteristics used for adjustment were those reported for after disease progression on BV. The methodology is described in full in Appendix D of the submission documents.

A 27. Priority question: According to the CS, "To be eligible for pembrolizumab through the CDF, patients had to have an Eastern Cooperative Oncology Group (ECOG) performance status (PS) of 0 or 1 and not have received prior treatment with an anti-PD-1, anti-PD-L1, anti-PD-L2, anti-CD137, or anti-cytotoxic T-lymphocyte-associated antigen-4 (CTLA-4) antibody" which is similar to the inclusion criteria for KEYNOTE-087 regarding prior treatments. How does this compare to the treatments received by patients in the observational study?

MSD have no details on prior lines of treatment for the SACT cohort, except that they must have satisfied the mentioned criteria to have been treated with pembrolizumab through the CDF.

A 28. Priority question: Two sources of comparative evidence for pembrolizumab vs BV are used in ITC analyses and further used in the cost-effectiveness analysis. These sources are the trial KEYNOTE-204 and the Markov trace from the BV vs SoC model used in NICE TA524.

a) The data from KEYNOTE-204 used in the CS have not been reported publicly. Please provide the CSR for KEYNOTE-204 for the data cutoff in question (16th Jan 2020).

The CSR for KEYNOTE-204 has been provided. As the CSR presents data on an interim analysis that is not intended for publication, MSD request that all data contained in the CSR be considered commercial in confidence.

b) Please provide the full methods used as well as the analysis for the ITC models involving KEYNOTE-204.

Overall survival data from the subgroup of patients who were treated at third line is reported in Appendix P. The hazard ratio (HR) from this analysis was used in a simple

Confidential

Bucher indirect treatment comparison with the HR from NICE TA524 (see response to A27c below). As detailed in section B3.3.1.3 of the CS, the HR for pembrolizumab vs SoC was calculated by multiplying the HR for pembrolizumab vs BV and the HR for BV vs SoC together. The standard error of the log HR was obtained by taking the square root of the sum of the two variances. The upper and lower confidence limits on the natural scale were the exponentials of the limits on the log scale.

c) Please provide the full methods used and the full analysis results for the Markov trace from the BV vs SoC model used in NICE TA524.

The method is already described in full in section B.3.3.1.3 (page 101) of the CS. Briefly, we digitised the OS Markov trace to 4 years, created a new curve that applied a constant HR to the SoC Markov trace and varied this constant HR until the visual fit to the BV Markov trace was optimised. We observed a close visual fit across the 4-year time horizon using a HR of 0.62 and because of this, the proportional hazards assumption appeared to be reasonable. We then confirmed with clinicians at the advisory board that an HR of 0.62 was reasonable in their experience. Of course, it had already been considered reasonable by implication by the NICE Committee assessing BV and recommending BV in TA524.

d) The company is using a model output (Markov trace from the BV vs SoC model used in NICE TA524) as an input to a further model. Please justify this methodology/methods by citing relevant references or guidance.

We were unable to find references establishing a precedent for this method but, conceptually, it is sensible. In the absence of actual data on the SoC (a bundled comparator mostly consisting of generic chemotherapies), the Markov trace from TA524 represents the best available evidence on the treatment effect of BV on OS vs. SoC in R/R cHL. The approach has the advantage that, in an evidence light area, the comparator arm has been drawn from a model based on assumptions that have previously been accepted by NICE. Furthermore, clinicians at an advisory board confirmed that the estimated HR for the comparison between BV and SoC was plausible. MSD acknowledge that this approach is associated with uncertainty, therefore, several alternative scenarios were presented, the strengths and limitations of which are elaborated on further in the CS and our response to B12 below.

Confidential

e) BV is not a comparator defined either in NICE scope or in the company's decision problem. We suggest that you remove this input from any analysis as it is not relevant to this submission.

BV is relevant to this submission in that, prior to the approval of pembrolizumab, it was the standard of care in R/R cHL and comprises one arm of the only source of randomised evidence in this setting. It is also relevant in that "3L+" and "4L+" trials in R/R cHL are not tightly prescriptive by treatment line; for example, 29% of patients in Eyre 2017 were 4L+, Cheah 2016 has a median of 3 but a range of 0 to 9 prior therapies listed, in KEYNOTE-204 the percentage who were 3L vs 4L+ was 21% vs 16%, in KEYNOTE-087 the proportion of patients who were 4L vs 5L+ was 46% vs 54%. Consequently, marketing authorisations in the area typically cover multiple lines. It is also relevant in that BV is considered to be a more effective treatment than standard chemotherapy in R/R cHL.

If pembrolizumab is more effective than BV then it must, by implication, be at least that much more effective than SoC. This would only not be true if failure on BV would alter the patient characteristics such that the treatment effect of pembrolizumab would diminish but the opposite is more likely to be true. In preparing our responses, we asked a clinician to comment on the direction of bias in using the effectiveness of pembrolizumab vs BV in the 3L setting as a surrogate for effectiveness of pembrolizumab vs. SoC in the 4L (BV failed) population and he commented that the observed effectiveness would likely be greater. His reasoning was that a similar proportion of patients would respond well to pembrolizumab whereas a much lower proportion would respond to SoC than had responded to BV. This is because the 4L population is effectively an even more chemotherapy-insensitive population (BV is a chemotherapy based regimen). He commented that failure to respond to BV is unlikely to meaningfully affect a patient's ability to respond to pembrolizumab because it is a different mechanism of action. This is supported by the data from the trials; very similar CR and PR rates to pembrolizumab are observed in the no-prior SCT groups of KN204 and KN087 (CR = 27% and 26%, respectively and PR = 35% and 38%, respectively).

The treatment effect of pembrolizumab vs BV is useful information to incorporate in the economic model because it allows decision makers to examine how various levels of plausible effectiveness affect the cost-effectiveness results using the effectiveness

Confidential

of pembrolizumab vs BV as a reference. KEYNOTE-204 was a 3L+ trial and this setting is 4L+ among patients who have failed BV, this is a source of indirectness of population which biases against pembrolizumab (for the reasons discussed above) rather than total lack of applicability of this evidence.

A 29. Please provide evidence on the likely extent of error due to unaccounted for covariates, in relation to the observed relative treatment effect, as specified in NICE Decision Support Unit (DSU) technical support document (TSD) 18.

MSD consider that it is not possible to provide evidence on the likely extent of error for unaccounted variates. TSD18 advises that the easiest way to quantify residual systematic error introduced from unobserved prognostic variables and effect modifiers is by comparing observed and predicted outcomes for the intervention of interest in several studies in the target population. As acknowledged in TSD18, quantification of the error is not feasible if other studies in the target population are not available. MSD note that KEYNOTE-087 and the data from the SACT cohort are the only data for pembrolizumab in the population relevant to the decision problem, and we consider that it would not be appropriate to attempt to provide an empirical estimate on the extent of error based on results from only two studies. As discussed in the CS, there are differences in the HR for OS associated with pembrolizumab compared with SoC, but, in all ITCs, the direction of effect and the direction of observable bias favours pembrolizumab and the differences reach statistical significance.

Adverse events

A 30. Priority question: The CS reports very high rates of adverse events (AEs) and drug-related AEs in cohort 2 of KEYNOTE-087.

a) How do these results relate to the comparators?

b) How does the type of AEs experienced by the participants of cohort 2 of KEYNOTE-087 relate to the AEs experienced by patients treated with a comparator?

As stated in the CS, in general, pembrolizumab was well tolerated by patients in Cohort 2 of KEYNOTE-087, with a manageable safety profile. The safety profile of pembrolizumab is considered acceptable in the context of alternative therapies, such

Confidential

as standard chemotherapy regimens. Clinicians in the NHS have considerable experience with pembrolizumab in a variety of indications, including cHL. No safety signal was identified in KEYNOTE-087 that differs from the large portfolio of pembrolizumab trials that have already reported. The data presented from KEYNOTE-087 show that most AEs experienced were low grade, and did not result in study discontinuation. We also spoke to a clinician to elicit further information for this CQ response who confirmed that pembrolizumab is typically a much better tolerated treatment than standard chemotherapy regimens. The AEs from comparator studies are available in the cost effectiveness section of the CS, the clinical effectiveness section focuses on the safety profile of pembrolizumab from KEYNOTE-087.

A 31. Priority question: There are no data provided for AEs experienced by patients treated with any of the comparators. Please provide the necessary data and an appropriate analysis and comparison.

These data are provided in the cost-effectiveness section of the CS. Grade 3+ AE rates are estimated separately for all of the comparators and weighed averages are used in the economic model along with utility decrements and durations to estimate QALY losses for each treatment arm. In the base case, QALY losses were -0.0009 for pembrolizumab and -0.0039 for the weighted SoC.

A 32. The company has chosen to present the AEs data for all the cohorts of KEYNOTE-087. Please justify this decision.

MSD consider AE data for Cohort 2 to be the most relevant, particularly in the economic evaluation, but decided to present AE results for all three cohorts within the clinical section to demonstrate the consistency of the safety profile of pembrolizumab across the groups of patients with R/R cHL.

Confidential

Section B: Clarification on cost-effectiveness data

General

B 1. Priority question: The company collected new evidence, in line with the recommendations for data collection outlined in the FAD. A key recommendation was the collection of long-term follow-up data, including overall survival (OS) data. The collected OS data from the SACT show a significantly lower survival rate in SACT versus KEYNOTE-087 at every timepoint (see Table 19 of CS).

a) Please comment on the potential impact on cost-effectiveness, if these OS data were implemented in the original model.

Given time to SCT among patients who actually had an SCT in SACT has been observed at roughly 24 weeks, we assume it is the discussion of the 24-week model that is of interest. The 12-week model also estimates ~30% of patients on SoC get an SCT, which is implausible.

OS from KEYNOTE-087 is only used for the first 24 weeks of the model. Thereafter, patients are divided into those who have SCT and those who do not. OS on SCT is based on a parametric extrapolation from 13 patients and is fairly long at ~15 years in either arm. OS without SCT is effectively based on a parametric extrapolation from PFS (not OS) resulting in a mean undiscounted LYs for this group in the pembrolizumab and SoC arms of 3.6 and 3.1 respectively. After 24 weeks, OS drops very sharply because of this structural choice.

OS in the pembrolizumab arm of the TA540 model is estimated at 43% at 4 years, which is ~10% lower than what was observed in the RWD from SACT and very far below the eventual observed OS in KEYNOTE-087. We conclude that the old model does not have overly optimistic OS estimates for pembrolizumab. In fact, mean LYs to 4 years on pembrolizumab are remarkably similar (2.8 years in the old model and 2.7 years in the new one). However, the trajectory for OS in the old model is much lower

Confidential

than that from SACT because it is heavily influenced by a surrogate relationship with PFS.

The OS for SoC is likely to be overestimated in the old model. Mean survival for patients who do not have an SCT on SoC is 3.1 years, which is also likely to be an overestimate. For example, the old model calculates that 60% of the non-SCT patients will be alive after a further 20 months follow up, which is markedly more than in the Eyre study, where OS was just 25% among patients who did not have an SCT at 20 months follow-up.

b) Please explain why, despite the less favourable OS data, the cost-effectiveness results are more favourable in this new submission, compared with the submission in TA540, where the company's base-case incremental cost-effectiveness ratios (ICERs) were £36,950 (24-week model) and £55,628 (12-week model) per quality-adjusted life year (QALY) gained (according to the FAD).

There are several reasons why the old model has a less favourable ICER than the new one. As discussed in part a) above, they have little to do with the difference in OS between KEYNOTE-087 and SACT. Some are data related and some are structural:-

1. The cost of SCT is reduced from £110,374 to £66,569, more accurately reflecting NHS unit costs, now we know that a third of SCTs are autologous. This would reduce the ICER because more patients have an SCT in the pembrolizumab arm.

2. The PAS discount on pembrolizumab has changed significantly.

3. The old model assigned a very low HRQoL value to patients with progressed disease (0.46), whereas the HRQoL for disease progression in both KEYNOTE-087 and KEYNOTE-204 trials remain above 0.8 in the pembrolizumab arm. It is also not supported by the fact that the utility estimates in the trials are drawn from many patients who are at later lines. Due to the short PFS in KEYNOTE-087, patients spend the majority of their time in the PD health state, which leads to very low average utilities across both arms. Adjusting these data to more appropriate values would reduce the ICER because OS is higher on pembrolizumab.

Confidential

4. There is no cure point implemented post SCT in the previous model. This means that patients continue to relapse for decades and actually spend the majority of their post-SCT alive time in the PD health state with its associated low utility of 0.46. Implementing more appropriate data would likely significantly reduce the ICER because many more patients have an SCT in the pembrolizumab arm.

5. The severity modifier of 1.2 is also applied in the new model. With this removed, the company's new ICERs range from [redacted].

Adjusting all these data simultaneously would bring the £36k ICER down significantly although some structural problems would still remain, particularly the likely underestimation of OS on pembrolizumab vs. the observed data overestimation of OS on SoC vs. clinical expectation.

c) Please implement all the collected data in the original cost-effectiveness model, present the results and submit the model file.

As outlined in the CS and in part in our response to b) above, the old model has several structural problems and no longer provides useful information for decision-making. Given the outcome data observed in SACT for the last 4 years and the significant clinical experience that has built up during this time, the various surrogate relationships that were used in it are no longer necessary. For these reasons, we have not done this.

Review

B 2. Table 46 in Appendix H1.2 outlines the inclusion and exclusion criteria for HRQoL systematic literature review. Ten studies were excluded during the full-text review because the outcomes were not of interest to the review. Please elaborate on the health states that were considered of interest for this review, based on which the ten studies were excluded.

All health states relevant to R/R cHL population were considered of interest to the review with no restrictions in the inclusion or exclusion criteria regarding health states, as provided in Appendix H1.2 (Table 46) and summarised below in Table 1 . The ten studies excluded during the full-text review stage due to reporting outcomes not of

Confidential

interest to the review are provided in the table below. For all of these studies, no utility values were reported regardless of health state and therefore the studies were not of interest to the review. Von Tresckow 2017[4] reports the overall change in utility in a phase 2b trial, but not the utility value itself by health state. In addition, the geography was not clearly stated in either Dada 2018[5] or Ionova 2021a[6], but appear not to be relevant to the UK context.

Table 1. Studies excluded from the 2023 utility SLR due to 'outcomes not of interest'

ID	Journal	Volume: Page	Title	Full text/abstract only
Dada 2018[5]	HemaSphere	2: 42	Nivolumab in relapsed/refractory classic Hodgkin lymphoma: Experience with ten patients	Abstract only
Engert 2017[7]	Blood	130	Effect of nivolumab on patient-reported outcomes in patients with relapsed/refractory classical hodgkin lymphoma after autologous transplantation: Results from the multicohort phase 2 checkmate 205 study	Full text
Husson 2018[8]	Journal of Clinical Oncology	36	Independent prognostic value of the EORTC QLQ-C30 summary score on all-cause mortality: Results from the population-based PROFILES registry	Full-text
Ionova 2019[9]	Blood	123: 5296	Response to Brentuximab Vedotin and Quality of Life in Patients with Relapsed/Refractory Hodgkin Lymphoma (RR	Full-text

Confidential

			HL) in the Real World Setting	
Ionova 2021a[6]	Hematological Oncology	39: 442-443	Brentuximab vedotin for treatment in patients with relapsed/refractory classical Hodgkin lymphoma in a real world setting: Clinical outcomes and impact on quality of life	Full-text
Ionova 2021b[10]	HemaSphere	5: 572	Outcomes of brentuximab vedotin as ≥3 line treatment in patients with relapsed/refractory classical hodgkin lymphoma: Physician's and patient's perspective	Full-text
Lepik 2019[11]	Hematological Oncology	37: 495	Response to nivolumab as ≥3rd line therapy in pts with relapsed/refractory classical Hodgkin's lymphoma (CHL) and its impact on quality of life in responders and nonresponders	Abstract only
Stadtbaeumer 2021[12]	Oncology Research and Treatment	44: 38	Predicting cancer-related fatigue of Hodgkin Lymphoma survivors: Identification of risk factors	Abstract only
Stadtbaeumer 2022[13]	HemaSphere	6: 36	Predicting the health-related quality of life of Hodgkin lymphoma survivors: identification of risk factors	Abstract only
Von Tresckow 2017[4]	Blood	130	Patient-reported outcomes in patients with classical hodgkin lymphoma treated with pembrolizumab	Full text

Confidential

			monotherapy, results of a phase 2 study	

B 3. Table 50 in Appendix I1.1 presents the inclusion and exclusion criteria for the cost and healthcare resource use review.

a) The inclusion criteria were expanded to include studies from US, Canada, Europe, Germany, or Denmark, in case of limited UK evidence. Eight relevant studies were identified, seven of which were multicenter studies in the US, and one was a single-center study in France. Two of these studies specifically focused on R/R cHL. However, the company did not use the costs and resources data from these studies in the economic model. Please justify why the costs and resource use data have been excluded, although the inclusion criteria were extended to other countries due to limited UK evidence

b) Please provide further justification on why cost evaluation studies have been excluded from costs and healthcare resource use review

Eight of the included cost and resource use studies identified were multicentre studies in the US. Cost data from the US cannot be generalisable to the UK setting and therefore no cost data extracted was utilised in the economic model. With regards to resource use from the US studies, the data reported were either too granular to be of use in the economic model or were not specific to line of treatment and therefore would give an inaccurate perception of resource use. The company therefore felt that clinical expert opinion would better reflect resource use in the UK context with regards to the population of interest. One single-centre study in France was identified meeting the inclusion and exclusion criteria, however, this study reported no resource use and only cost data specific to Nivolumab treatment in the French hospital setting. This cost is not generalisable to the UK context due to a difference in the healthcare systems and reimbursement structures.

The inclusion and exclusion criteria for the 2023 systematic literature review update was in line with that from 2017, whereby cost evaluation and costs and resource use were looked at in isolation. If cost evaluation studies were to be included, no cost evaluation studies were identified in the 2017 search, with only one study identified in the 2023 literature search (Jones, 2017[14]). Jones 2017 is available only as an

Confidential

abstract and not full-text, and therefore does not provide any detail on cost or resource use data with no specification of the sources used. In conclusion, adding cost evaluation to the cost and resource use review would provide no additional data of use in the economic model.

Model structure

B 4. Priority question: In the CS it is states "Allowing the OS curves to continue to four years has the major advantage that the SCT related events, their treatment effects and short term outcomes do not have to be estimated explicitly but rather are implicit parts of the OS curves". This statement indicates that the current model structure is potentially inconsistent with good modelling practices. Firstly, it is suboptimal form a transparency perspective as outcomes are not estimated explicitly. Secondly, it violates the homogeneity within health states assumption: "states need to be homogeneous with respect to both observed and unobserved (i.e., not known by the decision maker) characteristics that affect transition probabilities" (http://dx.doi.org/10.1016/j.jval.2012.06.014). Please comment on the chosen model structure in this light.

We understand the EAG's concern but do not think that this limitation could plausibly have a decision-important effect on the ICER. Some degree of within-state heterogeneity is almost always present in cohort level health economic models. For example, a standard 3-state advanced cancer partitioned survival model (PSM) of the sort that has been utilised in many previous NICE appraisals of pembrolizumab (e.g., TA428,[15] TA531,[16] TA683,[17] TA772,[18] and many more) consists of two assumed-homogenous alive health states, Progression Free Survival (PFS) and Progressed Disease (PD) and a (truly homogenous) death state. This model structure typically attracts little objection in NICE Technology Appraisals. However, within the PFS health state there are patients who have Complete Response, Partial Response and Stable Disease. Within the Progressed Disease health state there are patients who never responded and those who are progressing from Complete Response, Partial Response and Stable Disease. In both health states there are patients of a wide variety of ages with different prognoses and comorbidities, yet a common age and utility are

Confidential

assumed. ToT is often costed externally to the model structure. Various subsequent treatments are utilised yet are typically modelled only as cost payoffs rather than either affecting transition probabilities or being induced by transitions. Utility might truly be expected to change over time to reflect the changing composition of the health states among these response-related sub-states, yet it is rarely modelled this way in practice due to lack of evidence. The reason why such heterogeneity is deemed acceptable that to implement more granular structures is cumbersome, data-intensive and unlikely to meaningfully drive the decision, which is the ultimate purpose of the model. Whether we had explicitly accounted for transitions to SCT or not within the first 4 years would not affect OS, since this would have to be estimated in the same way (using the SACT data for pembrolizumab and an OS HR to estimate OS for SoC). The effect on costs and QALYs of introducing these transitions would be minimal and would favour pembrolizumab, as we illustrate elsewhere in our responses. Relatedly, the proportion getting SCT, the associated costs and QALY disbenefits are explicitly accounted for as payoffs outside the OS curves, so are transparent and modifiable. In evidence light areas such as this one, our view was that it was better to adhere to the principles of model parsimony by building a simple model that captured the principal costs and benefits, *then* investigating whether any structural limitations could affect the decision via the use of extensive sensitivity analyses. This was preferred to attempting to build a more granular model that required additional parameters that would have relied on a series of highly uncertain data and assumptions.

B 5. Priority question: In the original TA540 CS it is stated that, "it is expected that pembrolizumab monotherapy will be used as a "bridge" to alloSCT". Similarly, in the current CS it is stated that "pembrolizumab is a bridge to SCT" and "Clinicians from the MSD UK Advisory Board stated among patients who received prior BV and are autoSCT ineligible, pembrolizumab would be the preferred treatment option to help achieve a better or durable response to bridge them to SCT". In addition, the recommendations for data collection (TA540 FAD) were mainly focused on (time to) SCT. Given the above, the mechanism through which pembrolizumab affects patient outcomes is through increasing the probability of (curative) SCT. However, the company additionally assumed that pembrolizumab would also improve both OS and health-

Confidential

related quality of life, both pre landmark (all patients) and post-landmark (in case of no/failed SCT).

a) Please comment on the statement that the main mechanism through which pembrolizumab affects patient outcomes is through increasing the probability of (curative) SCT, providing supporting evidence.

Increasing the probability of (curative) SCT is not the only mechanism through which pembrolizumab improves patient outcomes. While pembrolizumab is likely to increase both the probability of SCT and the probability that SCT will be curative, it is also likely to increase OS in patients who do not receive SCT and to increase HRQoL independent of SCT. >60% of patients in both KEYNOTE-204 and KEYNOTE-087 achieved CR or PR. The SACT dataset is divided roughly 60/40 into patients who are potentially candidates for SCT and those who are not e.g. through age/comorbidity. Pembrolizumab can control cHL among these patients as well as among those who are thought to be candidates for SCT. The ToT data in SACT show that although the vast majority of SCTs had occurred by 9 months, ~25% of patients remained on treatment with pembrolizumab at 1 year and ~11% at 2 years. Given the Blueteq criteria that pembrolizumab should be stopped if there was not ongoing clinical benefit, this is evidence that there is a significant cohort of patients who do not get an SCT but continue to benefit from pembrolizumab. Clinicians at the MSD UK advisory board were clear that patients who did not have an SCT can have good control and HRQoL even years after stopping pembrolizumab, outcomes which far exceeded their expectations on SoC. The advisors noted that while durability of response depended on the individual patient, that most patients with durable responses maintained these after stopping pembrolizumab. One advisor indicated that quality of life improvements with pembrolizumab treatment meant that some patients even opted not to undergo planned alloSCT. The evidence on all treatment effects is detailed in the CS but MSD would like to draw the EAG's attention to the OS HR among patients who never had an SCT from KEYNOTE-204 ([redacted]). This is relevant (given this is pembrolizumab vs. BV, a better treatment than SoC, given in a population that is likely to be more chemo-sensitive in which a roughly equivalent proportion of patients achieved SCT between the arms, this is theoretically biased against pembrolizumab) and the data from Eyre among patients who did not have an SCT (OS=25% at 20 months) and Cheah

Confidential

(OS=~30% at 2 years) vs. the SACT data (OS=54% at 2 years). Patients in the Eyre data have a median age 20 years younger than in SACT, one treatment line earlier and 100% were thought to be candidates for SCT vs. 60% in SACT so this comparison is also theoretically biased against pembrolizumab. Patients in Cheah are also 20 years younger and have a history of greater chemo-sensitivity so this comparison is also theoretically biased against pembrolizumab.

b) Please justify assuming increased OS pre landmark, in addition to the survival benefit yielded through the increased probability of (curative) SCT.

Please see our response to part a).

c) Please justify assuming increased OS post landmark, in addition to the survival benefit yielded through the increased probability of (curative) SCT and also considering this is ≥2 years after stopping pembrolizumab treatment (given the four-year landmark and two-year stopping rule).

Please see our response to part a). The continuation of an OS treatment effect is based on no data having ever been identified, including among longer term (5 year+) studies of the treatment effect of immunotherapies waning. We do understand the reason NICE Committees have typically imposed a waning effect in appraisals of pembrolizumab and the standard 3–5 years post treatment cessation (e.g. NICE TA885) has been included here as an optional scenario analysis to be consistent with precedent. Advisers at the UK advisory board were also clear that they see durable responses after stopping treatment on pembrolizumab.

d) Please justify assuming increased health-related quality of life pre landmark, in addition to the health-related quality of life benefit yielded through the increased probability of (curative) SCT.

The inclusion of this treatment effect is based on randomised evidence from KEYNOTE-204. A study in which the proportion of patients receiving an SCT was almost the same between the arms (40.9% vs. 39.1%). The underlying reason for this difference is explained in part a), namely that pembrolizumab elicits greater and more durable levels of response than SoC (or BV) even among patients who do not get an

Confidential

SCT. Within the model for this appraisal, the sub 4-year treatment effect is in fact biased against pembrolizumab. This is because it is taken from a study where there was no difference in SCT probabilities between the arms, whereas the difference is expected to be significant in this setting.

e) **Please justify assuming increased health-related quality of life post landmark, in addition to the health-related quality of life benefit yielded through the increased probability of (curative) SCT and also considering this is ≥2 years after stopping pembrolizumab treatment (given the four-year landmark and two-year stopping rule).**

There is no difference in HRQoL post landmark among patients who had a curative SCT. The difference assumed among the no/failed-SCT group is taken from the evidence in KEYNOTE-204 (see our response to part d) and from clinical expert opinion. The duration of benefit is uncertain and we examined removing this treatment effect entirely in sensitivity analysis. Another option would be to implement a similar 3-5 year post-cessation waning effect, which is explored in B20.

f) **Please provide the results of scenario analyses (and an updated version of the model), assuming no pembrolizumab specific OS and health-related quality of life benefit. Assuming differences between treatment are only driven by the increased probability of (curative) SCT.**

We do not consider this scenario helpful for decision-making. It is clear that pembrolizumab is expected to have an effect on OS over and above its ability to help patients reach a (curative) SCT. Please see our responses to earlier bullets in this section.

g) **Please provide the results of scenario analysis (and an updated version of the model), assuming no pembrolizumab specific OS and health-related quality of life benefit after the landmark point, assuming differences between treatment are only driven by the increased probability of (curative) SCT and increased OS and health-related quality of life pre landmark.**

This scenario can be achieved by setting the treatment waning parameters in the Model Settings sheet to “yes”, 4 and 0 (i.e. instant waning/equalisation to the SoC

Confidential

hazards at 4 years) and by setting the HRQoL of pembrolizumab after the landmark equal to that of SoC after the landmark in the 'References' tab. The ICER increases from [redacted] to [redacted] gained indicating a moderate influence on the ICER but not the decision.

B 6. Priority question: The current model structure includes a four-year landmark point, where patients are transited to either successful SCT or no/failed SCT. This assumption might be suboptimal and potentially has limited face validity.

a) Please comment on the face validity of this assumption compared to including the probability of transitioning to successful SCT or failed SCT every cycle.

Both structures have reasonable face validity but only one is feasible and parsimonious to implement with the data available. At 4 years there will be a proportion of those alive who have had a successful SCT and are, in effect, cured and a proportion who constitute the remainder. The same would be true in a model structure which allows transitions every cycle. The proportions of patients with SCT would be exactly the same in both model arms because they'd have been drawn from the same underlying datasets. The model after 4 years would therefore be exactly the same. The alternate structure proposed here provides some more nuance within the first 4 years but to implement this structure requires the estimation of an additional uncertain treatment effect and some additional calibration. The value of doing this is minimal, given that it would result in the same 4-year OS curves and an additional HRQoL benefit for pembrolizumab (because the pembrolizumab patients would spend a greater proportion of their pre-landmark time in the cured-SCT state than the SoC patients), which already has a very low ICER in this setting.

b) Please comment on why a landmark was necessary at all and justify its timepoint at 4 years, given that most patients will have had SCT in the first two years.

Fuller justification is provided in the CS but briefly, this was the timepoint at which clinicians confirmed that all SCT related events (procedures and relapses) would have resolved. This meant 4 years was the timepoint at which the population could be

Confidential

divided up into those who were cured and those who still had R/R cHL. The landmark model was chosen because it enabled a parsimonious analysis that could capture all the data that would meaningfully alter the ICER, while minimising the number of assumptions we needed to make and without imposing the need to estimate additional treatment effects (e.g. HR on "time to SCT or death"). In an appraisal that already contains no directly applicable outcome data on the standard of care, attempting to nuance the model further by modelling relapses after SCT explicitly would have been computationally complex and would have relied on even more uncertain, low quality evidence. Varying the landmark does not meaningfully alter the ICER. The landmark can be set to 2 years and it drops the ICER slightly to [redacted] gained. We note that the previous model also employed a landmark structure to try to overcome the computational complexity of transitioning patients to SCT then modelling their time-dependent outcomes, albeit in a slightly different way.

c) **Please elaborate on the implications in terms of differences in QALYs and costs of this assumption compared to including the probability of transitioning to successful SCT or failed SCT every cycle.**

The suggested structure would likely result in a limited incremental QALY gain for pembrolizumab due to patients spending more time in the cured-SCT health state pre-landmark than the SoC patients. All costs would remain the same with the exception of Health State Resource Use prior to the landmark, which would reduce for pembrolizumab more than SoC for the same reason as above. The model after 4 years would be exactly the same. Overall, life years would be the same, incremental QALYs would slightly increase and incremental costs slightly decrease so the ICER would slightly decrease.

d) **In addition to section B.3.2.4.4, please further justify that combining no and failed SCT is reasonable and elaborate on the implications in terms of differences in QALYs and costs of this assumption.**

The precise effect of combining the two groups (or, more specifically, using the data on the No SCT patients as a surrogate for the combined group of No/Failed SCT patients) is uncertain and should be the subject of sensitivity analysis, including hypothesising whether adding a small number of Failed SCT patients to the much

Confidential

larger group of No SCT patients meaningfully alters HRQoL, HSRU, OS transition probabilities and in which direction. This is a simplifying assumption that is necessary because no relevant data on which to base a more complex model was identified.

Within the SoC arm, ~92% of patients do not get an SCT and ~5% will fail one. In effect, the No/Failed group in the SoC arm are therefore quite homogenous. They all have R/R cHL and will be treated with generic chemotherapy to which they all have proven insensitivity. Since the proportion is small and the failure of SCT does not meaningfully alter the treatment options available, it may be reasonable to conclude that the inclusion of failed SCT patients in this population does not meaningfully alter outcomes. The No SCT OS estimates are therefore likely to be a good surrogate for No/Failed OS.

In the pembrolizumab arm, these conclusions are less certain and consequently the appropriateness of combining the groups is less certain. ~70% will not have an SCT and ~14% will fail one meaning that, although they are a significant minority, the Failed group could have a larger influence on outcomes. The group treated with pembrolizumab are also less homogenous in that some patients will have responded well to pembrolizumab, not had an SCT because of comorbidities etc. but have some lingering treatment effect. As a matter of theory, the group who failed SCT could be split out and modelled separately but this would add needless complexity to the model, given that outcomes among the no/Failed SCT group are not important drivers of the ICER. Any difference in outcomes that would be achieved via extending the model is already covered within the sensitivity analyses we've undertaken. See for example our response to question B5 (g) in which we illustrate that the total removal of all treatment effects post the 4-year landmark, which is highly conservative, only increases the ICER to [redacted] gained. The ICER could only be driven up further if there is reason to believe that OS and HRQoL would somehow be better in the "No/Failed" group who are alive at four years in the SoC arm than the pembrolizumab arm. Given the lack of effectiveness of generic chemotherapy options, we view this as implausible.

e) Time to SCT was recommended in the TA540 FAD for data collection in the CDF. Please justify that time to SCT was not explicitly incorporated in the model structure.

Confidential

As outlined in the CS, the most important outcomes have now been observed directly through pembrolizumab's time in the CDF and do not need to depend on a median SCT time point via a linked-evidence approach as they did in TA540. The reason this outcome was prioritised for data collection is because the 12-week and 24-week economic models that were considered during TA540 produced markedly different ICERs and the Committee were interested in which one was more appropriate. However, the reason for this structural sensitivity was because a proportion of PFS patients were transitioned to SCT at either a 12-week or a 24-week landmark. The denominator was, of course, markedly lower at 24 weeks than 12 weeks. Now that pembrolizumab has spent several years in the CDF and we have data on 215 patients treated in UK clinical practice, we know what proportion get an SCT, what proportion are autoSCT or alloSCT, what OS is and that PFS is not strongly predictive of any of those outcomes. A more accurate and less sensitive model structure is now possible and time to SCT as a model-driving parameter is not necessary.

f) An alternative model structure would consist of three health states:

i. no/failed SCT (without distinction before/after an arbitrary landmark)

ii. successful SCT

iii. death

This model structure is simple, transparent, adheres to best modelling practices and does not require tunnel states. Moreover, compared with the company's model structure, it would not necessitate using an arbitrary landmark point and allow including the probability of transitioning to successful SCT every cycle. Please justify why this approach was not adopted in the CS.

Theoretically, we have no objections to the suggested structure and we were initially interested in implementing something like this, which is why we requested the "time to SCT or death" data from SACT.

We also discuss this earlier in our response but principally we have two reasons not to implement it. The first reason is that it would make very little difference to the

Confidential

decision and the second is that it would require the estimation of an additional uncertain treatment effect or some form of model calibration to estimate outcomes on the SoC arm. It would make little difference to the decision because OS would be the same up to 4 years and this model would be *exactly* the same from year 4 onwards, where the large majority of the incremental benefit is observed. There would be a small incremental net health benefit for pembrolizumab prior to 4 years because there would be more cured-SCT patients in the pembrolizumab arm and these patients have lower health state resource use and higher HRQoL.

The second reason is that it requires the estimation of a treatment effect on "time to SCT or death" (time to SCT alone would not work because the implied health state is "alive and hasn't had or has failed an SCT"), which is an outcome that does not exist in any trials and is a difficult composite end-point for clinicians to help estimate.

g) Please provide the results of scenario analyses (and updated version of the model), using the model structure as described in the previous sub-question.

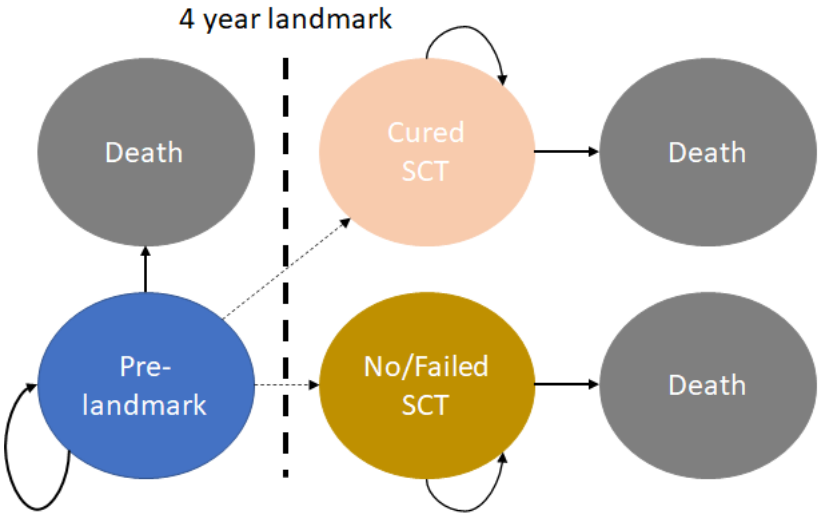
Due to the reasons outlined in our response to question 5f above, MSD does not believe that implementing this model structure will impact on the decision, has several key disadvantages compared with the landmark model and is not feasible to implement in the timeframe of clarification questions responses.

B 7. Schematic representation of the model structure. Please provide a schematic representation of the model structure, illustrating the model health states as well as transitions (this is not clear from CS Figure 12), similarly as provided in CS Figure 11 for the original TA540 model.

In the pre-landmark time period, all patients exist in the blue health state in the diagram below. Their transitions are governed by the observed survival data in SACT (or these data multiplied by an HR in the SoC arm). At the landmark the probability of curative SCT is calculated by %SCT multiplies by %curative. These percentages are both specific by intervention. These patients transition to the "Cured SCT" health state and the remainder (the alive patients minus the cured proportion) transition to the "No/Failed SCT" health state. There is no difference in transition probabilities in the "Cured SCT" health state between the arms. The transitions to death are governed by general population mortality, which can be multiplied by an arbitrarily set SMR. The

No/Failed group's transitions to death are governed by parametric extrapolation from the "No SCT" group in SACT. This is multiplied by the OS HR among "No SCT" patients from KN204 to obtain the transitions for the SoC arm. When treatment waning is imposed, the transitions for the No/Failed SCT group in the pembrolizumab arm become equal to those in the SoC rather than the other way around (i.e. the pembrolizumab patients' hazards increase rather than the SoC hazards decreasing).

Figure 5. Schematic of the model structure



Population, intervention, comparators

B 8. Priority question: NICE final scope included single or combination chemotherapy such as gemcitabine, vinblastine and cisplatin, as a comparator. Furthermore, the company mentions that patients with R/R cHL are likely to receive single-agent chemotherapy after BV. However, MSD have deviated from the final scope by creating a blended comparator based on Cheah 2016, Eyre 2017 and expert opinion, where the proportion of all the treatments in the SoC combination were assumed to be equal (Table 50 in the CS).

a) Please provide further justification for using the blended comparator using the different sources to inform the model.

The blended comparator reflects the interventions in use in UK clinical practice, which are highly variable. We were unable to obtain firm proportions from the advisory board.

Confidential

Adjusting its composition does not affect health outcomes in the model because these are not drawn from studies that reflect the efficacy of single comparators but instead affects only SoC treatment costs. We recognise this is an area of uncertainty so undertook sensitivity analyses that indicated a fairly small effect on the ICER. For example, assuming 100% receive bendamustine (a relatively cheap monotherapy) results in an ICER of [redacted].

b) Please provide scenario analysis that incorporates single or combination chemotherapy including drugs such as gemcitabine, vinblastine and cisplatin in the comparator arm in line with the NICE final scope.

We added a DSA where gemcitabine, vinblastine and cisplatin monotherapy were added to the SoC composition (again, equal weights between every option were assumed) but would note that cisplatin monotherapy is not in use in the NHS. In general, our view is that while the true breakdown of SoC treatments is uncertain, the results of plausible breakdowns are already covered within the envelope of sensitivity analyses already undertaken.

c) Please further justify the assumption that the composition of all treatments in the SoC combination is equal and explain the direction of potential bias on cost-effectiveness results.

The justification for why the SoC regimens in the composition are equal is made in the CS (B.3.2.5) but will be explained briefly. We listed the treatments from Cheah et al. and Eyre et al. (2017) and presented them to clinical experts at our advisory board. The 8 clinicians we consulted with were able to rule out some options but could not confidently provide estimates for the proportion of patients that would receive each of the remaining regimens as there is no agreed SoC. This uncertainty was due to several factors: the availability of pembrolizumab in 4L R/R cHL, patient heterogeneity (clinical history (e.g. response to other lines, patient preference, fitness levels, age), heterogeneity of chemotherapy options in use in difference centres and finally, the rarity of the population in question.

For simplicity, we assumed an equal proportion for each treatment. In addition, we tested alternative SoC compositions including one scenario where bendamustine was

Confidential

set to 100% and a second scenario where the SoC treatment costs were halved as outlined in the CS. Although the direction of the bias is unknown (given the inherent uncertainty of these estimates), the results from the scenario analyses conducted indicate that even extreme assumptions would produce a minimal impact given the ICERs were [redacted] and [redacted] respectively [redacted] and [redacted] difference from the basecase ICER).

d) Please justify why the approach for the SoC composition from TA540, and TA462 have not been used to inform the economic model.

MSD note the TA540 comparator composition methodology considered the treatments from Cheah et al., TA462, the clinical advisory board from TA540 and a series of assumptions. Consistent with the TA540 methodology, we obtained a list of candidate treatments from a number of sources and validated them with clinical experts. Some treatments were removed from the list and others added (for a full discussion of which treatments were removed or included please see CS B.3.2.5). The differences would suggest the composition in TA540 and TA462 does not reflect current clinical practice for R/R cHL, although MSD note that the exact composition is a matter of considerable uncertainty.

e) Please explain why historical data from SACT were not used to inform the composition of SoC.

We do not have access to any historical data from SACT.

f) Please provide scenario analysis with historical data from SACT and TA540 informing the comparators

We do not have access to historical data from SACT. Due to time constraints, we have not implemented this scenario but would not that it would certainly have a small effect on the ICER, given the results of other scenarios we have run around SoC composition.

B 9. The baseline characteristics for the population model were derived from the baseline characteristics of both the SACT database and KN-087 Cohort 2, Please justify why these two sources were used and discuss how this could potentially bias the outcomes of the analysis.

Confidential

The SACT database was the preferred source of inputs for the economic model, where data are available, as it better reflects the R/R cHL population of interest in the UK real-world clinical setting. Therefore 'Baseline age' and 'Proportion female' were derived from the SACT data. Data on 'Weight' and 'Body surface area (BSA)' are not reported in the provided SACT data and consequently could not be derived for the model, therefore these baseline characteristics were taken from the KEYNOTE-087 Cohort 2 ASaT population. The table below summarises the available baseline characteristics from both SACT and KEYNOTE-087, showing that the SACT population were older than Cohort 2 from KEYNOTE-087 (51 years vs 42.3). Changing the baseline characteristics in the model to all be derived from KEYNOTE-087 has little impact on the ICER, and would produce an ICER more favourable to the intervention.

Table 2. Comparison of available baseline characteristics between SACT and KEYNOTE-087

Characteristic	**Mean**	
	SACT	**KEYNOTE-087 Cohort 2 ASaT**
Baseline age (years)	51	42.3
Proportion female	0.40	0.47
Weight (kg)	NR	73.73
Body Surface Area (BSA)	NR	1.85
Abbreviations: ASaT, all-participants-as-treated SACT, systemic anti-cancer treatment		

Treatment effectiveness

B 10. Priority question: The company informed some required data inputs using structured expert elicitation. The data collection recommendations made in the FAD (TA540) include data that was now elicited, such as proportion of people on SoC having an SCT.

a) Please explain why historical data from SACT were not used.

b) Please provide these historical data from SACT to inform the comparator OS, time to SCT and the composition of SoC.

Confidential

MSD do not have access to any historical data from SACT, nor are we aware that any have been published.

B 11. Priority question: To be in line with reported guidance from NICE DSU TSD 14 and 21 on survival analysis, please provide the following, separately for pre-landmark OS, post-landmark OS, time to SCT for patients, OS for patients with and without successful SCT, and time to treatment discontinuation (TTD) and separately for the intervention and comparator (which can be informed using historical SACT data as recommended above):

a) Tables with the numbers of patients at risk, per 3 months.

All KM plots supplied already come with N at risk tables where data are available (KEYNOTE trials and SACT [please see the SACT report where these tables were reported separately]). In the case of Cheah, these data are not supplied at all in the study and in Eyre they are only supplied every 20 months.

b) To examine the proportional hazard assumption:

i. Plot the scaled Schoenfeld residuals versus time (all survival curves)

These are now provided in the ITC documentation (please see separate document supplied as part of MSD's response to clarification. We note that while there is no strong evidence to indicate non-proportional hazards in the ITCs, a trend towards equal hazards is to be expected in these comparisons. The reason for this is that a large percentage of patients (~30%) in all datasets receive SCT interventions which are potentially curative. We note that because of the nature of the studies, the proportion in Cheah and Eyre that receive SCT is far higher than is expected in the control arm of our economic model. Furthermore, there are few patients at risk in these studies after 2 years. Thus, information from Schoenfeld residuals and smoothed HR should not be used to conclude that hazards between pembrolizumab and SoC in the economic model should narrow over time. Within both KEYNOTE-204 cohorts (ITT population and those who did not have an SCT, there are very few events after the first 18 months so it is difficult to draw any conclusions about how long a treatment effect might persist.

Confidential

ii. Plot the log cumulative hazard versus log time

The log cumulative hazard versus log time plots are provided in a separate file titled 'ID5084_response to B11 b [noCON}'.

c) To examine the heuristics of the hazard function over time, please plot the smoothed hazards over time.

Please see response to part d) below.

d) To examine diagnostics of parametric survival models (using the observed data):

i. Plot the cumulative hazard versus time

ii. Plot the log smoothed hazard versus time

iii. Plot the standard normal quartiles versus log time

iv. Plot the log survival odds versus log time

There are two parametric models used in the economic model. The first is used simply for convenience rather than extrapolation and relates to OS among pembrolizumab patients in SACT for the first four years. It can be seen from Figure 13 in the CS that the exponential curve provides a reasonable fit to this data based on AIC. Visual examination reveals slightly higher hazards in year one and slightly lower hazards in year four. Broadly the hazards are reasonably constant but the dataset is, of course, immature with median OS not having been reached. Given that 30% of patients have received SCT and that a minority of patients are expected to respond very well to immunotherapy, as is normal in pembrolizumab trials. It would be reasonable to select a non-exponential curve with a decreasing hazards property were extrapolation being undertaken for the economic model. The second parametric model is used to extrapolate outcomes for the No/Failed SCT group using SACT data on those who didn't have an SCT. The exponential curve fits the data well indicating that hazards are broadly constant within the observed period (Figure 15 in the CS). We selected a log-logistic model in our base case to fit with the clinical expectation of declining hazards among patients treated with pembrolizumab but considered the exponential curve a reasonable, more conservative alternative.

Confidential

e) **To examine the validity of the extrapolation beyond the data, please provide supporting evidence that the extrapolations are consistent with relevant external data and/or expert opinion. In case of expert opinion, please provide a full description of the methods and results of the expert consultation conducted.**

The data that were extrapolated in the model was only overall survival for the group of patients who were unable to have an SCT or who had progressed after SCT. This was initially assessed via comparison to external evidence. Following clarification, we have consulted one clinical expert to assess the validity of the choices of survival curve in the standard of care arm (exponential versus log-logistic). We asked the clinical expert whether they would expect any patients treated on standard of care to remain alive after 8-10 years. The clinician estimated that close to zero of these patients would be alive which better reflected the exponential curve. Based on this response we also presented the expert with landmark estimates at 5, 8 and 10 years from the pembrolizumab arm of the model using the exponential extrapolation where treatment waning had either been applied or omitted. The clinical expert indicated that [the survival estimates where waning was applied were more implausible because all patients had died at 10 years. He indicated that there would be a small but significant proportion of patients who responded well to pembrolizumab and that the “no-waning” scenario looked plausible. Based on this feedback, we consider that it may be more appropriate to update our base case analysis to include the more conservative exponential curve instead of the log-logistic for the “No/Failed SCT patients”. This is already handled within the sensitivity analyses provided and has a moderate effect on the ICER. Another potential amendment could include an exponential curve for the SoC and a log-logistic curve for pembrolizumab but we didn’t have time to implement this.

f) **Please justify the selection of the approaches to estimate and extrapolate OS, PFS, and TTD, taking into account the responses to the preceding questions as well as the *"Survival Model Selection Process Algorithm"* provided in NICE DSU TSD 14**

Survival analysis curve selection is detailed in section B.3.3.1.6 of the company’s submission. We do not have strong views about the most appropriate parametric

Confidential

model based on visual fit, AIC and clinical expectation. We consider that all models are helpful for decision making to show the effect of uncertainty.

g) As suggested in NICE DSU TSD 14, please provide *"substantial justification"* in case different types of parametric models are used for different treatment arms.

There are no parametric survival extrapolations that use different models between the arms.

h) Please provide time to SCT data and perform survival analysis adhering to TSD 14 and 21 guidance also for scenario asked in B6.g).

There is no need to provide parametric survival extrapolations for time to SCT data as no additional SCTs are expected beyond the observed follow-up period.

B 12. Priority question: The company highlight serious limitations surrounding all data sources considered for the indirect comparison of treatment effectiveness. For example, the chosen Bucher ITC relies on data from a population in a different treatment line (where the relative effectiveness may differ and it is unknown in which direction) and another comparison of brentuximab vedotin versus SoC, for which it is unclear whether SoC is the same as is considered in this appraisal. It furthermore relies on digitized plots. Please provide a tabular overview detailing the advantages and disadvantages associated with all data sources considered for the indirect comparison, considering the appropriateness of the population (age, treatment line, …) and the definition of the treatment.

A tabular overview of the key limitations of the OS estimates was presented in CS Section B.3.3.1.4, table 33 (page 103). An additional summary is provided below (Table 3).

MSD agrees that the process of digitising published Kaplan–Meier curves for the purposes of conducting ITCs yields pseudo-IPD that may not exactly match the original IPD of the comparator study, but notes that there is no clear reason to believe that any mismatching would bias results in favour of any particular

Confidential

comparator and that the magnitude of any bias is likely to be small. This method is standard in NICE technology appraisals.

Confidential

Table 3. Summary of the key strengths and limitations of the indirect treatment comparisons

Estimate	HR (CI)	Strengths	Limitations
1) Bucher ITC (KN204 and TA524)	[redacted]	• TA524 evidence previously accepted by NICE as representative of standard of care at 3rd line • Calculated implied HR of 0.62 for BV vs SoC was validated by clinicians at an advisory board • Comparison of pembro vs BV based on data from randomised controlled trial (see next row) • If generalisability bias exists it will favour SoC because pembrolizumab response is not expected to alter whereas response to BV or chemo would likely be worse in our population • Based on the only randomised trial in R/R cHL	• KEYNOTE-204 and TA524 data are both from 3rd line patients (ie one line of therapy earlier than the population of interest) • Patients in KEYNOTE-204 younger and fitter than patients in SACT. Baseline characteristics for TA524 not reported. • No appropriate method to calculate standard error for BV vs SOC from TA524 à assigned arbitrarily large value • Unclear direction of bias in TA524 • SoC from TA524 unlikely to closely match our SoC • High and equal number of SCTs in both arms in KEYNOTE-204 may influence outcomes
2) KN204 OS	[redacted]	• Based on the only randomised trial in R/R cHL • Based on stratified subgroup of patients who had no prior SCT • Analysis uses IPD from trial • If generalisability bias exists it will favour SoC because pembrolizumab response is not expected to alter whereas response to BV or chemo would likely be worse in our population	• BV used as proxy for SoC • KEYNOTE-204 data are from 3rd line patients (ie one line of therapy earlier than the population of interest) • Patients in KEYNOTE-204 younger and fitter than patients in SACT • High and equal number of SCTs in both arms in KEYNOTE-204 may influence outcomes

Confidential

3) Bucher ITC (SACT vs Eyre and TA524)	0.41 (0.22 - 0.77)	• TA524 evidence previously accepted by NICE as representative of standard of care at 3rd line • SACT data reflects real-world outcomes on pembrolizumab • Eyre data based on patients seen in UK clinical practice	• Eyre and TA524 data are both from 3rd line patients (ie one line of therapy earlier than the population of interest) • 100% of patients in Eyre are fit for transplant • Patients in KEYNOTE-204 younger and fitter than patients in SACT. • Baseline characteristics for TA524 redacted/not reported. • No appropriate method to calculate standard error for BV vs SOC from TA524 à assigned arbitrarily large value • Unclear direction of bias in TA524 • SoC from TA524 unlikely to closely match our SoC
4) ITC SACT vs Eyre	0.66 (0.44 - 0.98)	• SACT data reflects real-world outcomes on pembrolizumab • Eyre data based on patients seen in UK clinical practice • Committee welcomed analysis based on Eyre in TA540 (while acknowledging limitations)	• BV used as proxy for SoC • Eyre data from 3rd line patients (ie one line of therapy earlier than the population of interest) • 100% of patients in Eyre are potentially fit for transplant (pending response) vs. 61% in SACT • Patients in Eyre younger and fitter than patients in SACT. • Unadjusted ITC
5) ITC SACT vs Cheah	0.59 (0.4 - 0.86)	• SACT data reflects real-world outcomes on pembrolizumab	• BV used as a proxy for SoC • Cheah is based on patients treated at a single institution in the US • 71% of patients in Cheah had a prior SCT and 32% received unspecified investigational agents following treatment with BV

Confidential

			• Patients in Cheah younger and fitter than patients in SACT • Complete records not available in Cheah for all patients who were treated with BV • Unadjusted ITC
6) MAIC KN087 vs Eyre	0.23 (0.12 - 0.42)	• Data from KEYNOTE-087 was reweighted to match age, sex and ECOG status (0 vs 1) reported in Eyre • Committee welcomed analysis based on Eyre in TA540 (while acknowledging limitations) • KEYNOTE-087 has long follow-up	• BV used as a proxy for SoC • Eyre data from 3rd line patients (ie one line of therapy earlier than the population of interest) • 100% of patients in Eyre are potentially fit for transplant (pending response) • Patients in Eyre younger and fitter than patients in SACT.
7) MAIC KN087 vs Cheah	0.24 (0.14 - 0.4)	• Data from KEYNOTE-087 was reweighted to match age, sex and ECOG status (0 vs 1) reported in Cheah • Comparison versus Cheah used in base case analysis in TA540 (while acknowledging limitations) • KEYNOTE-087 has long follow-up	• BV used as a proxy for SoC • Cheah is based on patients treated at a single institution in the US • 71% of patients in Cheah had a prior SCT and 32% received unspecified investigational agents following treatment with BV • Patients in Cheah younger and fitter than patients in SACT • Complete records not available in Cheah for all patients who were treated with BV

B 13. Priority question: Elaborate justification and supporting evidence is missing for some of the modelling choices. Please provide justification and evidence to support:

a) Survival after successful SCT is assumed to be equal to that of the general population.

This is based on feedback at the clinical advisory board that there are no directly applicable long-term data but that long term mortality is expected to be equal or close to the general population. We acknowledged this uncertainty in the CS by including an arbitrary SMR parameter which can be applied to the cycle-by-cycle hazard of death to help understand the decision uncertainty associated with this assumption.

b) The choice of the log-logistic based on decreasing hazards, given that the decreasing hazards observed on patients treated with pembrolizumab are likely also influenced by it being a bridge treatment to SCT, and given that treatment with pembrolizumab occurred at least 2 years prior to the application of this curve.

The log-logistic curve is fitted only to patients that have not had an SCT so the hazard function is not influenced by SCTs. Declining hazards are consistently observed across pembrolizumab trials because a proportion of patients respond very well to immunotherapy and these patients constitute an increasing proportion of the alive cohort as time progresses. We also included exponential function in scenario analysis. Upon consultation with a clinician during writing of these responses, it may be that the exponential function is more appropriate for the SoC arm (see response to B30). An exponential function could also be imposed on the pembrolizumab arm although if treatment waning is also imposed, this would lead to 8- and 10-year survival estimates described as implausibly low by the same clinician.

c) Treatment effect applied in the no/failed SCT state, given that this occurs at least 2 years after pembrolizumab treatment discontinuation.

Despite a 35-cycle stopping rule, no data have ever been identified that suggest the treatment effect of pembrolizumab wanes over time, including in longer term (5 year+) trials. We acknowledge that it is normal for NICE Committees to impose a treatment

waning assumption of 3-5 years post cessation of pembrolizumab (e.g. the recent NICE TA885) and implemented this as a scenario analysis.

d) Please also provide a scenario analysis in which this treatment effect is set to 1.

Please see our answer to question B5 (g). Removal of all treatment effects post 4-years increases the ICER to [redacted].

B 14. Priority question: No treatment waning is assumed in the base-case. However, given that pembrolizumab treatment is stopped at 2 years max, this assumption is questionable. The treatment waning scenario is only applied to the no-/failed SCT patients.

a) To give an indication of potential treatment waning on pembrolizumab, please provide smoothed hazard plots over time with patient numbers at risk as observed in the Kaplan Meier data for the randomized KEYNOTE-204 evidence and the KEYNOTE-087 evidence.

As KEYNOTE-087 does not have a comparator arm we did not consider that this would provide any meaningful evidence to examine treatment waning compared with standard of care.

Given that there are very few events among the no-SCT group after 18 months in KEYNOTE-204 and OS remains very high in both arms, smoothed hazard ratio plots are unlikely to help with this assessment. Nevertheless, this is provided below and should be interpreted with caution. We note that both HRs are low and statistically significant. The OS in the ITT population of KEYNOTE-204 is, of course, heavily influenced by the high number of SCTs in both arms. There would be very few SCTs on SoC in the 4L setting so any signal of hazards equalising over time in KEYNOTE-204 does not indicate treatment waning is expected in the 4L setting comparing pembrolizumab with SoC (not that we would infer that it is happening in KEYNOTE-204 from these data anyway).

Figure 6. Smoothed OS hazard ratio among patients who never had an SCT in KEYNOTE-204 including confidence intervals

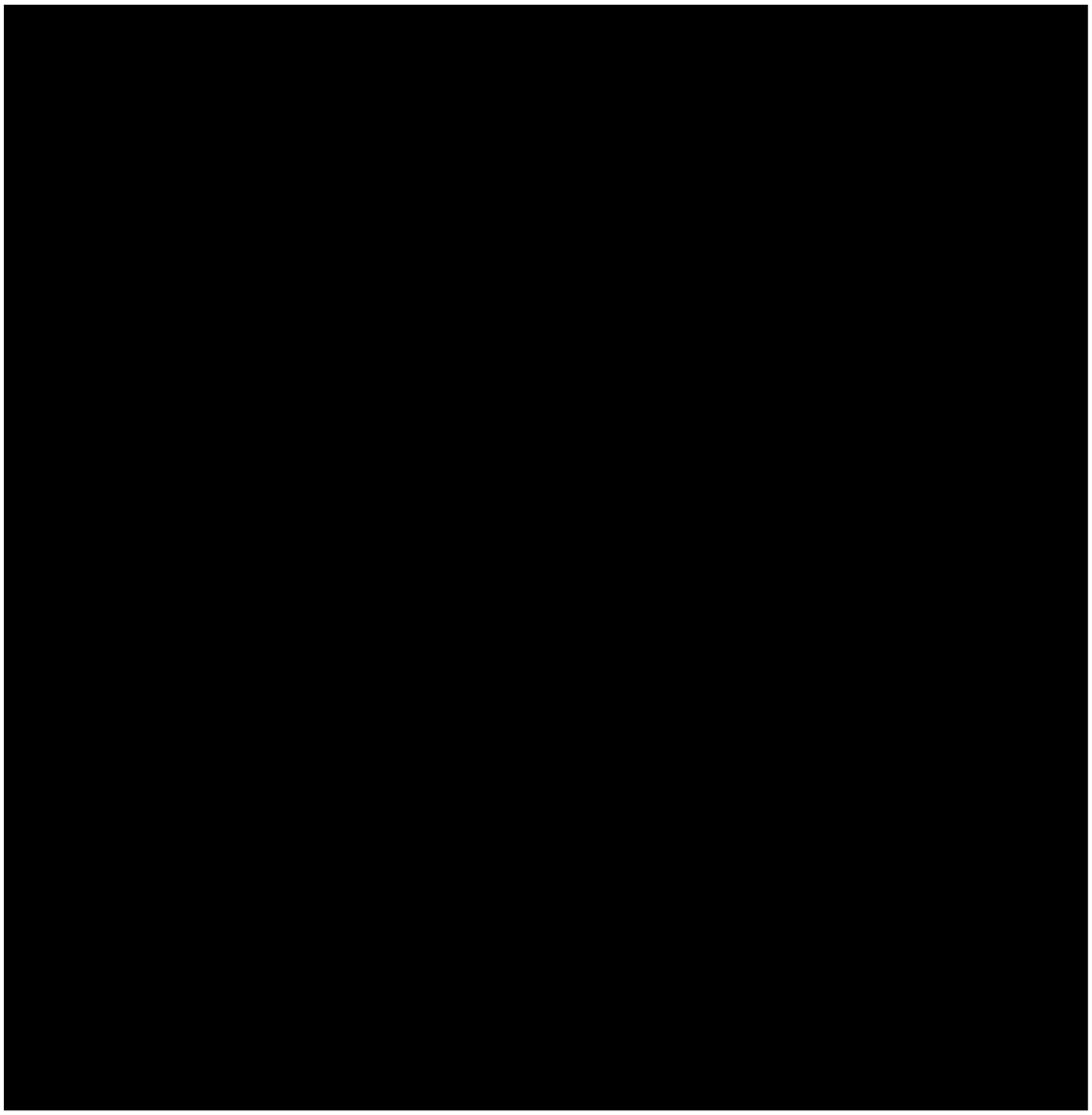


Figure 7. Smoothed OS HR in ITT population of KEYNOTE-204 including confidence intervals

b) Please justify the assumption of no treatment waning, i.e. that there is a lifetime difference in PFS and OS based on the initial treatment

Please see the response to B13c above. Extended difference in OS is only applicable in the no/failed SCT group as patients who are “cured” follow general population mortality in both treatment arms. There is no extended difference in PFS as there is no progression-free health state in the model. The economic model includes the ability to wane both the mortality and utility treatment effects in sensitivity analysis.

c) Please provide a scenario in which treatment waning starts at the time that treatment is stopped (2 years), i.e. hazard ratio of 1 in the pre-landmark OS

We have not provided this as it is contrary to established precedent in previous NICE appraisals of pembrolizumab (e.g. TA885[19]), where hazards equalise from 3 years

after treatment cessation to 5 years (i.e. years 5-7 in model time). Before the landmark, HR is inclusive of both the effect of ongoing pembrolizumab treatment and of the effect of some patients receiving curative SCTs and therefore this analysis does not conceptually fit with the model structure.

B 15. A structured expert elicitation was performed to inform key model inputs. Please elaborate on whether the methods used are aligned with the Medical Research Council (MRC) SEE protocol (https://pubmed.ncbi.nlm.nih.gov/34105510/).

A full description of the methodology of the SEE is described in Appendix N and summarised in section B.3.3.2 (page 108) of the CS. A pre-specified protocol was developed, the methods of which were aligned with the MRC protocol, as summarised in table 36 of the CS.

Deviations from the MRC protocol were noted and reported in Appendix N. In summary:

- A full pilot exercise of the SEE exercise with clinicians using the Excel based tool was not feasible. Instead, one clinician reviewed the questions in text format. Based on this feedback questions were modified to improve clarity.
- One expert was unable to provide individual responses to the SEE exercise due to IT issues but contributed to the group discussion session.

These were considered minor deviations from the protocol that are unlikely to influence the results of the exercise.

Adverse events

B 16. Please comment on the face validity of the much higher AE rates in the SoC arm of the model compared with the pembrolizumab arm (Table 40 vs Table 38 of the CS), providing supporting evidence. Please include a scenario in the model file in which the AE rates are set equal in both arms.

As KEYNOTE-087 is a single arm trial, the AE incidences for SoC were sourced from the literature in line with the TA540 methodology. This slightly deviated in the CDF exit following the SoC composition update from Eyre et al. (2017) and the Clinical Advisory Board validation as stated in section B.3.2.5 of the CS. The AE incidences listed in

Table 39 are not commented on individually for face validity given the number of studies and the very minimal effect these data have on the ICER. However, the sample size and number of AEs experienced are summarised in the aforementioned table. In addition, this methodology was originally accepted by the EAG and received no further comment by the NICE Committee in TA540. As stated in the CS, Table 2, the higher AEs in Table 40 is expected given pembrolizumab has a well-tolerated safety profile.

Health-related quality of life

B 17. Priority question: Regarding the use of KEYNOTE-204 for estimating health-related quality of life in the CS base-case. As mentioned in the CS, there are multiple limitations regarding the use of KEYNOTE-204, as it does not reflect the target population, nor the appropriate comparator and assumptions are required to estimate the impact of SCT (see question B22). This is partly reflected in the CS statement that *"We note that the data are from an earlier line of therapy not directly related to the health states within the economic model and therefore these estimates are a source of uncertainty. This is counter-acted to some extent by BV being a conservative surrogate for SoC"*. Please justify the appropriateness of using KEYNOTE-204 for estimating health-related quality of life in the CS base-case, given the limitations mentioned above.

The company acknowledge that KEYNOTE-204 is in an earlier of therapy, 3L as opposed to 4L in the decision problem (although approximately 37% were actually 4L+ in the study). However, KEYNOTE-204 provides value in that it is the only RCT in this population thus providing a direct treatment comparison. Patients in KEYNOTE-087 have not received pembrolizumab, or any other immunotherapy, in a previous line of therapy and thus you could expect a similar response to this new mechanism of action. This is demonstrated in the similar overall response rates observed between KEYNOTE-204 (3L+ with no prior SCT) and KEYNOTE-087 (Cohort 2), 61.8% vs. 64.2%, respectively. This has been confirmed and explained by a clinical expert as patients with immunotherapy resistance will not yet have been selected out of the pathway. On the other hand, patients in 4L SoC would have previously failed BV and therefore any additional chemotherapy would be expected to be even less effective than that experienced in 3L as these patients would already have chemoresistance

selected from previous lines of chemotherapy. In conclusion, patients receiving pembrolizumab in 4L would be expected to respond similarly to those in 3L, whereas patients receiving chemotherapy in 4L would expect a worse response than in 3L. As such, the effect on quality-of-life demonstrated in KEYNOTE-204 between pembrolizumab and SoC can reasonably be assumed to be at least as strong in the 4L.

B 18. Priority question: CS Table 42 reports the utility values based on KEYNOTE 087.

a) Please describe in detail the procedure used to estimate these utility values. Including an overview of the data included, how missing data were handled, how diagnostics of the regression model were assessed, how the (candidate) covariates as well as interaction terms were selected (with rationale) and how the regression model accounted for nesting effects.

The EQ-5D-3L was collected i) 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, ii) on treatment discontinuation and iii) 30 days post treatment discontinuation. The EQ-5D-3L was generated using the UK value set derived by Dolan et al. (1997).[20] The utility values were estimated using the means and not a regression model. MSD confirm the KEYNOTE-087 utilities were only collected from non-missing data and no imputation strategies were implemented at the time of the analysis.

b) Please provide the results of utility values estimated based on KEYNOTE 087 using a mixed effects model (including relevant covariates if available such as SCT status), as well as a detailed description of the procedure used to estimate these utility values (see previous sub-question).

The KEYNOTE-087 utilities were not used in the economic model in either the base case or as a scenario analysis. We do not have any linear mixed effects modelling data to share but sensitivity analyses can be undertaken if anything about the likely the magnitude and direct of effect of such an analysis is known.

B 19. Priority question: CS Table 43 reports the utility values based on KEYNOTE 204.

a) Please describe in detail the procedure used to estimate these utility values. Including an overview of the data included, how missing data were handled, how diagnostics of the regression model were assessed, how the (candidate) covariates as well as interaction terms were selected (with rationale) and how the regression model accounted for nesting effects.

The values were simple naïve means. We have attached a report where the results of many more advanced regression modelling techniques are reported. We note that in the final regression model the only variable that appears to statistically significantly predict utility is study arm. The results of a more parsimonious linear mixed effects model including only study arm as a predictive variable are also presented within this report and give coefficients of 0.73 for BV and +0.085 for pembrolizumab. These data suggest that estimates of average utility by study arm using this model would be similar to, although slightly more conservative than, the naïve means, which were 0.742 and 0.837.

b) Please provide the results of utility values estimated based on KEYNOTE 204 using a mixed effects model (including relevant covariates if available such as SCT status), as well as a detailed description of the procedure used to estimate these utility values (see previous sub-question).

See also response to part a) above. MSD have attached a report exploring utility regression models that included various covariates such as age, continuous age (centralised at 40 years of age), gender, post treatment SCT status, treatment and grades 3-5 AEs. A final multivariate model is also presented including continuous age (centralised at 40 years of age), SCT status post-treatment and grade 3-5 AE status with alternative final models for exploratory analyses. The only independent variable that is statistically significant in the final model is treatment arm.

B 20. In the CS it is stated “because it was established during NICE TA772 and TA540 that pembrolizumab has a treatment effect on utility as well as disease progression and because the assumption of a persistent utility treatment effect was validated by clinicians at the UK clinical advisory board.” Please provide supporting evidence for this statement and justify that this is applicable in the current decision problem.

Please see our previous response to question B.5.a) regarding clinical expert validation of the assumption of persistent utility benefit for pembrolizumab for patients who do not receive an SCT.

In TA540 a small utility difference of 0.011 was applied between pembrolizumab and SoC in the progression-free health state.

Based on statistically significant data from KEYNOTE-204, it was accepted in TA772 that there was a utility benefit for pembrolizumab pre-progression. The Committee also concluded that there was likely to be a post-progression utility benefit for pembrolizumab compared with BV, but that this was unlikely to persist for the whole period of progression. The Committee favoured using the EAG base case where no utility benefit was assumed but noted this may be conservative. TA524 established BV as a more effective treatment than SoC. It follows that pembrolizumab must have a benefit on utility vs. SoC.

To explore the uncertainty in this area further we have conducted a scenario analysis where utility benefit for pembrolizumab wanes between 3 and 5 years post treatment cessation (in line with scenarios on OS HR treatment waning). This results in an ICER of [redacted], compared with [redacted] in the base when no utility waning is assumed and [redacted] when the utility benefit is removed entirely.

B 21. In CS Table 42, the ‘ongoing treatment’ utility was reported to be lower than the ‘completed or discontinued treatment’ utility (0.836 versus 0.860).

a) Please elaborate on the possible mechanism explaining the potentially higher ‘ongoing treatment’ utility compared with ‘completed or discontinued treatment’ utility population utility.

Table 42 also shows that the estimate for 'completed or discontinued treatment' is drawn from 21 EQ-5D forms as opposed to 383 EQ-5D forms for ongoing treatment. It may also be that patients who stopped treatment and filled in a form where more likely to have achieved an SCT, which might increase their utility but the direction of bias that might have influenced this small number of responders is not clear.

b) Please justify the pembrolizumab utility increment, given the utility increases when pembrolizumab was stopped.

See response above. We would not draw that conclusion from these data. The pembrolizumab utility increment is a statistically significant finding from a large and relevant RCT.

B 22. In the CS it is assumed that progressed disease utilities could be used for the group who had either never received or relapsed after SCT after the landmark. Please provide supporting evidence for this assumption, also considering the statement in CS section B.3.2.4.2 *"On review of the pembrolizumab datasets, PFS appeared not to be a reliable surrogate for either OS or whether patients receive an SCT"* which appears to contradict this assumption.

This is not a contradiction as the reliability of PFS as a surrogate for either OS or for receiving an SCT does not have any bearing on the health-related quality-of-life for the specific group of patients who never receive an SCT or relapse after SCT.

Patients who have failed an SCT (i.e. relapsed after an SCT) can reasonably be characterised as having had a PFS event. For those who have not had an SCT, it may be reasonable to assume most of them have progressed by 4 years as it can be seen in KEYNOTE-087 that PFS is relatively short (median <12 months and ~15% at 4 years) and that trial censors patients who have an SCT so the real values may be lower among the non-SCT cohort.

MSD had no source of direct evidence for utility in this group and acknowledge that there are several limitations to using the progressed disease utility values from KEYNOTE-204. The main limitation is that the data from KEYNOTE-204 are from an earlier line of therapy not directly related to the health states within the model. Nevertheless, these data present the best source of evidence to estimate the quality-

of-life benefit for pembrolizumab and despite uncertainties may be a conservative estimate given that they are based on BV as a surrogate for standard of care. Clinical experts consulted by MSD noted that they observed clinically significant quality-of-life improvements for patients treated with pembrolizumab compared with standard of care that can persist for many years after treatment.

B 23. Please comment on the extent to which the utility values applied for patients treated with pembrolizumab and SoC already capture the impact of adverse events (for example, because patients with adverse events were included in the utility values obtained from the respective studies). Please provide a scenario where no disutilities are applied for adverse events.

The possibility that including AE-related disutilities may lead to a small amount of double-counting for the reasons outlined above was acknowledge in the company submission (section B3.5.5, page 132). A scenario removing AE disutilities was already provided and results in an ICER of [redacted] per QALY gained. This scenario can be replicated by setting the “include adverse event disutility” option to “exclude” in the “model settings” tab.

B 24. General population utility values (gender and age matched) are used for patients in the “cured-SCT” health state. Considering these patients did have (had) multiple treatments, including SCT, and cHL. Using general population utility values implicitly assumes the consequences of the treatments and health condition on health-related quality of life are fully reversible. Please elaborate on the appropriateness of this assumption and provide supporting evidence.

As discussed in Section B.3.5.3, patients in the cured-SCT health state after the landmark were assigned age-matched general population utility, irrespective of health condition, in line with the Ara and Brazier utility set. This general population utility accounts for underlying health conditions in the population. Therefore, by implementing this utility it does not implicitly assume a full reversal of the health condition and treatment received, but instead acknowledges that a patient’s utility can be impacted by any health condition and previous treatment. It is worth noting that patients with R/R cHL in this line of treatment do not appear to have HRQoL that is markedly lower than the general population and require only a modest improvement to return to population norms. Given that patients in this health state are assumed

cured of R/R cHL, an improvement in utility compared to the pre-landmark health state would be expected, as is demonstrated by utility of 0.837 pre-landmark vs 0.864 post landmark for the cured-STC health state. The implementation of an improvement of this small magnitude does not seem unreasonable. Furthermore, as noted in Section B.3.5.3, the baseline age for deriving the general population utility was taken from the KEYNOTE-087 cohort 2 population, as opposed to the population who received an SCT who are younger and fitter, and therefore is a conservative estimation and impacts the ICER not in favour of the intervention.

B 25. The SCT related disutility is based on TA540. Please elaborate how transferable the calculated disutility (in CS Table 48) is to the current decision problem, reporting on differences in population, standard of care as well as (expected) proportions auto-SCT versus allo-SCT.

MSD would firstly like to clarify that the source of SCT-related disutility in the company submission, as presented in Table 48 of the CS, was adapted from TA524, not TA540.

The rational and methodology for adapting SCT related disutility from TA524 is discussed in Document, Section B.3.5.6. As noted in priority question A7, the current decision problem represents a sub-set of patients from TA524, R/R cHL patients post-BV and ineligible for SCT and therefore MSD considers the population to be transferable. The disutility is specific to the decrement experienced whilst undergoing SCT, regardless of baseline utility, and therefore little variation might be expected between 3L and 4L. Whilst there is no gold standard SoC in this line of treatment, both TA524 and ID5084 SoC comprises of a blend of chemotherapy regimens. Similarly, the SCT-related disutility is specific to the SCT procedure itself and therefore is unlikely differ substantially by chemotherapy regimen. As noted in the company submission, the adapted SCT related disutility did not account for proportions of allogeneic vs autologous SCT. The base case for the economic model used SCT proportions from the SACT data (65% allogeneic vs 35% autologous), similarly in TA524 the data collected from the CDF showed that 45 patients underwent alloSCT and 33 underwent autologous SCT which equates to 58% allogeneic and 42% autologous.[21] The breakdown is not provided for patients who received salvage chemotherapy after BV and before SCT. From the available data, the proportions are fairly similar, with a slightly higher proportions of allogeneic SCTs in the current

decision problem who are more likely to experience SCT-related difficulties or adverse events such as chronic graft vs host disease. We haven't accounted for this explicitly in the model but the effect on the ICER would be very small. In this evidence light area, whereby no alternative utilities specific to R/R cHL were identified in the literature, MSD consider this SCT-related disutility derived from TA524 to both transferable and the best available to inform the economic model. It is not an important driver of the ICER but may be examined in sensitivity analysis.

Cost and resource use

B 26. As per the marketing authorisation, patients treated with pembrolizumab are treated until disease progression is confirmed, if unacceptable toxicities occur or if they reach NHS England's 35-cycle (24 months) stopping rule. Explain why few patients in the model continued the treatment beyond the 24 months stopping rule.

ToT data for pembrolizumab in the model is derived from the SACT database whereby a few patients remained on treatment beyond 24-months. In pembrolizumab trials, this usually captures patients who are taking a treatment break but resume treatment again afterwards, so are still within the 35-cycle stopping rule, and is usually the cause of a small tail on the ToT KM curve after 2 years and this is a plausible explanation for why a few patients remain on treatment beyond 24 months in the SACT. However, within the economic model there is a hard stopping rule in place at 24 months (104 weeks) within the economic model which is a simplified method that prevents receiving pembrolizumab after this point in-line with NHS England guidelines. This is evident in the model trace for the intervention in tab 'trace_treatment1' where the ToT (column E and G) drops to '0.00' at cycle 105 (row 120). As a result, no patients in the intervention arm receive pembrolizumab treatment after this point.

B 27. Radiotherapy was added in the SoC composition and applied as a one-time cost in the cost-effectiveness model based on the Advisory Board recommendation. However, radiotherapy was not included in the NICE final scope for treating patients with R/R cHL. Furthermore, data about resources and costs of radiotherapy in cHL is limited, and the MSD based the general approach on published guidelines for other cancer sites (Lung Cancer (NG122) and the non-

small cell lung cancer (NSCLC)). Please justify the appropriateness of including radiotherapy in the SoC and provide supporting evidence.

We do not have any additional justification to provide and suggest that the composition of the blended SoC should be the subject of sensitivity analysis.

B 28. A lower subsequent therapy proportion was assumed for the pembrolizumab arm based on feedback from experts in the advisory board and the proportions in KEYNOTE-204. However, the treatment composition in the subsequent therapy was based on TA306 along with data derived from SACT and clinical experts and not the composition from KEYNOTE-204. Please justify the use of different sources to inform the composition and proportions of subsequent therapy in the model.

It is normal in NICE Technology Appraisals to cost subsequent therapies as those that are available and in use in the UK NHS and we have followed that methodology here.

B 29. In the SoC arm only, nivolumab was included as a subsequent treatment based on NICE TA462 which led to a significantly a higher cost for subsequent treatments in the SoC arm when compared to the pembrolizumab arm. Please justify the reason behind not adding nivolumab as a subsequent treatment to the pembrolizumab arm.

Following the NICE TA462 guidance that recommends Nivolumab for use in the R/R cHL in patients' who have previously received BV and failed autoSCT, Nivolumab is included as a subsequent treatment in the SoC arm for the proportion of patients of have received, and failed, auto SCT. Given the few patients that receive SCT in the SoC arm, this equates to only 0.07% of the subsequent treatment distribution and thus has little impact on the total subsequent treatment cost in the SoC arm. Nivolumab has not been included as a subsequent treatment in the intervention arm as it is unlikely to be used in clinical practice following pembrolizumab given that the clinical evidence for NICE TA462 was based on the CheckMate 205 trial whereby the exclusion criteria included 'checkpoint inhibitor at any time before nivolumab treatment'.[22] Therefore, the clinical evidence does not support the use of Nivolumab following BV, autoSCT, and pembrolizumab, in any given order.

In fact, the transparency benefits that arise from excluding nivolumab (and its confidential PAS price) from the model entirely might outweigh the benefits of including it for this tiny group in the SoC arm.

Validation

B 30. Priority question: Further external validation of modelled effectiveness would be desirable.

a) Please report on the face validity assessment of the model structure, model assumptions, model inputs, intermediate outcomes as well as final outcomes in more detail (including what aspects were assessed and what were the considerations as well as conclusions).

OS to 4 years and the proportion getting SCT are based on every patient treated with pembrolizumab in UK clinical practice in the last 4 years. The ITCs produce variable estimates of OS for SoC at 4 years but all are within the confidence interval of the combined estimate given by experts during the SEE (6.9%-38.3%). The SoC estimates are also validated as conservative (i.e. potentially overestimated) by the "no-SCT" groups in Eyre and Cheah; given that very few patients have an SCT in our model, these sub-populations might be considered representative and yet have similar or worse OS to the model despite being 20 years younger on average than our patients. The data from these studies suggest SoC OS might be overestimated in the model. The proportion receiving SCT on SoC was elicited directly from a pool of experts using SEE. The fact that it is so low reflects the chemo-insensitive nature of this patient population, who have already failed three lines of chemotherapy-based treatment. Cure proportions were also elicited using SEE. Outcomes for cured patients (OS/HRQoL) being in line with the general population were confirmed at the advisory board. Extrapolations for the no/Failed SCT group were discussed with one expert who stated that OS would be ~0% at 8-10 years among these patients but expected a number of patients who had received pembrolizumab to still be alive at this time, even though they had not had an SCT. This suggests the exponential curve would be the best extrapolation to use for this group whereas either an independent log-logistic curve or an exponential curve without treatment waning might be appropriate for the pembrolizumab arm. Unfortunately we did not have time to implement separate parametric extrapolations for the separate arms in the economic model. The treatment

effect on HRQoL prior to the landmark was validated with a clinical expert who confirmed it is conservative as BV is a more effective treatment than SoC. The continuation of HRQoL benefit for pembrolizumab in the "No/Failed SCT" group was also confirmed as plausible by the same advisor but it was stated that it is not known how long this lasts. As described in the CS section B.3.14.1 the model produces OS estimates that are similar to simply extrapolating the observed SACT OS using a log-normal curve. The model is conceptually superior to a simple parametric extrapolation because it explicitly takes account of the patients who are cured by SCT but it is reassuring that there is a simple and methodologically orthodox method produces the same results.

b) Please comment on the company's modelled estimate for OS on SoC at 4 years being much lower and not in line with estimated mean by the experts reported in Table 37.

This is already addressed in the CS section B.3.14. Briefly, we felt that the estimation of 4-year OS via comparative evidence was conceptually superior to that elicited from experts. It was also validated by the data from Eyre on patients who failed BV and didn't get an SCT where OS was just 25% at 20 months. Our SoC arm are patients who failed BV and only 8% got an SCT so match relatively closely. OS is 28% at 20 months in our model. This Eyre subgroup were from a study with median age ~20 years younger than those receiving pembrolizumab in SACT and 100% were considered potentially fit for transplant vs. only 60% in the SACT dataset. This Eyre subgroup (and consequently our model, given the data are similar) therefore potentially overestimates survival on SoC compared to the theoretical SoC cohort we are trying to model. We also note that the OS estimates produced by the model are within the confidence interval obtained via the SEE (6.9% - 38.3%). That said, we recognise there are uncertainties in the evidence and feel that sensitivity analysis on the OS HR is helpful for decision-making. ITCs with HRs around 0.4 produced 4-year SoC OS in line with expert opinion.

c) Please assess the external validity of model inputs, intermediate outcomes as well as final outcomes using:

i. evidence used to develop the economic model (e.g., evidence used in the base-case)

ii. evidence not used to develop the economic model using alternative real world evidence data sources, if available (e.g., evidence used in scenarios).

Part (i) of this question is covered extensively in the CS and in our answers to other Clarification Questions. For part (ii), we are not aware of any additional evidence that could/should have been included in our model.

B 31. It appears that no technical verification was performed.

a) Please provide a detailed description of the validity assessment performed as well as the results.

b) Please complete the TECH-VER checklist (Büyükkaramikli et al. 2019, https://pubmed.ncbi.nlm.nih.gov/31705406/) and provide the results.

We have attached an internal technical verification checklist. No technical issues were identified.

B 32. Please provide cross validations, i.e., comparisons with other relevant NICE TAs focused on similar, potentially relevant, diseases (e.g., related NICE recommendations and NICE Pathways listed in the final scope) and TA540 and elaborate on the identified differences regarding:

a) Model structure and assumptions

b) Input parameters related to:

i. Clinical effectiveness

ii. Health state utility values

iii. Resource use and costs

c) Estimated (disaggregated) outcomes per comparator/ intervention

i. Life years

ii. QALYs

iii. Costs

A summary of the key features of the economic analysis is provided in Document B (Table 30. Features of the economic analysis) for the company submission and the following NICE TAs: TA462 (Nivolumab for treating relapsed or refractory classical Hodgkin lymphoma), TA524 (Brentuximab vedotin for treating CD30-positive Hodgkin lymphoma), TA540 (Pembrolizumab for treating relapsed or refractory classical Hodgkin's lymphoma), and TA772 (Pembrolizumab for treating relapsed or refractory classical Hodgkin lymphoma after stem cell transplant or at least 2 previous therapies).

Candidate model structures are also discussed narratively in the Model Structure section of the CS. Briefly, there is no standard model structure in either cHL appraisals or in haematology-oncology appraisals more generally. Complex model structures have been criticised for being untransparent (due to the computational complexity of incorporating many downstream time-dependent transition probabilities) and relying on highly uncertain indirect evidence and linked evidence approaches[23]. There are very few studies that provide directly relevant data, as highlighted by the dearth of information in the SLRs conducted for this appraisal. A fuller explanation of the differences between the new model and the one considered during TA540 is provided in response to an earlier question.

The disaggregated outcomes, as available in the Committee papers, for the NICE TAs listed above provided in the table below by intervention and comparator. Only life-years (LYs) are consistently reported although sometimes they are discounted and sometimes not, which makes comparisons difficult. The total LYs on SoC being lowest in our model is reasonable, given that patients in our line of treatment are the most chemo-refractory. Among the scenario analyses we presented in section B.3.11.3 of the CS, undiscounted life years can be as high as 4.14 on SoC and as low as 7.78 on pembrolizumab. R/R cHL is an area where the specific proportions of patients getting an SCT has a heavy influence on LYs gained as a cured patient may live for 30+ years.

Table 4. Comparison of disaggregated costs, LYs, and QALYs with previous NICE TAs

Outcome	TA462*[23]	TA524[21]	TA540[1]	TA772[18]	Company submission
Total costs (£)					
Intervention	NR	NR	86,855*	NR	[redacted]*
Comparator	21,090*	NR	32,217*	NR	[redacted]*
Total LYs					
Intervention	5.013*	12.39	7.94 5.43*	4.98*[x]	10.96 7.47*
Comparator	2.110*	4.48	4.36 3.24*	4.98*[x]	2.24 1.81*
Total QALYs					
Intervention	NR	NR	3.15*	4.11*	[redacted]
Comparator	0.932*	NR	1.74*	3.52*	[redacted]

***Abbreviations**: LY, life-year; QALY, quality-adjusted life-year*

**Discounted*

[x]Equal OS assumed and taken from external source (no OS data was available for either arm and pembrolizumab was cost-effective without needing to model an OS benefit so these data are no more than placeholder value)

Cost-effectiveness results

B 33. Priority question: Please provide a model file with settings that replicate the company base-case as the current ICER in the model is significantly higher than the one reported in the report, and it is unclear what settings contribute to this.

The model will produce the base case once the confidential PAS discount for pembrolizumab is applied. This is located in cell C523 of the References sheet.

B 34. Please provide all results also excluding the severity modifier.

There is a switch in cell D6 on the All Results sheet which can be set to 1 to remove the severity modifier.

Section C: Textual clarification and additional points

C 1. Priority question: Table 53 in the supplementary materials I3 is not working.

MSD have supplied the table as a separate file in our response to the clarification questions.

References

1.NICE. 2018. Pembrolizumab for treating relapsed or refractory classical Hodgkin lymphoma. Last accessed: 17 May 2023. Available from: https://www.nice.org.uk/guidance/ta540/chapter/1-Recommendations.
2.MSD UK Ltd. Classical Hodgkin Lymphoma Advisory Board Meeting, London, Monday 10 July 2023. MSD data on file: 2023. 2023.
3.NICE. 2023. Pembrolizumab for treating relapsed or refractory classical Hodgkin lymphoma [Review of TA540] [ID5084]. Last accessed: 23 August 2023. Available from: https://www.nice.org.uk/guidance/topic-selection/gid-ta11317/documents.
4.von Tresckow B, Fanale M, Ardeshna K, Chen RW, Meissner J, Morschhauser F, et al. Patient-Reported Outcomes in Patients with Classical Hodgkin Lymphoma Treated with Pembrolizumab Monotherapy, Results of a Phase 2 Study. *Blood* 2017;130:2169.
5.Dada R, Zabani Y. Nivolumab in relapsed/refractory classic Hodgkin Lymphoma; experience with ten patients. *Hemasphere* 2018;2:42.
6.Ionova T, Andrievskikh M, Amdiev A, Barykah E, Chang V, Endakova A, et al. Brentuximab vedotin for treatment in patients with relapsed/refractory classical Hodgkin Lymphoma in a real world setting: clinical outcomes and impact on quality of life. *Hematol Oncol* 2021;39:442.
7.Engert A, Taylor F, Bennett B, Chen C, Cocks K, McDonald J, et al. Effect of Nivolumab on Patient-Reported Outcomes in Patients with Relapsed/Refractory Classical Hodgkin Lymphoma after Autologous Transplantation: Results from the Multicohort Phase 2 Checkmate 205 Study. *Blood* 2017;130:3441.
8.Husson O, Hadewijch de Rooij B, Kieffer JM, Oerlemans S, Mols F, Aaronson NK. Independent prognostic value of the EORTC QLQ-C30 summary score on all-cause mortality: Results from the population-based PROFILES registry. *J Clin Oncol* 2018;36(15):10070.
9.Ionova T, Alanassiev B, Andrievskikh M, Amdiev A, Barykah EA, Chang V, et al. Response to Brentuximab Vedotin and Quality of Life in Patients with Relapsed/Refractory Hodgkin Lymphoma (RR HL) in the Real World Setting. *Blood* 2019;134:5296.
10.Ionova T, Andrievskikh M, Amdiev A, Barykah E, Chang V, Endakova A, et al. Outcomes of brentuximab vedoting as >=3 line treatment in patients with relapsed/refractory classical Hodgkin Lymphoma: physician's and patient's perspective. *Hemasphere* 2021;5:572.
11.Lepik K, Mikhaylova N, Kondakova E, Tsvetkova L, Zalyakov Y, Borzenikova E, et al. Response to nivolumab as >= 3rd line therapy in pts with relapsed/refractory classical Hodgkin's Lymphoma (cHL) and its impact on quality of life in responders and nonresponders. . *Hematol Oncol* 2019;37:495.
12.Stadtbaemur N, Kreissl S, Borchmann P, Mayer A. Predicting cancer-related fatigue of Hodgkin Lymphoma survivors: Identification of risk factors. *Oncol Res Treat* 2021;44:V596.
13.Stadtbaemur N, Mayer A, Borchmann P. P078: Predicting the health-related quality of life of Hodgkin Lymphoma survivors: identification of risk factors. *Hemasphere* 2022;6:36.
14.Jones B, Ward T, Harrison JP, Hurst M, Tyas D, McEwan P, et al. The cost effectiveness of nivolumab for the treatment of people with relapsed or refractory classical Hodgkin Lymphoma following autologous stem cell transplant and brentuximab vedotin. *Value Health* 2017;20:A433.

15.NICE. 2017. Pembrolizumab for treating PD-L1-positive non-small-cell lung cancer after chemotherapy. Last accessed: 29 August 2023. Available from: https://www.nice.org.uk/guidance/TA428/chapter/1-Recommendations.
16.NICE. 2018. Pembrolizumab for untreated PD-L1-positive metastatic non-small-cell lung cancer. Last accessed: 27 October 2023. Available from: https://www.nice.org.uk/guidance/ta531/chapter/1-Recommendation.
17.NICE. 2021. Pembrolizumab with pemetrexed and platinum chemotherapy for untreated, metastatic, non-squamous non-small-cell lung cancer. Last accessed: 27 October 2023. Available from: https://www.nice.org.uk/guidance/ta683/chapter/1-Recommendations.
18.NICE. 2022. Pembrolizumab for treating relapsed or refractory classical Hodgkin lymphoma after stem cell transplant or at least 2 previous therapies. Last accessed: 17 May 2023. Available from: https://www.nice.org.uk/guidance/TA772/chapter/1-Recommendations.
19.NICE. 2023. Pembrolizumab plus chemotherapy with or without bevacizumab for persistent, recurrent or metastatic cervical cancer. Last accessed: 27 October 2023. Available from: https://www.nice.org.uk/guidance/TA885/chapter/1-Recommendations.
20.Dolan P. Modeling valuations for EuroQol health states. *Med Care* 1997;35(11):1095-108.
21.NICE. 2018. Brentuximab vedotin for treating CD30-positive Hodgkin lymphoma. Last accessed: 17 May 2023. Available from: https://www.nice.org.uk/guidance/ta524/chapter/1-Recommendations.
22.Armand P, Engert A, Younes A, Fanale M, Santoro A, Zinzani PL, et al. Nivolumab for Relapsed/Refractory Classic Hodgkin Lymphoma After Failure of Autologous Hematopoietic Cell Transplantation: Extended Follow-Up of the Multicohort Single-Arm Phase II CheckMate 205 Trial. *J Clin Oncol* 2018;36(14):1428-39.
23.NICE. 2017. Nivolumab for treating relapsed or refractory classical Hodgkin lymphoma. Last accessed: 28 July 2023. Available from: https://www.nice.org.uk/guidance/ta462/chapter/1-Recommendations.

Addendum to Clarification Question A28e (and context for several other questions) - Confidential

Here we provide some additional evidence from KEYNOTE-204, which was a trial of pembrolizumab vs BV in R/R cHL patients who are 3L+. The EAG have highlighted concerns about the generalisability of this trial to the 4L setting. Although the generalisability of evidence from this trial was confirmed at the MSD UK clinical advisory board, we thought this additional data might be helpful.

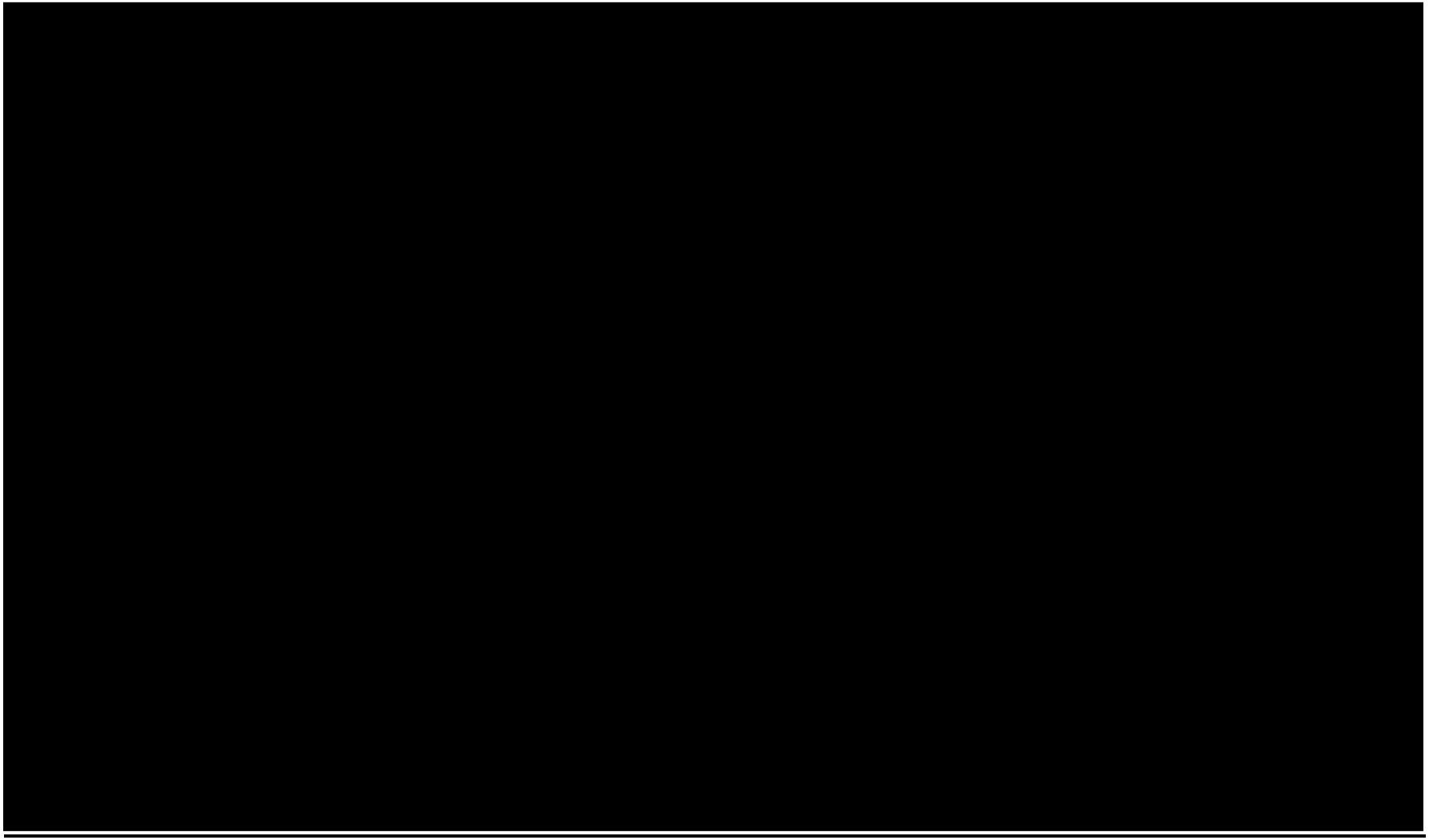
Table 1 and Figure 1 show OS among the 3L only group. Table 2 and Figure 2 show OS in the 4L+ group. Patients who are 4L+ have failed to respond to one additional line of chemotherapy when compared to patients who are 3L.

Line of therapy was not a stratification factor in the trial and the number of events is relatively low so results must be interpreted with caution. Despite this, we interpret these data as being indicative that outcomes on pembrolizumab are unlikely to be meaningfully worse among patients who have failed to respond to an additional line of chemotherapy. The data suggest it is possible that outcomes on BV are slightly worse among the 4L+ patients. This is clinically plausible as these patients are, on average, expected to be more chemorefractory, compared to 3L patients and BV is a chemotherapy based treatment. The relative effectiveness of pembrolizumab is therefore potentially greater.

While uncertain, these data are broadly supportive of the effectiveness from KEYNOTE-204 being applicable and likely conservative when used in the economic model in the 4L+ setting as a surrogate for the effectiveness of pembrolizumab vs chemotherapy, as was done in the CS.

Addendum to Clarification Question B11 on proportional hazards

The below graphs provide the requested model diagnostics for the proportional hazards assumption for the two HRs from KEYNOTE-204 that we used in the economic model (the overall OS HR in the trial and the HR among the subgroup who never had an SCT). We note that in neither case is the proportional hazards assumption obviously violated.



Keytruda (MK-3475)

KN087

A Phase II Clinical Trial of MK-3475 (Pembrolizumab) in Subjects with Relapsed or Refractory (R/R) Classical Hodgkin Lymphoma (cHL)

Indirect Treatment Comparisons (ITCs) & Diagnostics for MSD-UK

Pembrolizumab vs. Standard of Care (SoC) intervention(s)

Database Cutoff Date 15-MAR-2021

Version 0.1

BARDS HTA

SUMMARY OF CHANGES

Version	Date	Description of changes
2.0	**23 OCT 2023**	**Log HR over time plots added as per EAG request**

Table of Contents

1 INTRODUCTION

This Health Technology Assessment (HTA) report is the reference document for Indirect Treatment Comparisons (ITCs) as well as corresponding diagnostic plots of pembrolizumab monotherapy versus Standard of Care (SOC) intervention(s) in participants with relapsed or refractory (R/R) Classical Hodgkin Lymphoma (cHL) who have failed to respond or relapsed on brentuximab vedotin (BV) and are ineligible to auto-SCT (Cohort 2 of KN087). The analyses are performed by BARDS HTA in support of the UK Keytruda KN087 submission to NICE.

2 OBJECTIVE

To assess efficacy in participants with relapsed or refractory (R/R) Classical Hodgkin Lymphoma (cHL) who have failed to respond or relapsed on brentuximab vedotin (BV) and are ineligible to auto-SCT (Cohort 2 of KN087) via the use of indirect treatment comparisons (ITCs) and to compare pembrolizumab versus selected Standard of Care (SoC) interventions for Overall survival (OS).

Following ITCs will be presented in this:

- Pembrolizumab (based on KN087 (cohort 2)) vs. BV (based on Eyre et al. 2017)
- Pembrolizumab (based on KN087 (cohort 2)) vs. BV (based on Eyre et al. 2017) (adjusted analysis and unadjusted analysis as sensitivity analysis)
- Pembrolizumab (based on KN087 (cohort 2)) vs. Post BV tretaments (based on Cheah et al. 2016)
- Pembrolizumab (based on UK RWE data (SACT data)) vs. Post BV treatments (based on Cheah et al. 2016)

Diagnostic plots (Log cumulative hazard plots and Schoenfeld residual plots) will be generated to assess the proportional hazards assumptions for all aforementioned ITCs.

3 ANALYTICAL AND METHODOLOGICAL DETAILS

3.1 Set of studies

The ITCs and corresponding diagnostic plots will use the results from a systematic literature review (SLR) conducted in May 2023, that identified relevant studies. These studies have then been screened for suitable comparator studies and qualified if the studies passed a feasibility analysis. In addition, a retrospective study based on Real World Data derived from data routinely collected in the Systemic Anti-Cancer Therapy (SACT) database within all patients with an application for pembrolizumab for treating relapsed or refractory classical Hodgkin lymphoma in the Cancer Drug Fund (CDF) will also be used to provide ITCs.

Table 1 gives an overview of all comparators as well a Pembrolizumab data that have been selected to be used for ITCs.

Keytruda (MK-3475) KN087 Cohort 2
MSD-UK; ITCs of pembrolizumab vs. SoC interventions

Table 1

Summary of studies used in the analyses

Target Population	Internal Data source	External (Comparator) Data Source	Outcomes (effect measures)
• participants with relapsed or refractory (R/R) Classical Hodgkin Lymphoma (cHL) who have failed to respond or relapsed on brentuximab vedotin (BV) and are ineligible to auto-SCT (Cohort 2 of KN087)	• **Pembrolizumab:** IPD from KEYNOTE-087 Database cutoff: 15th MAR 2021	**BV from Eyre et al. 2017 study:** • Pseudo-IPD and summarized baseline characteristics **Pembrolizumab from UK RWE database study (SACT data):** • Pseudo-IPD and summarized baseline characteristics **Post BV treatments from Cheah et al. 2016 study:** • Pseudo-IPD and summarized baseline characteristics	• OS

Abbreviations *IPD, individual patient data; OS, Overall Survival; SCT, Stem cell transplantation.*

3.1.1 KN087

Protocol 087 is a multicentre, single arm, multi-cohort, nonrandomised Phase 2 trial of Pembrolizumab in subjects with refractory or relapsed classical Hodgkin lymphoma (rrcHL). Subjects meeting eligibility criteria were allocated to receive pembrolizumab 200 mg every 3 weeks (Q3W) within one of three cohorts, depending on their prior disease history and therapy:

- Cohort 1: subjects who failed to respond or progressed after auto-SCT therapy and relapsed or failed to respond after treatment with BV post auto-SCT.
- Cohort 2: subjects who were SCT-ineligible and relapsed after treatment with or failed to respond to BV.
- Cohort 3: subjects who did not respond or progressed after auto-SCT and had not received BV treatment post auto-SCT. These subjects could have received BV as part of primary treatment or salvage therapy.

3.1.2 UK RWE database (SACT data)

The underlying study, (National Disease Registration Service, 2021), is a retrospective study based on Real World Data derived from data routinely collected in the Systemic Anti-Cancer Therapy (SACT) database within all patients with an application for pembrolizumab for treating relapsed or refractory classical Hodgkin lymphoma in the Cancer Drug Fund (CDF). Patients were adults and had histologically documented classical Hodgkin lymphoma, had failed at least 2 lines of chemotherapy and treatment with brentuximab vedotin (BV). In addition, patients needed to be SCT naïve and to be ineligible for SCT. The patients had an Eastern Cooperative Oncology Group (ECOG) performance status (PS) of 0 or 1 and did not receive prior treatment with an anti-PD-1, anti-PD-L1, anti-PD-L2, anti-CD137, or anti-cytotoxic T-lymphocyte-associated antigen-4 (CTLA-4) antibody. Pembrolizumab was given as monotherapy and would commence at a fixed dose of 200mg per infusion. The patients received a maximum treatment duration with pembrolizumab of 2 years (or 35 x 3-weekly cycles of pembrolizumab or its equivalent if 6-weekly pembrolizumab monotherapy dosing was used). The corresponding UK RWE report is presented in section 6.
In total 215 patients were selected according to these criteria. The primary endpoint was Overall Survival (OS). Time to SCT and Progression-Free Survival were also reported but not considered for this report. The pembrolizumab arm has been used for the ITC as an experimental arm. The corresponding RWE report is provided in section 6.1.

Patients from the pembrolizumab arm were used to provide diagnostic plots within this report.

3.1.3 Eyre et al. 2017

The underlying study (Eyre, et al., 2017) is a retrospective, multicentre study in 99 SCT-naïve patients with relapsed or refractory cHL treated with BV monotherapy at 9 large UK centres between May 2011 and July 2016. Patients were eligible if they were transplant naïve, had

received at least 2 prior lines of treatment, and had received BV with the intention of undergoing subsequent SCT consolidation. Patients were deemed fit for transplant but had an insufficient remission to proceed. Patients were treated with BV monotherapy dosed at 1.8 mg/kg once every three weeks until SCT consolidation, progression, toxicity or death from any cause. PFS was calculated from the initiation of BV to the time of relapse, disease progression, death, or censored at the date of last follow-up. Overall survival (OS) was calculated from the initiation of BV monotherapy to date of death and censored at the date of last follow-up. Responses to treatment were analysed by either PET-CT or standard CT according to local investigator discretion.

Patients from the BV arm were used to indirectly compare pembrolizumab vs. BV using a MAIC, unadjusted ITCs as well as to provide corresponding diagnostic plots.

3.1.4 Cheah et al. 2016

The underlying study (Cheah, et al., 2016) is a retrospective observational study conducted at the MD Anderson Centre in the USA between June 2007 and January 2015 in initially 89 patients with relapsed or refractory cHL after treatment BV monotherapy. In total 79 patients who had a 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 were considered eligible for conclusion. The individual treating clinician determined post-progression treatment strategy in each case. The focus of this study was determining clinical outcomes (Progression-Free Survival (PFS) and Overall Survival (OS)) following disease relapse after BV therapy.

Data provided on patients treated after progression on BV were used to indirectly compare pembrolizumab vs. Post BV treatments using both adjusted and unadjusted ITCs as well as to provide corresponding diagnostic plots.

3.2 Analysis Populations

The All Subjects as Treated (ASaT) population is used for the analyses in KN087 which includes all participants who received at least 1 dose of pembrolizumab. This approach is consistent with CSR approach.

The population of interest for this report is Cohort 2, which includes subjects who were SCT-ineligible and relapsed after treatment with or failed to respond to BV, as defined in the protocol (details in section 3.1.1).

For the comparator studies as well as the UK RWE data (pembrolizumab arm), data were used as provided in each of the publications. This population was called modified All-Participants-as-Treated (modified APaT) population which refers to the analysis population that was used to report results in the corresponding publication. In contrast to the ASaT population for KN087, the populations that were used to report results for the comparator studies only differ in the treatment received, which needed to be a relevant comparator treatment for the UK submission and could not simply be any treatment received.

3.2.1 Efficacy Endpoints

3.2.1.1 Overall Survival

In KN-087, overall survival (OS) time is defined as the time from first dose of study treatment to death due to any cause. Participants without documented death at the time of analysis will be right censored at the date of last known alive. Participants who had a survival update after the data cutoff date are censored at the cutoff date. OS is expressed in months.

For comparator studies, data were used as reported in each of the publications.

3.2.2 Extraction of outcomes from the competing trials

For OS, Kaplan Meier curves are reported. To allow a treatment effect estimation in OS within the context of a time-to-event analysis, a digital software has been used to extract the data from the available Kaplan-Meier curves reported in the competing trials' publications. The number of patients at risk over time alongside the digitized Kaplan-Meier curve are derived and the method developed by (Guyot, Ades, Ouwens, & Welton, 2012) is applied to derive pseudo-IPD. The process of the method uses iterative numerical methods to solve inverted Kaplan-Meier equations and goal of this approach is to use all the information reported in the Kaplan-Meier curve to help identify the censoring pattern. This procedure has been undertaken by an external vendor.

3.3 Statistical Methods

3.3.1 Unadjusted ITC for time-to-event outcomes

A naïve ITC without adjustment for effect modifiers was performed based on cox proportional hazard models using both pseudo IPD of selected comparator arm as well as the KN087 (cohort 2) data, UK RWE data respectively. It is provided as a sensitivity analysis to enable a comparison against results of adjusted analyses.

The model includes treatment as a single covariate. Results for each analyzed time-to-event endpoint include hazard-ratios, corresponding 95% confidence intervals and p-value, median survival time, corresponding 95% confidence interval as well as number and percentage of events by treatment arm (Pembrolizumab vs. comparators). If there are zero events in one of the treatment groups, the two-sided Wald test will be replaced with a two-sided Score test. In addition, KM curves as well as log cumulative hazard and Schoenfeld residual plots will be provided. In addition, the approach of (Grambsch & Therneau, 1994) is used to perform a Grambsch & therneau test and to estimate the time-dependent log(hazard ratio) $\beta(t)$, which is then plotted against event time.

3.3.2 Adjusted ITC for time-to-event outcomes

Data from KN087 (pembrolizumab, cohort 2) patients were re-weighted to match the average baseline characteristics of patients retrieved from the Cheah et al. publication. An additional Matching re-weighting patients from cohort 2 of KN087 to match average baseline characteristics of patients reported in the Eyre et al. publication was also completed.

To adjust for differences in baseline characteristics between studies, individual participants from the internal studies with available IPD are re-weighted to match the mean baseline characteristics reported for the external studies with only AgD, as described in (Signorovitch, Sikirica, Haim Erder, & et al, 2012) and (Signorovitch, Wu, Yu, & et al, 2010).

Each study corresponds to a unique treatment, and a study participant can be characterized by the random quadruple (X, T, Y), where:

· X is a vector of baseline characteristics (e.g., age, sex, baseline disease severity, etc.),

· T indicates the treatment received (e.g., T = 0 for primary treatment of interest and T=1 for control),

· Y is the outcome of interest

The observed data are based on realizations of the random quadruple (x_i, t_i, y_i) i = 1, …, n, but we observe the IPD (x_i, t_i, y_i) only when $t_i = 0$. When $t_i = 1$, the IPD (x_i, t_i, y_i) are not observed individually, but the summary baseline characteristics $\bar{x_1}$ (e.g., mean or median, proportion)

from the published literature ((Signorovitch, Sikirica, Haim Erder, & et al, 2012) and (Signorovitch, Wu, Yu, & et al, 2010)).

Given these observed data, the causal effect of treatment T = 0 versus T = 1 on the outcome of Y can be estimated as:

$$\log(hazard)_{T=0}\{\text{using estimated weights}\} - \log(hazard)_{T=1} \quad (1)$$

where the weight used defined as $w_i = \Pr(T_i = 1 \mid x_i) / \Pr(T_i = 0 \mid x_i)$ is the odds that participant *i* receives treatment $T = 1$ versus $T = 0$ (i.e. enrols in trial 1 versus trial 0) given baseline characteristics x_i. Thus, participants receiving treatment $T = 0$ are re-weighted to match the distribution of participants receiving $T = 1$: participants more likely to have received $T = 1$ versus $T = 0$ will be up-weighted to compensate for their under-representation in the $T = 0$ sample; participants less likely to have received $T = 1$ versus $T = 0$ will be down-weighted to compensate for their over-representation in the $T = 0$ sample.

To apply this estimator, first the estimate w_i for each participant with $t_i = 0$ from the observed data must be obtained. As in matching methods based on propensity scores, the w_i may be assumed to follow the logistic regression model:

$$w_i = \exp(\alpha + x_i'\beta) \quad (2)$$

However, since IPD are not available for comparator study, the usual maximum likelihood approach cannot be used to estimate the parameters of the propensity score model. Instead, a method of moments can be used. To apply the method of moments to estimate as shown in equation (2) the IPD of participants receiving T=0 (e.g., primary treatment of interest arm of the internal study) is re- weighted to exactly match their mean baseline characteristics to the aggregated data available in the literature (e.g., control). β is estimated solving the following equation:

$$0 = \frac{\sum_{i:t_i=0} x_i \exp(x_i'\hat{\beta})}{\sum_{i:t_i=0} \exp(x_i'\hat{\beta})} - \overline{x_1} \quad (3)$$

The matching approach described above can be performed in a straightforward manner using optimization techniques (minimization) within SAS procedure PROC NLP (nonlinear programming).

By applying these weights, the patient characteristics of KN087 should perfectly match the aggregate data retrieved from the external publication.

In a subsequent step, the final treatment effect estimate as well as corresponding confidence intervals and p-values can be obtained using SAS procedures (e.g., PROC PHREG) with

participant weights entered through the WEIGHT option and ties were addressed by Efron's method (TIES=EFRON). Standard errors for MAIC estimates are calculated using a robust sandwich estimator. Sandwich estimators are derived empirically from the data, and account for the fact that the weights are estimated rather than fixed and known. The robust sandwich estimator can be obtained through the use of COVSANDWICH option within PROC PHREG. In addition, KM curves as well as log cumulative hazard and Schoenfeld residual plots will be provided both before and after matching.

To determine the impact of re-weighting on the available statistical information in the IPD, an effective sample size can be computed as the square of the summed weights divided by the sum of the squared weights. The maximum effective sample size occurs when all participants have equal weight. The occurrence of a small effective sample size can indicate that some participants are receiving extreme weights, and there may be little statistical power to detect differences between treatments.

The following baseline characteristics were identified and selected as potential effect modifiers based on clinical expert opinion retrieved via an expert interview before the ITCs were conducted:

1. Age
2. Sex
3. ECOG (0 vs 1)

3.3.3 Baseline Characteristics

Patient characteristics at baseline are summarised for cohort 2 of KN087, UK RWE data respectively, and the comparator study.

Characteristics that were available for KN087 (cohort2) are presented in baseline table for KN087 (cohort2) and the comparator study, where available.

Descriptive summaries are provided for the effect modifiers used for the re-weighting process. Summaries are displayed for patients before and after weighting. The effective sample size (ESS; measure to assess the impact of re-weighting) is also displayed for KN087 (cohort2).

3.4 Software

All analyses are conducted using SAS 9.4 for Linux operating system.

3.5 Limitations

As with any indirect comparison, conclusions from the analyses described above are limited by the extent to which the set of included trials meet the assumptions of the proposed methodology. Both unadjusted and adjusted comparisons are valid only if there are no (further) effect modifiers or prognostic factors in imbalance. This assumption is critical, and largely considered impossible to meet. Failure of this assumption leads to an unknown amount of bias in the estimate.

Whenever possible and available, descriptive summaries of key disease characteristics, including effect modifiers, will be presented to assess possible imbalances between treatment arms / data sources.

Small sample sizes mean that results should be treated with caution.

The process of digitizing published Kaplan-Meier curves yields pseudo-IPD that may not exactly match the original IPD of the comparator study.

In addition, the interpretation of diagnostic plots to assess the proportional hazards (PH) assumption may be subjective and may not provide a clear result of the assessment of the PH assumption. A probability to assess how likely the PH assumption may hold or not cannot be provided based on diagnostic plots. Throughout this report the term “potentially weak violation of proportional hazards assumption” is used whenever the diagnostic plots only provided little evidence for a violation. Further explanation is provided in section 4 whenever this term was applied. Whenever the PH assumption does not hold this may lead to biased results.

Differences in follow-up durations between both treatment arms may occur and lead to biased estimates in case of the proportional hazard assumption is not met.

In addition, baseline characteristics used for matching in the ITC of pembrolizumab (based on cohort 2 of KN087) vs. Post BV treatments (based on Cheah et al. 2016) correspond to the number of patients who had a progression after treatment with BV (n=89) for the Cheah et al. study, whereas the corresponding pseudo IPD are derived from patients who actually received treatment after progression on BV (n=79). Hence, the baseline characteristics may not be totally representative for the set of patients used in the efficacy analyses which may lead to biased estimates within the corresponding ITC.

Furthermore, the eligibility criteria of the studies that were used to provide the ITC of Pembrolizumab derived from cohort 2 of KN087 vs. BV based on the Eyre et al. publication data did not match completely. Although it had been clinically assessed that the bias would rather not be in favour of the experimental (pembrolizumab) arm which is the reason the ITC could be considered as a very conservative comparison, these analyses should be interpreted with extreme caution, as this aspect is only one out of many aspects to be considered to judge about the bias of the resulting ITC (see limitations mentioned above) and could still leave

some potential for residual bias. The patients in the comparator study appeared to be healthier (See for example differences in age (median age: 40 years (in KN087 cohort 2) vs. 32 (in Eyre.et al 2017)) and number of prior lines (median number of prior lines: 4 (in KN087 cohort 2) vs. median number of prior lines: 2 (in Eyre et al.2017) in Table 6) as compared to the pembrolizumab arm from the cohort 2 of KN087. In addition, differences in sex were observed (Female: 46.91% (in KN087 cohort 2) vs. Female: 54.55% (in Eyre et al. 2017)). Please also note that this assessment is limited to a few common characteristics reported both, the publication and KN087. The patient characteristics and assessment of risk of bias was performed with input from a clinical expert, prior to conducting the indirect treatment comparison analysis.

All other comparisons were feasibile in terms of comparability of eligibility criteria based on clinical expert input.

4 RESULTS

Given the analyses in this report are unadjusted or adjusted only based on known and available effect modifiers, the results should be interpreted with extreme caution.

4.1 Comparison against BV and diagnostic plots

Baseline characteristics comparing cohort 2 of KN087 vs. the Eyre et al. study are presented in Table 2. Results for OS are presented in Table 3 and Figure 1.

Baseline characteristics showing results before and after matching cohort 2 of KN087 data compared to the summarized baseline characteristics from the Eyre et al. study are presented in Table 4.

Results for OS comparing the KN087 Cohort 2 data (pembrolizumab) vs. BV derived from Eyre et al. 2017 are presented in Table 5 and Figure 2.

Corresponding diagnostic plots before matching are presented in Figure 3, Figure 5 and Figure 7.

Corresponding diagnostic plots after matching are presented in Figure 4, Figure 6 and Figure 8.

Graphical investigation based on Schoenfeld residual plots and the Log-cumulative hazard plots, both before and after matching, did not show a clear violation of proportional hazards assumption. The inspection of the log (HR) over time plot showed only little evidence for violation of the proportional hazards assumption, as the curve deviates from the horizontal line, both, before and after weighting although the p-value obtained from the Grambsch & Therneau test is larger than 0.05 which suggests there is no evidence for a violation of the PH-assumption.

Baseline characteristics comparing pembrolizumab based on the UK RWE data vs. the selected comparator study (Eyre et al.) are given in Table 6. Results for OS are presented in Table 7 and Figure 9, with associated diagnostic plots Figure 10, Figure 11 and Figure 12.

Graphical investigation based on Schoenfeld residual plots and the Log-cumulative hazard plots shows potentially weak violation of proportional hazards assumption, which aligned with what can be seen in the corresponding KM plots. A slight violation was observed based on the Schoenfeld residual plot where the Schoenfeld residuals did not clearly vary around zero (solid black horizontal line) randomly, unlike what is expected under the assumption of proportional hazards. This was indicated at any given timepoints where the CI of the estimated LOESS curve did not completely cover the value zero (solid black horizontal line). Moreover, the corresponding KM curves were crossing each other which also indicates a violation of the PH assumption. The inspection of the log (HR) over time plot showed no evidence for violation of the proportional hazards assumption, as the curve does not deviate

from the horizontal line. The p-value obtained from the Grambsch & Therneau test is larger than 0.05 also indicates no violation of the assumption.

4.2 Comparisons against Post BV Treatments and diagnostic plots

Baseline characteristics comparing cohort 2 of KN087 vs. the Cheah et al. study are presented in Table 8.

Baseline characteristics showing results before and after matching cohort 2 of KN087 data compared to the summarized baseline characteristics from the Cheah et al. study are presented in Table 9.

Results for OS are given in Table 10 and Figure 13.

Corresponding diagnostic plots before matching are presented in Figure 14, Figure 16 and Figure 18.

Corresponding diagnostic plots after matching are presented in Figure 15, Figure 17 and Figure 19.

Graphical investigation based on Schoenfeld residual plots and the Log-cumulative hazard plots, both before and after matching, did not show a clear violation of proportional hazards assumption. However, the inspection of the log (HR) over time plot showed some evidence for violation of the proportional hazards assumption, as the curve clearly deviates from the horizontal line, both, before and after weighting. For the after matching result this is further supported by a p-value smaller than 0.05 which suggests a departure from the proportionality assumption.

Baseline characteristics comparing pembrolizumab based on the UK RWE data vs. the Cheah et al. study are presented in Table 11.

Results for OS are presented in Table 12 and Figure 20 with corresponding diagnostic plots presented in Figure 21, Figure 22 and Figure 23.

Graphical investigation based on Schoenfeld residual plots and the Log-cumulative hazard plots shows potentially weak violation of proportional hazards assumption, which aligned with what can be seen in the corresponding KM plots. A slight violation was observed based on the Schoenfeld residual plot where the Schoenfeld residuals did not clearly vary around zero (solid black horizontal line) randomly, unlike what is expected under the assumption of proportional hazards. This was indicated at any given timepoints where the CI of the estimated LOESS curve did not completely cover the value zero (solid black horizontal line). Moreover, the corresponding KM curves were crossing each other which also indicates a violation of the PH assumption. The inspection of the log (HR) over time plot showed no evidence for violation of the proportional hazards assumption, as the curve does not deviate from the horizontal line. The p-value obtained from the Grambsch & Therneau test is larger than 0.05 also indicates no violation of the assumption.

5 TABLES AND FIGURES

5.1 ITC against BV

5.1.1 Pembrolizumab (KN087) vs. BV (Eyre et al. 2017)

5.1.1.1 Baseline Characteristics

Table 2
Baseline Characteristics
Cohort 2
(All-Participants-as-Treated Population)

	Pembrolizumab (N^a = 81)	Brentuximab vedotin (N^b = 99)
Sex, n (%)		
Male	43 (53.09)	45 (45.45)
Female	38 (46.91)	54 (54.55)
Age (Years), n (%)		
< 65	66 (81.48)	NR (NR)
≥ 65	15 (18.52)	NR (NR)
Mean (SD)	42.31 (17.35)	NR (NR)
Median (Q1; Q3)	40.00 (26.00; 55.00)	32 (NR; NR)
Min; Max	20.00; 76.00	13; 70
Race, n (%)		
American Indian Or Alaska Native	1 (1.23)	NR (NR)
Asian	4 (4.94)	NR (NR)
Black Or African American	2 (2.47)	NR (NR)
White	73 (90.12)	NR (NR)
Missing	1 (1.23)	NR (NR)
Ethnicity, n (%)		
Hispanic Or Latino	5 (6.17)	NR (NR)
Not Hispanic Or Latino	63 (77.78)	NR (NR)
Not Reported	9 (11.11)	NR (NR)
Unknown	4 (4.94)	NR (NR)
Geographic Region, n (%)		
US	20 (24.69)	NR (NR)
Ex-US	61 (75.31)	NR (NR)
Disease Subtype, n (%)		
Classical Hodgkin Lymphoma- Nodular Sclerosis	65 (80.25)	75 (75.76)
Classical Hodgkin Lymphoma- Mixed Cellularity	10 (12.35)	12 (12.12)
Classical Hodgkin Lymphoma- Lymphocyte Rich	1 (1.23)	1 (1.01)
Classical Hodgkin Lymphoma- Lymphocyte Depleted	4 (4.94)	1 (1.01)

Baseline Characteristics
Cohort 2
(All-Participants-as-Treated Population)

	Pembrolizumab (N^a = 81)	Brentuximab vedotin (N^b = 99)
Missing	1 (1.23)	10 (10.1)
ECOG Performance Status, n (%)		
0	44 (54.32)	45 (55.56)
1	37 (45.68)	36 (44.44)
Prior Lines of Therapy Group, n (%)		
< 3	3 (3.70)	70 (70.71)
≥ 3	78 (96.30)	29 (29.29)
Prior Lines of Therapy		
Mean (SD) Median (Q1; Q3) Min; Max	4.00 (1.63) 4.00 (3.00; 4.00) 1.00; 11.00	NR (NR) 2 (NR; NR) 2; 4
Refractory or Relapsed After 3 or More Lines, n (%)		
Yes	81 (100.00)	NR (NR)
Brentuximab Use, n (%)		
Yes	81 (100.00)	NR (NR)
Prior Radiation, n (%)		
Yes	21 (25.93)	NR (NR)
No	60 (74.07)	NR (NR)
Bulky Lymphadenopathy, n (%)		
Yes	5 (6.17)	NR (NR)
No	76 (93.83)	NR (NR)
Baseline B Symptoms, n (%)		
Yes	27 (33.33)	33 (37.5)
No	54 (66.67)	55 (62.5)
Baseline Bone Marrow Involvement, n (%)		
Yes	5 (6.17)	NR (NR)
No	75 (92.59)	NR (NR)

Baseline Characteristics
Cohort 2
(All-Participants-as-Treated Population)

	Pembrolizumab (N^a = 81)	Brentuximab vedotin (N^b = 99)
Missing	1 (1.23)	NR (NR)

a: KEYNOTE-087, Database Cutoff Date: 15MAR2021, participants of Cohort 2.
b: Number of participants: Based on Eyre et al. 2017, mAPaT population.
The number of reported characteristics within the external comparator arm may differ between characteristics and may be smaller than the total number of patients in included to the study.
CI: Confidence Interval; Q1: First Quartile; Q3: Third Quantile; Min: Minimum; Max: Maximum; SD: Standard Deviation.

5.1.1.2 Overall Survival

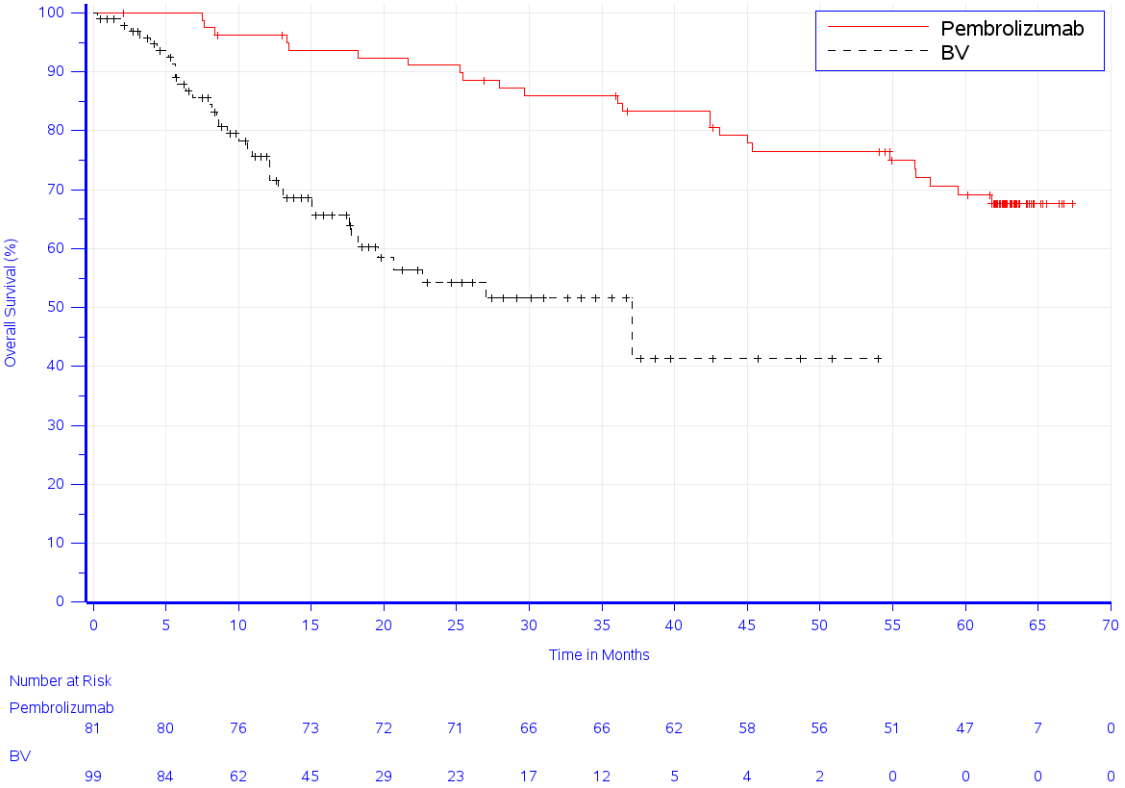
Table 3
Analysis of Overall Survival
Unadjusted Indirect Comparison Analysis of Pembrolizumab vs. Brentuximab Vedotin
(All-Participants-as-Treated Population)

	Pembrolizumab			BV			Pembrolizumab vs. BV	
	N[a]	Participants with Event n (%)	Median Time[b] in Months [95 %-CI]	N[c]	Participants with Event n (%)	Median Time[b] in Months [95 %-CI]	Hazard Ratio [95 %-CI][d]	p-Value[d,e]
Overall Survival	81	24 (29.6)	Not reached [-; -]	99	37 (37.4)	37.0 [18.2; -]	0.23 [0.12; 0.42]	< 0.001

a: KEYNOTE-087, Database Cutoff Date: 15MAR2021, participants of Cohort 2
b: From product-limit (Kaplan-Meier) method
c: Number of participants: Based on Eyre et al. 2017, mAPaT population
d: Based on Cox regression model with treatment as a single covariate
e: Two-sided p-value using Wald test (Score test in case of zero event in one treatment group)
BV: Brentuximab Vedotin; CI: Confidence Interval; mAPaT: modified APaT population

Figure 1

Kaplan-Meier Curve for Overall Survival
Comparison of Pembrolizumab vs. Brentuximab Vedotin
(All-Participants-as-Treated Population)



KEYNOTE-087, Database Cutoff Date: 15MAR2021, participants of Cohort 2
Comparator data based on Eyre et al. 2017
BV: Brentuximab Vedotin.

5.1.2 Pembrolizumab (KN087) vs. BV (Eyre et al. 2017) – Sensitivity Analysis

5.1.2.1 Baseline Characteristics

Table 4
Baseline Characteristics
Unanchored Matching-Adjusted Indirect Comparison of Pembrolizumab vs. Brentuximab Vedotin
Cohort 2
(All-Participants-as-Treated Population)

	Eyre 2017 (N[c]=99)	Study: KEYNOTE 087[a]	
		Before Matching (N[b]=81)	After Matching (N=74.63[d])
Age			
Median	32.0	40.0	31.0
ECOG (%)			
0	55.6	54.3	55.6
Sex (%)			
Female	54.5	46.9	54.5

a: Database Cutoff Date: 15MAR2021
b: KEYNOTE-087, Database Cutoff Date: 15MAR2021, participants of Cohort 2
c: Number of participants: Based on Eyre et al. 2017, mAPaT population
d: Effective sample size computed as the square of the summed weights divided by the sum of the squared weights; Weighted according to baseline characteristics presented in Eyre et al. 2017
ECOG: Eastern Cooperative Oncology Group; mAPaT: Defined as analysis populations used to report comparator study results;

Please note that a small deviation occurred when matching baseline characteristics of patients from cohort2 of KN087 to aggregated patient characteristics from Eyre et al. 2017. The median age does not exactly match, as cohort 2 of KN087 did not contain any patients with the median age presented for Eyre et al. 2017. However, the derived binary age matching variable (dichotomized at age threshold of 32 years corresponding to median age presented in Eyre et al. 2017) matched to 0.5 after matching. Hence, the matching procedure actually delivered an exact match on the median age.

5.1.2.2 Overall Survival

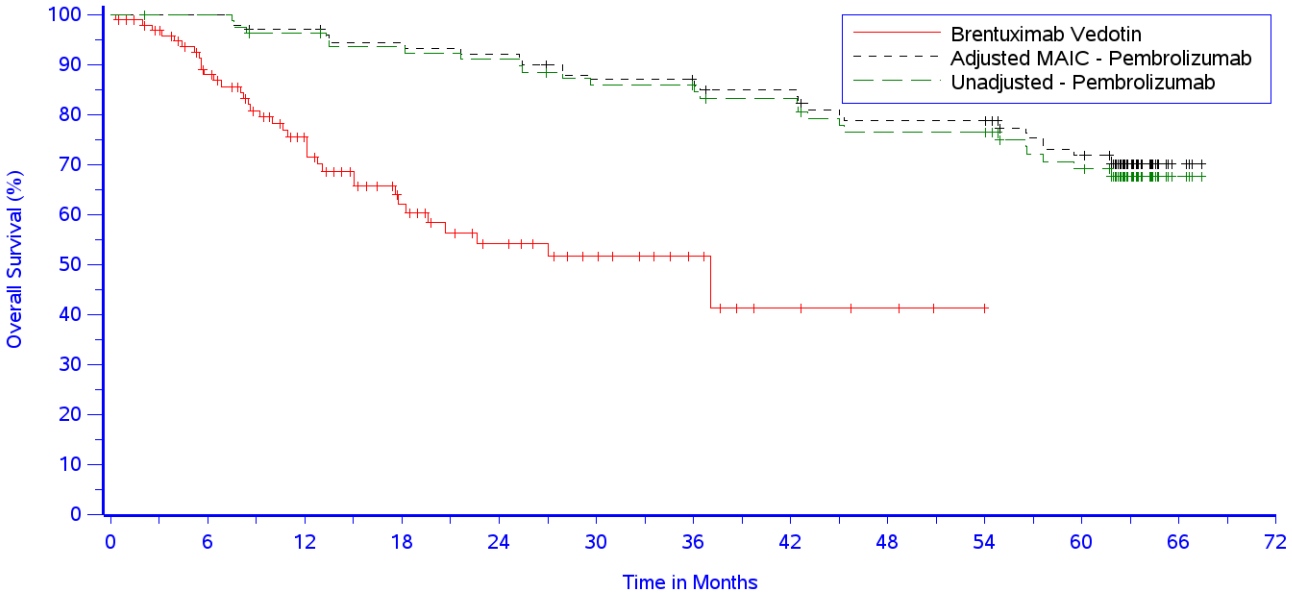
Table 5
Analysis of Overall Survival
Unanchored Matching-Adjusted Indirect Comparison of Pembrolizumab vs. Brentuximab Vedotin
Cohort 2
(All-Participants-as-Treated Population)

Study: KEYNOTE 087[a]	Pembrolizumab			Brentuximab Vedotin			Pembrolizumab vs. Brentuximab Vedotin	
	N[b]	Participants with Event, n (%)	Median Time[c] in Months [95%-CI]	N[d]	Participants with Event, n (%)	Median Time[c] in Months [95%-CI]	Hazard Ratio [95%-CI][e]	p-Value[e,f]
Before Matching	81	24 (29.6)	Not reached [-; -]	99	37 (37.4)	37.0 [18.2; -]	0.23 (0.12, 0.42)	< 0.001
After Matching[g]	77.7[h]	21 (27.0)	Not reached [-; -]	99	37 (37.4)	37.0 [18.2; -]	0.21 (0.12, 0.37)	< 0.001

a: Database Cutoff Date: 15MAR2021
b: KEYNOTE-087, Database Cutoff Date: 15MAR2021, participants of Cohort 2
c: From product-limit (Kaplan-Meier) method
d: Number of participants: Based on Eyre et al. 2017, mAPaT population
e: Based on Cox regression model with treatment as a single covariate
f: Two-sided p-value using Wald test (Score test in case of zero event in one treatment group)
g: Matching to baseline characteristics of comparator study was conducted using: Age (Median), Sex and ECOG
h: Effective sample size computed as sum of weights
CI: Confidence Interval; ECOG: Eastern Cooperative Oncology Group; mAPaT: Defined as analysis populations used to report comparator study results;

Figure 2

Kaplan-Meier Curve for Overall Survival
Unanchored Matching-Adjusted Indirect Comparison of Pembrolizumab vs. Brentuximab Vedotin
Cohort 2
(All-Participants-as-Treated Population)



Number at Risk

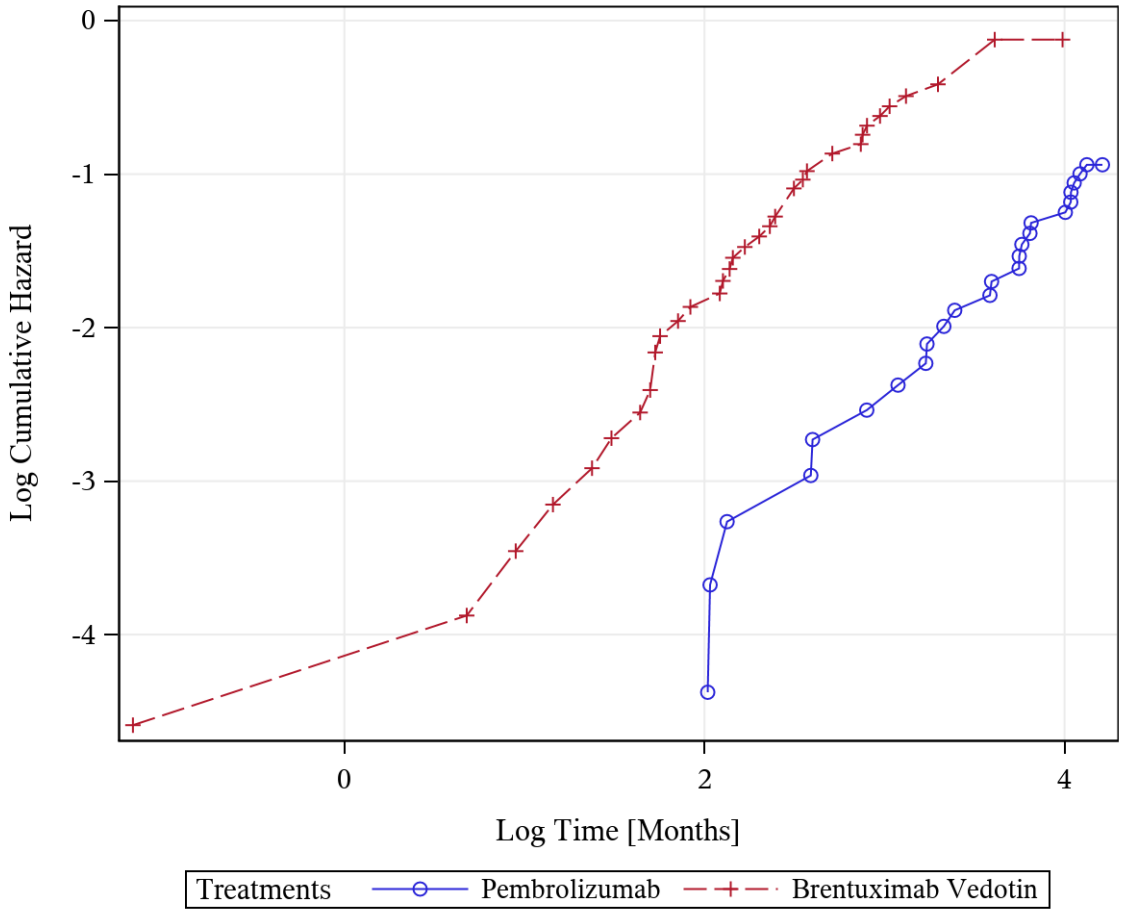
Brentuximab Vedotin	99	77	55	35	24	17	11	5	3	0	0	0	0
Adjusted MAIC - Pembrolizumab	77.7	76.7	73.8	70.8	68.9	64.3	62.9	59.8	54.5	54.5	45.8	3.5	0.0
Unadjusted - Pembrolizumab	81	80	76	73	71	66	65	62	56	56	47	4	0

KEYNOTE-087, Database Cutoff Date: 15MAR2021, participants of Cohort 2
Comparator data based on Eyre et al. 2017

5.1.2.3 Diagnostics

Figure 3

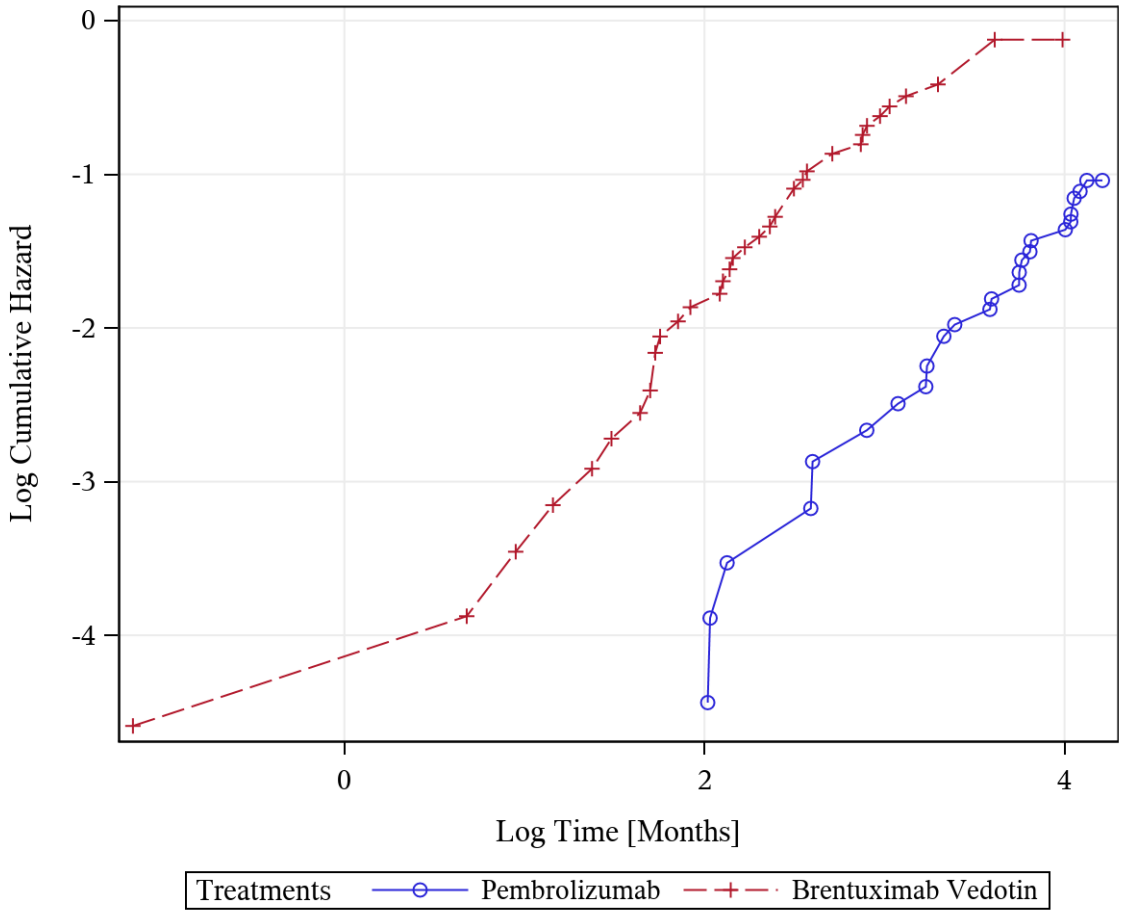
Diagnostic Plot
Log Cumulative Hazard vs. Log Time for Overall Survival
(Before Matching)
Cohort 2
(All-Participants-as-Treated Population)



KEYNOTE-087, Database Cutoff Date: 15MAR2021, participants of Cohort 2
Comparator data based on Eyre et al. 2017

Figure 4

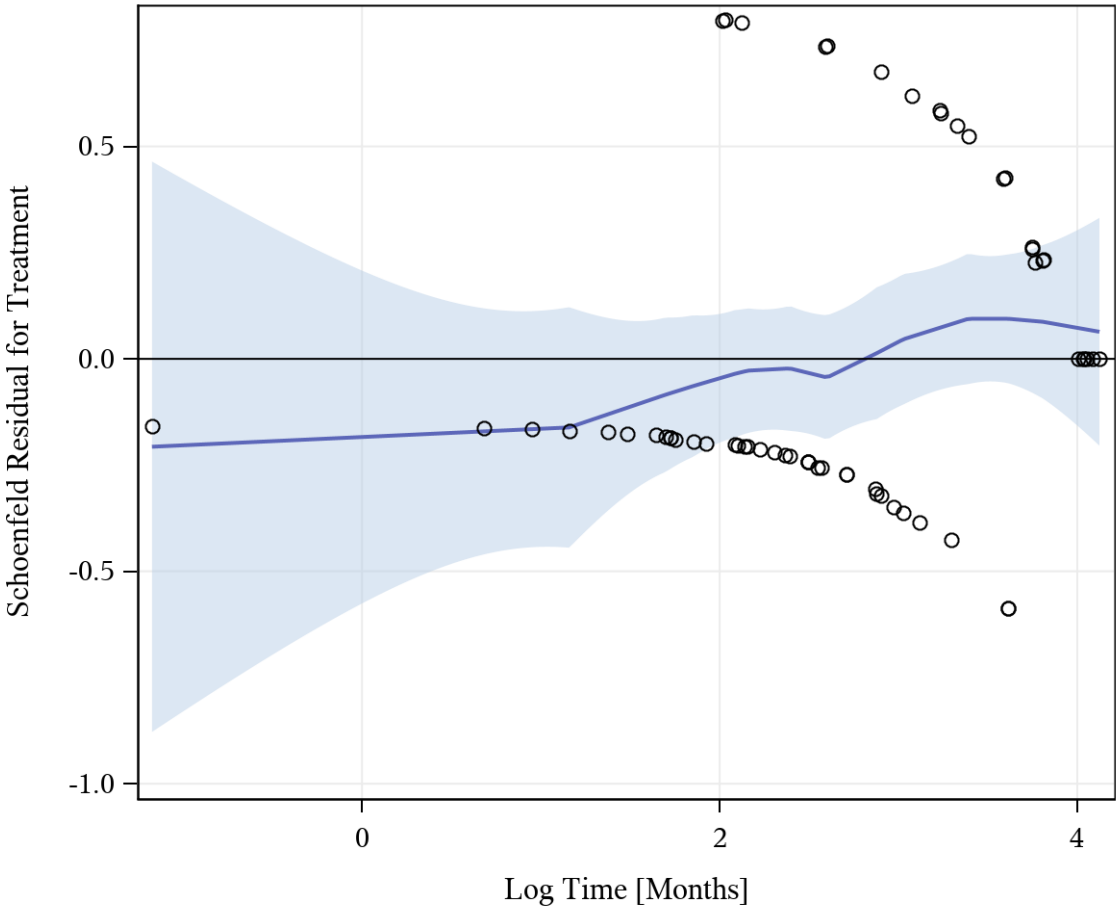
Diagnostic Plot
Log Cumulative Hazard vs. Log Time for Overall Survival
(After Matching)
Cohort 2
(All-Participants-as-Treated Population)



KEYNOTE-087, Database Cutoff Date: 15MAR2021, participants of Cohort 2
Comparator data based on Eyre et al. 2017

Figure 5

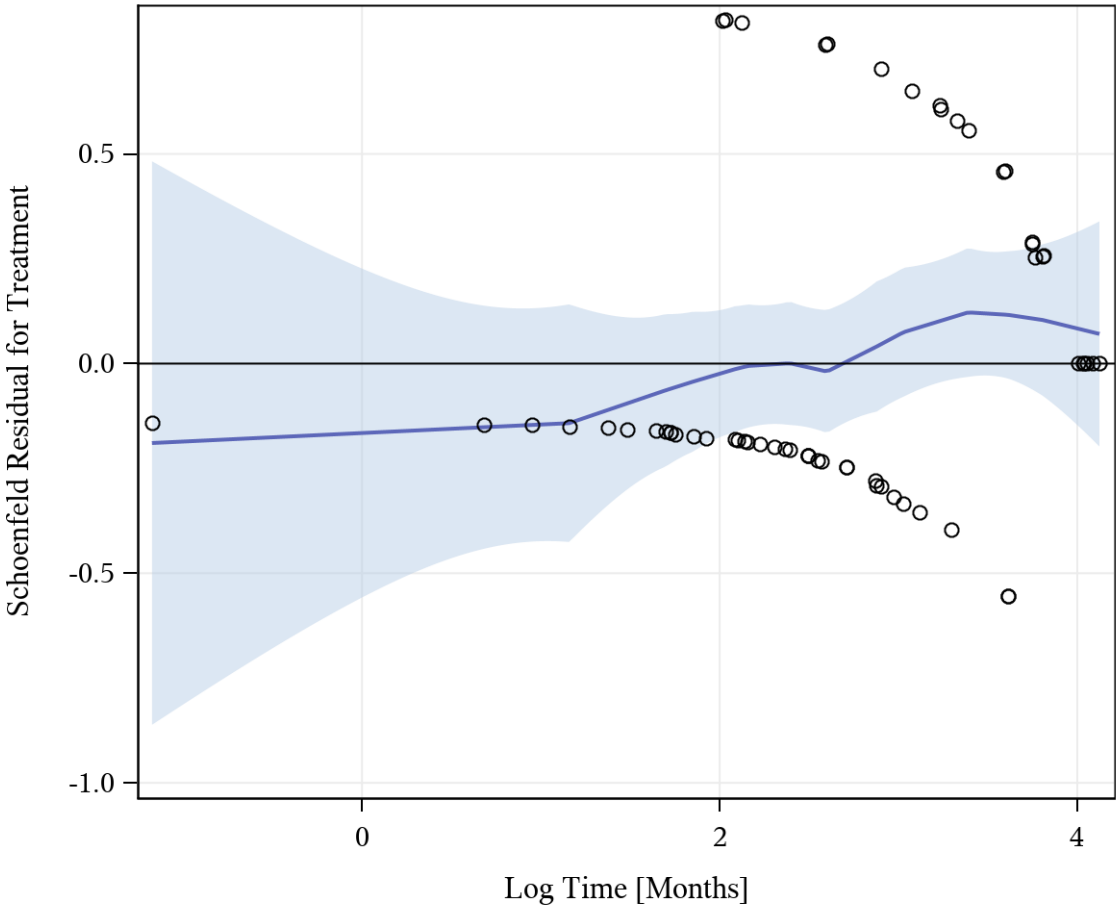
Diagnostic Plot
Schoenfeld residuals for treatment vs. Log Time for Overall Survival
(Before Matching)
Cohort 2
(All-Participants-as-Treated Population)



KEYNOTE-087, Database Cutoff Date: 15MAR2021, participants of Cohort 2
Comparator data based on Eyre et al. 2017
solid line represents LOESS (Locally Estimated Scatterplot Smoothing) curve. Area around the solid line shows 95% CI.

Figure 6

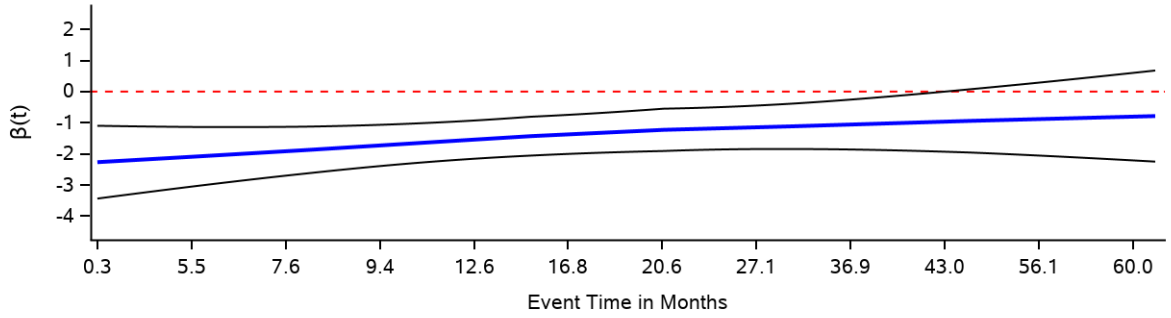
Diagnostic Plot
Schoenfeld residuals for treatment vs. Log Time for Overall Survival
(After Matching)
Cohort 2
(All-Participants-as-Treated Population)



KEYNOTE-087, Database Cutoff Date: 15MAR2021, participants of Cohort 2
Comparator data based on Eyre et al. 2017
solid line represents LOESS (Locally Estimated Scatterplot Smoothing) curve. Area around the solid line shows 95% CI.

Figure 7

Log Hazard Ratio β(t) Over Time
(Before Matching)
Cohort 2
(All-Participants-as-Treated Population)



ρ = 0.189

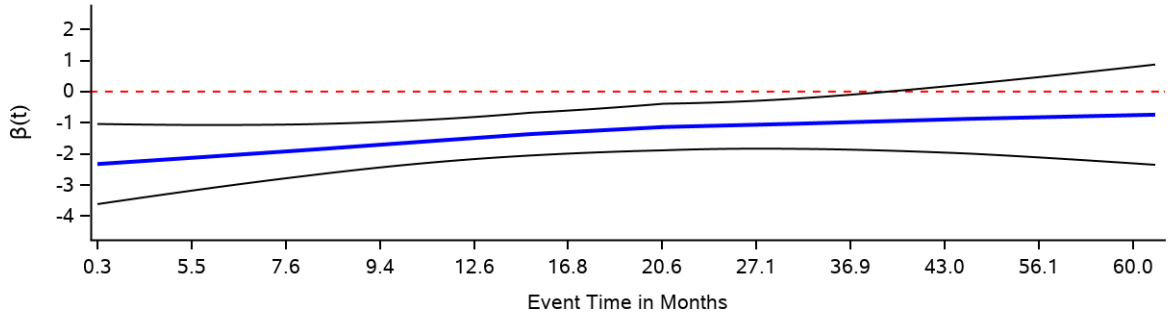
χ^2 = 1.813

p-value = 0.178

Database Cutoff Date: 15MAR2021
KEYNOTE-087, Database Cutoff Date: 15MAR2021, participants of Cohort 2
Comparator data based on Eyre et al. 2017
Estimation of hazard ratios over time is based on scaled Schoenfeld residuals

Figure 8

Log Hazard Ratio β(t) Over Time
(After Matching)
Cohort 2
(All-Participants-as-Treated Population)



ρ = 0.185

χ^2 = 1.912

p-value = 0.167

Database Cutoff Date: 15MAR2021
KEYNOTE-087, Database Cutoff Date: 15MAR2021, participants of Cohort 2
Comparator data based on Eyre et al. 2017
Estimation of hazard ratios over time is based on scaled Schoenfeld residuals

5.1.3 Pembrolizumab (SACT) vs. BV (Eyre et al. 2017)

5.1.3.1 Baseline Characteristics

Table 6
Relevant patient characteristics SACT (UK RWE data) and Eyre et al., 2017

Characteristics	Variable as measured in studies		SACT (n=215)	Eyre et al., 2017 (n=99)
Age	Age, n/N (%)	≥ 65 years		--
		< 65 years		--
		≥ 60 years	81 (38)	--
		< 60 years	134 (62)	--
	Age, median (range)		47	32 (13-70)
Sex	Male, n (%)	Yes	130 (60)	45/99 (45)
Disease status	Disease status, n/N (%)	Refractory	--	--
		Relapsed < 12 months	--	--
		Relapsed ≥ 12 months	--	--
Number of prior lines of therapy	Number of prior lines, n/N (%)	2	--	70/99 (71)
		3	--	24/99 (24)
		4		5/99 (5)
	Number of prior lines, median (range)		--	2 (2-4)
Prior auto-SCT	Prior auto-SCT, n (%)	Yes		0 (0)
Prior treatment	Radiation therapy, n (%)	Yes		(7-14)
Presence of B symptoms	Presence of B symptoms, n (%)	Yes		33/88 (38)
Performance status	ECOG, n/N (%)	0	59 (27)	45/86 (52)

Characteristics	Variable as measured in studies		SACT (n=215)	Eyre et al., 2017
		>=1	104 (49)	41/86 (48)
		Missing	52(24)	
Presence of bulky disease	Bulky disease, n (%)	Bulky disease		20/95 (21)

Summary statistics are reported as described in both Real World Evidence SACT report (National Disease Registration Service, 2021) as well as publication of relevant comparator study (Eyre, et al., 2017).

5.1.3.2 Overall Survival

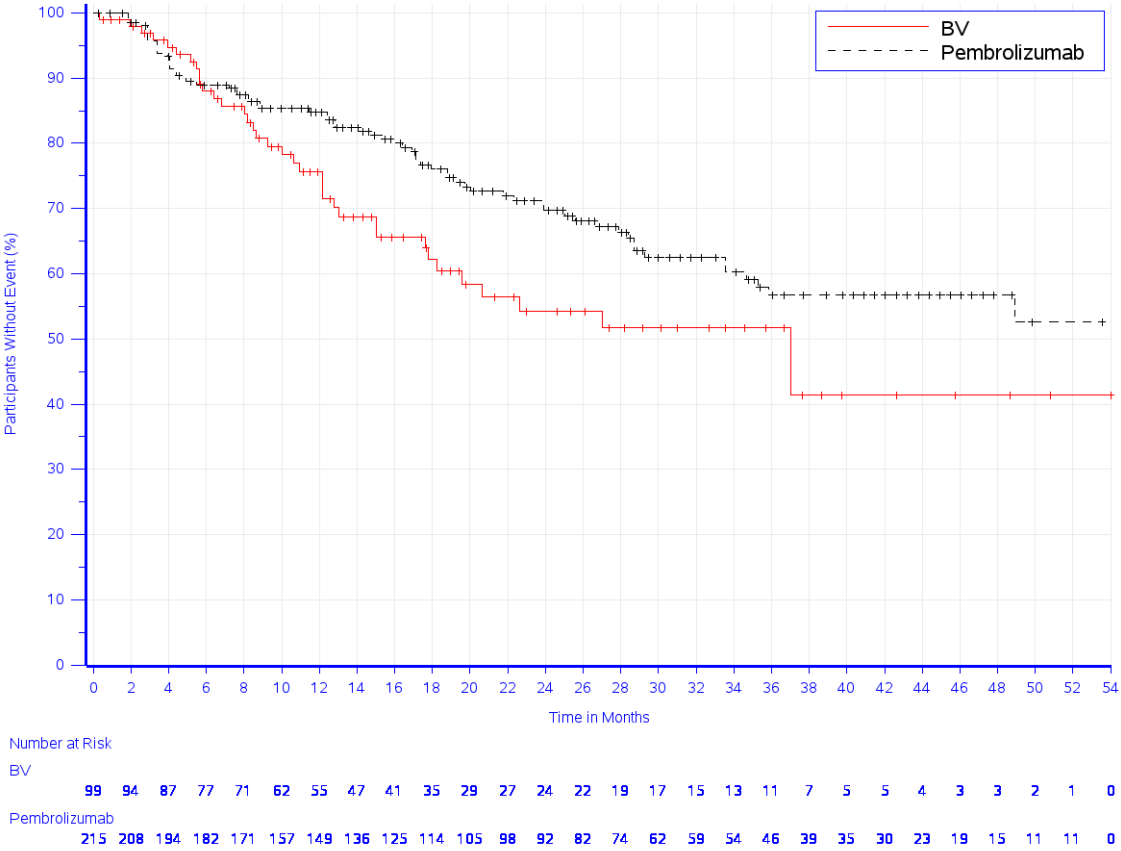
Table 7
Analysis of Overall Survival
Unadjusted Indirect Comparison Analysis of Pembrolizumab vs. Brentuximab Vedotin
(modified All-Participants-as-Treated Population)

	Pembrolizumab			BV			Pembrolizumab vs. BV	
	N[a]	Participants with Event n (%)	Median Time[b] in Months [95 %-CI]	N[c]	Participants with Event n (%)	Median Time[b] in Months [95 %-CI]	Hazard Ratio [95 %-CI][d]	p-Value[d,e]
Overall survival	215	68 (31.6)	Not reached [35.27; -]	99	37 (37.4)	37.05 [18.23; -]	0.66 [0.44; 0.98]	0.040

a: Number of participants: Based on Real World Evidence source (SACT database) in participants relapsing after treatment with or failing to respond to BV and ineligible for SCT, mAPaT
b: From product-limit (Kaplan-Meier) method
c: Number of participants: Based on Eyre et al. 2017, mAPaT population
d: Based on Cox regression model with treatment as a single covariate
e: Two-sided p-value using Wald test (Score test in case of zero event in one treatment group)
BV: Brentuximab Vedotin; CI: Confidence Interval; mAPaT: modified APaT population; SCT: Stem Cell Transplantation

Figure 9

Kaplan-Meier Curves of Overall Survival
Comparison of Pembrolizumab vs. Brentuximab Vedotin
(modified All-Participants-as-Treated Population)

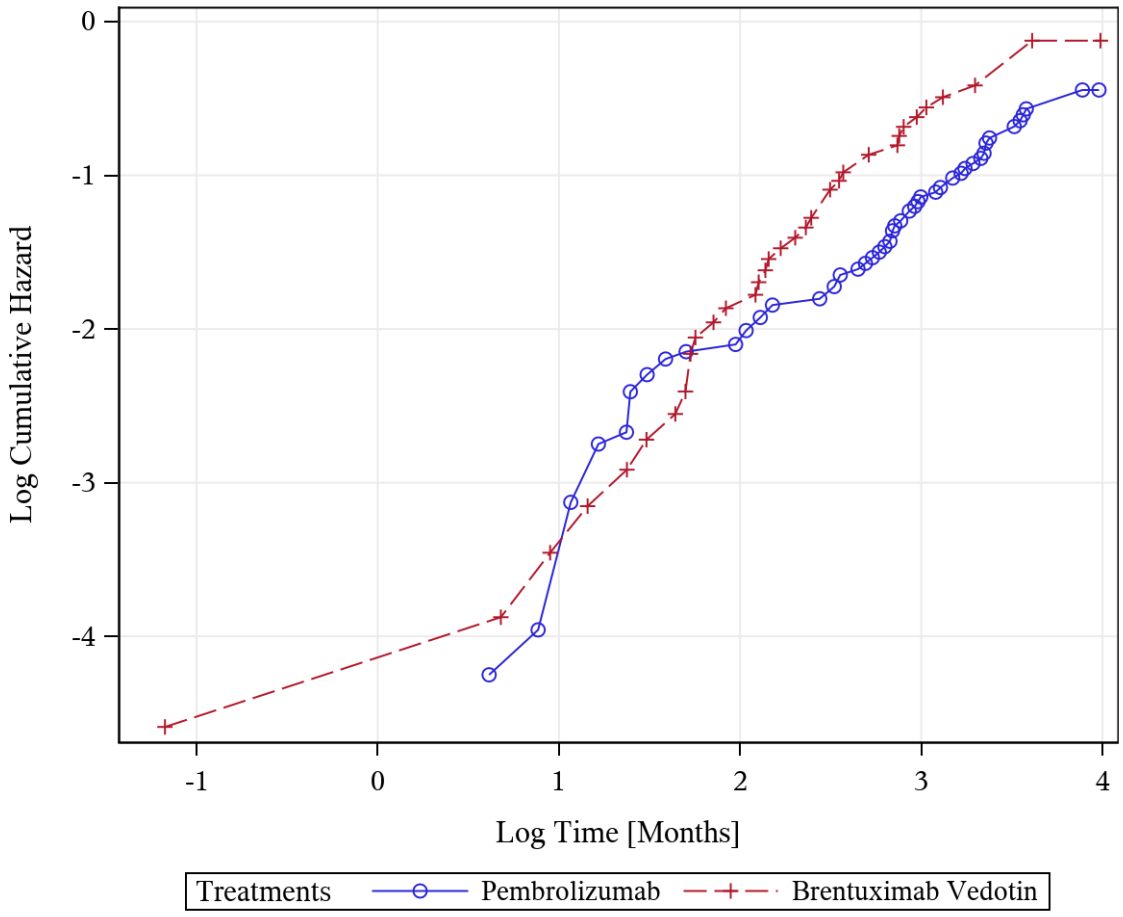


Pembrolizumab arm based on Real World Evidence source (SACT database) in participants relapsing after treatment with or failing to respond to BV and ineligible for SCT
Comparator data based on Eyre et al. 2017
BV: Brentuximab Vedotin; SCT: Stem Cell Transplantation;

5.1.3.3 Diagnostics

Figure 10

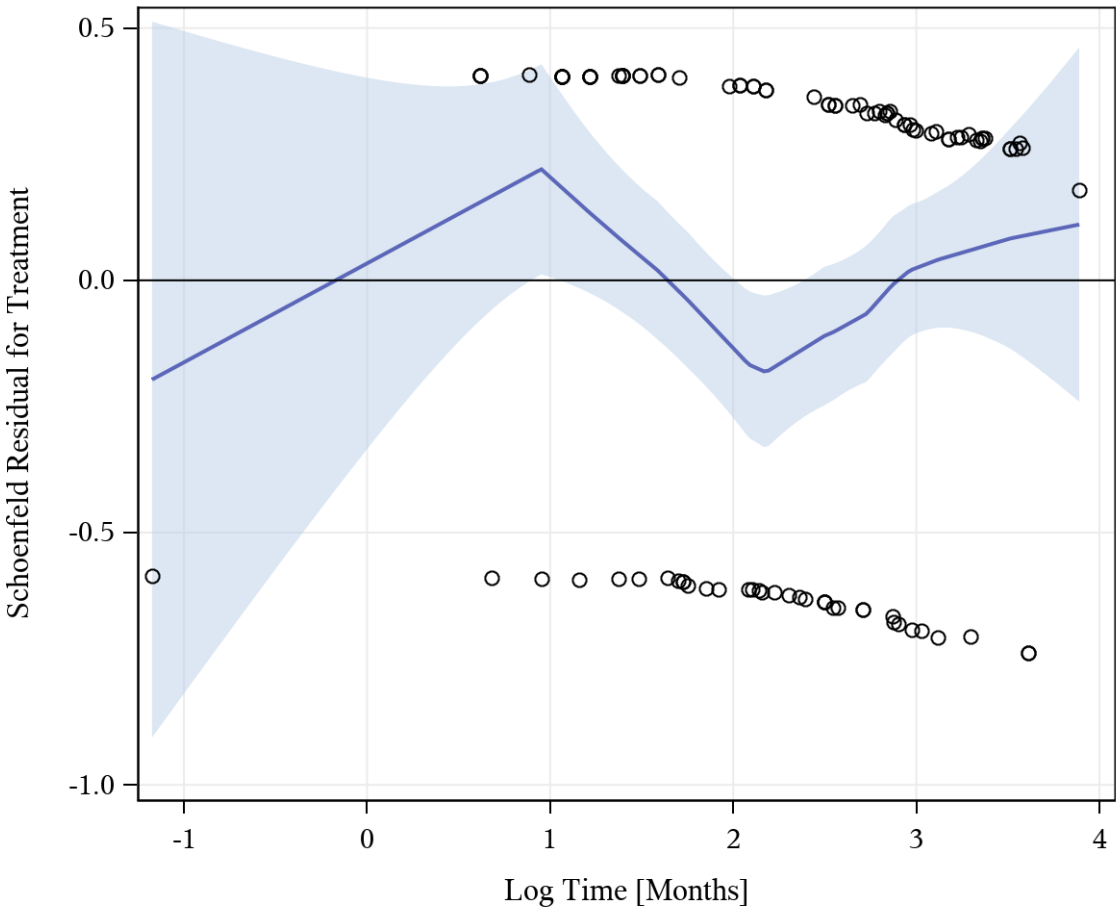
Diagnostic Plot
Log Cumulative Hazard vs. Log Time for Overall Survival
(modified All-Participants-as-Treated Population)



Pembrolizumab arm based on Real World Evidence source
(SACT database) in participants relapsing after treatment with or failing to respond to BV and ineligible for SCT
Comparator data based on Eyre et al. 2017
BV: Brentuximab Vedotin; SCT: Stem Cell Transplantation

Figure 11

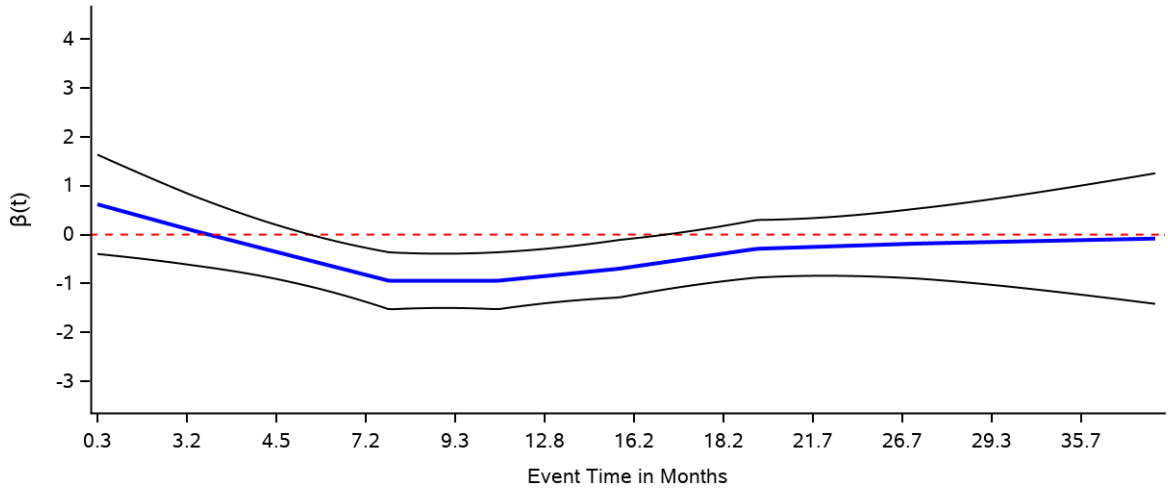
Diagnostic Plot
Schoenfeld residuals for treatment vs. Log Time for Overall Survival
(modified All-Participants-as-Treated Population)



Pembrolizumab arm based on Real World Evidence source
(SACT database) in participants relapsing after treatment with or failing to respond to BV and ineligible for SCT
Comparator data based on Eyre et al. 2017
solid line represents LOESS (Locally Estimated Scatterplot Smoothing) curve. Area around the solid line shows 95% CI.
BV: Brentuximab Vedotin ; SCT: Stem Cell Transplantation

Figure 12

Log Hazard Ratio β(t) Over Time
(modified All-Participants-as-Treated Population)



ρ = -.018

χ^2 = 0.035

p-value = 0.852

Database Cutoff Date: 15MAR2021
Pembrolizumab arm based on Real World Evidence source (SACT database) in participants relapsing after treatment with or failing to respond to BV and ineligible for SCT
Estimation of hazard ratios over time is based on scaled Schoenfeld residuals. Comparator data based on Eyre et al. 2017

5.2 ITC against Post BV Treatments

5.2.1 Pembrolizumab (KN087) vs. Post BV (Cheah et al. 2016)

5.2.1.1 Baseline Characteristics

Table 8
Baseline Characteristics
Cohort 2
(All-Participants-as-Treated Population)

	Pembrolizumab (N^a = 81)	Post BV Treatments (N^b = 89)
Sex, n (%)		
Male	43 (53.09)	51 (52.58)
Female	38 (46.91)	46 (47.42)
Age (Years), n (%)		
< 65	66 (81.48)	NR (NR)
≥ 65	15 (18.52)	NR (NR)
Mean (SD)	42.31 (17.35)	NR (NR)
Median (Q1; Q3)	40.00 (26.00; 55.00)	32 (NR; NR)
Min; Max	20.00; 76.00	18; 84
Race, n (%)		
American Indian Or Alaska Native	1 (1.23)	NR (NR)
Asian	4 (4.94)	NR (NR)
Black Or African American	2 (2.47)	NR (NR)
White	73 (90.12)	NR (NR)
Missing	1 (1.23)	NR (NR)
Ethnicity, n (%)		
Hispanic Or Latino	5 (6.17)	NR (NR)
Not Hispanic Or Latino	63 (77.78)	NR (NR)
Not Reported	9 (11.11)	NR (NR)
Unknown	4 (4.94)	NR (NR)
Geographic Region, n (%)		
US	20 (24.69)	NR (NR)
Ex-US	61 (75.31)	NR (NR)
Disease Subtype, n (%)		
Classical Hodgkin Lymphoma- Nodular Sclerosis	65 (80.25)	NR (NR)
Classical Hodgkin Lymphoma- Mixed Cellularity	10 (12.35)	NR (NR)
Classical Hodgkin Lymphoma- Lymphocyte Rich	1 (1.23)	NR (NR)
Classical Hodgkin Lymphoma- Lymphocyte Depleted	4 (4.94)	NR (NR)

Baseline Characteristics
Cohort 2
(All-Participants-as-Treated Population)

	Pembrolizumab (N^a = 81)	Post BV Treatments (N^b = 89)
Missing	1 (1.23)	NR (NR)
ECOG Performance Status, n (%)		
0	44 (54.32)	33 (40.74)
≥ 1	37 (45.68)	48 (59.26)
Prior Lines of Therapy Group, n (%)		
< 3	3 (3.70)	NR (NR)
≥ 3	78 (96.30)	NR (NR)
Prior Lines of Therapy		
Mean (SD)	4.00 (1.63)	NR (NR)
Median (Q1; Q3)	4.00 (3.00; 4.00)	NR (NR; NR)
Min; Max	1.00; 11.00	NR; NR
Refractory or Relapsed After 3 or More Lines, n (%)		
Yes	81 (100.00)	NR (NR)
Brentuximab Use, n (%)		
Yes	81 (100.00)	NR (NR)
Prior Radiation, n (%)		
Yes	21 (25.93)	NR (NR)
No	60 (74.07)	NR (NR)
Bulky Lymphadenopathy, n (%)		
Yes	5 (6.17)	NR (NR)
No	76 (93.83)	NR (NR)
Baseline B Symptoms, n (%)		
Yes	27 (33.33)	7 (8.14)
No	54 (66.67)	79 (91.86)
Baseline Bone Marrow Involvement, n (%)		
Yes	5 (6.17)	NR (NR)
No	75 (92.59)	NR (NR)

Baseline Characteristics
Cohort 2
(All-Participants-as-Treated Population)

	Pembrolizumab (N^a = 81)	Post BV Treatments (N^b = 89)
Missing	1 (1.23)	NR (NR)

a: KEYNOTE-087, Database Cutoff Date: 15MAR2021, participants of Cohort 2.
b: Number of participants: Cheah et al. 2016, mAPaT population.
The number of reported characteristics within the external comparator arm may differ between characteristics and may be smaller than the total number of patients in included to the study.
BV: Brentuximab Vedotin; CI: Confidence Interval; Q1: First Quartile; Q3: Third Quantile; Min: Minimum; Max: Maximum; SD: Standard Deviation.

Please note that the total number of patients reported for Post BV treatments (n=89) corresponds to the number of patients who had a progression after treatment with BV. This set of patients is different from the number of patients used for the corresponding efficacy analyses (n=79) which refers to the number of patients with progression after treatment with BV and also treated after progression on BV.

Table 9
Baseline Characteristics
Unanchored Matching-Adjusted Indirect Comparison of Pembrolizumab vs. Post BV Treatments
Cohort 2
(All-Participants-as-Treated Population)

	Cheah 2016 (N[c]=89)	Study: KEYNOTE 087[a]	
		Before Matching (N[b]=81)	After Matching (N=70.28[d])
Age			
Median	32.0	40.0	32.0
ECOG (%)			
0	41.0	54.3	41.0
Sex (%)			
Female	47.0	46.9	47.0

a: Database Cutoff Date: 15MAR2021
b: KEYNOTE-087, Database Cutoff Date: 15MAR2021, participants of Cohort 2
c: Number of participants: Based on Cheah et al. 2016, mAPaT population
d: Effective sample size computed as the square of the summed weights divided by the sum of the squared weights; Weighted according to baseline characteristics presented in Cheah et al. 2016
ECOG: Eastern Cooperative Oncology Group; mAPaT: Defined as analysis populations used to report comparator study results;

Please note that the total number of patients reported for Post BV treatments (n=89) corresponds to the number of patients who had a progression after treatment with BV. This set of patients is different from the number of patients used for the corresponding efficacy analyses (n=79) which refers to the number of patients with progression after treatment with BV and also treated after progression on BV.

5.2.1.2 Overall Survival

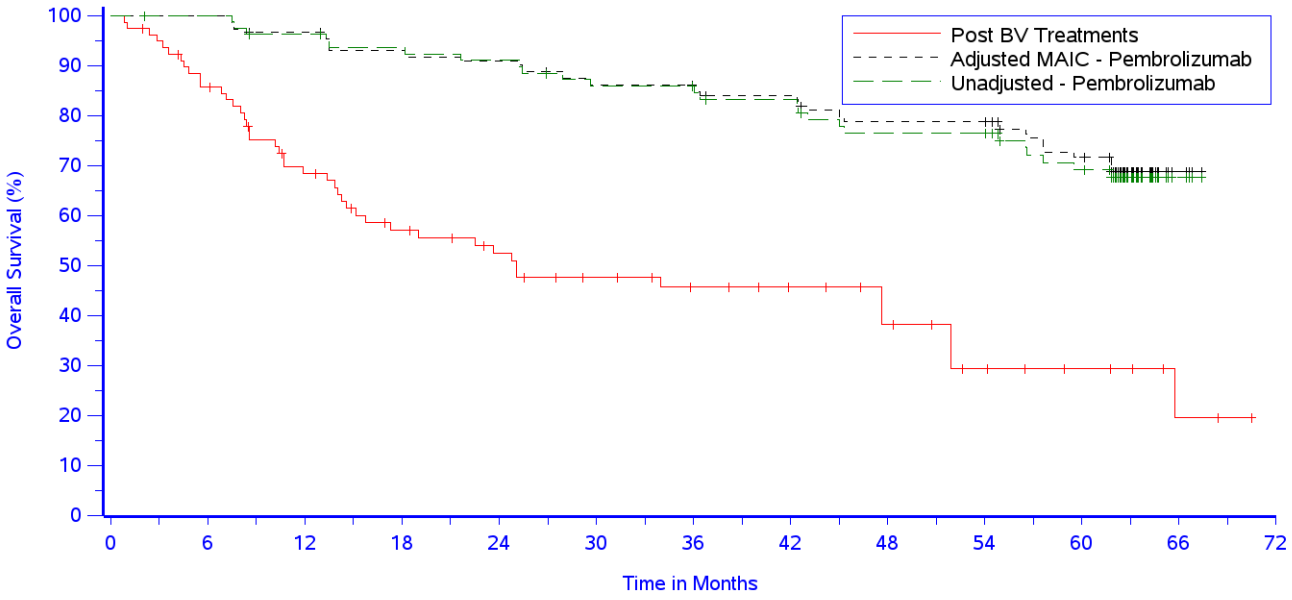
Table 10
Analysis of Overall Survival
Unanchored Matching-Adjusted Indirect Comparison of Pembrolizumab vs. Post BV Treatments
Cohort 2
(All-Participants-as-Treated Population)

Study: KEYNOTE 087[a]	Pembrolizumab			Post BV Treatments			Pembrolizumab vs. Post BV Treatments	
	N[b]	Participants with Event, n (%)	Median Time[c] in Months [95%-CI]	N[d]	Participants with Event, n (%)	Median Time[c] in Months [95%-CI]	Hazard Ratio [95%-CI][e]	p-Value[e,f]
Before Matching	81	24 (29.6)	Not reached [-; -]	79	46 (58.2)	25.1 [14.6; 51.9]	0.25 (0.15, 0.42)	< 0.001
After Matching[g]	75.4[h]	21 (27.9)	Not reached [-; -]	79	46 (58.2)	25.1 [14.6; 51.9]	0.24 (0.14, 0.40)	< 0.001

a: Database Cutoff Date: 15MAR2021
b: KEYNOTE-087, Database Cutoff Date: 15MAR2021, participants of Cohort 2
c: From product-limit (Kaplan-Meier) method
d: Number of participants: Based on Cheah et al. 2016, mAPaT population
e: Based on Cox regression model with treatment as a single covariate
f: Two-sided p-value using Wald test (Score test in case of zero event in one treatment group)
g: Matching to baseline characteristics of comparator study was conducted using: Age (Median), Sex and ECOG
h: Effective sample size computed as sum of weights
BV: Brentuximab Vedotin; CI: Confidence Interval; ECOG: Eastern Cooperative Oncology Group; mAPaT: Defined as analysis populations used to report comparator study results;

Figure 13

Kaplan-Meier Curve for Overall Survival
Unanchored Matching-Adjusted Indirect Comparison of Pembrolizumab vs. Post BV Treatments
Cohort 2
(All-Participants-as-Treated Population)



Number at Risk

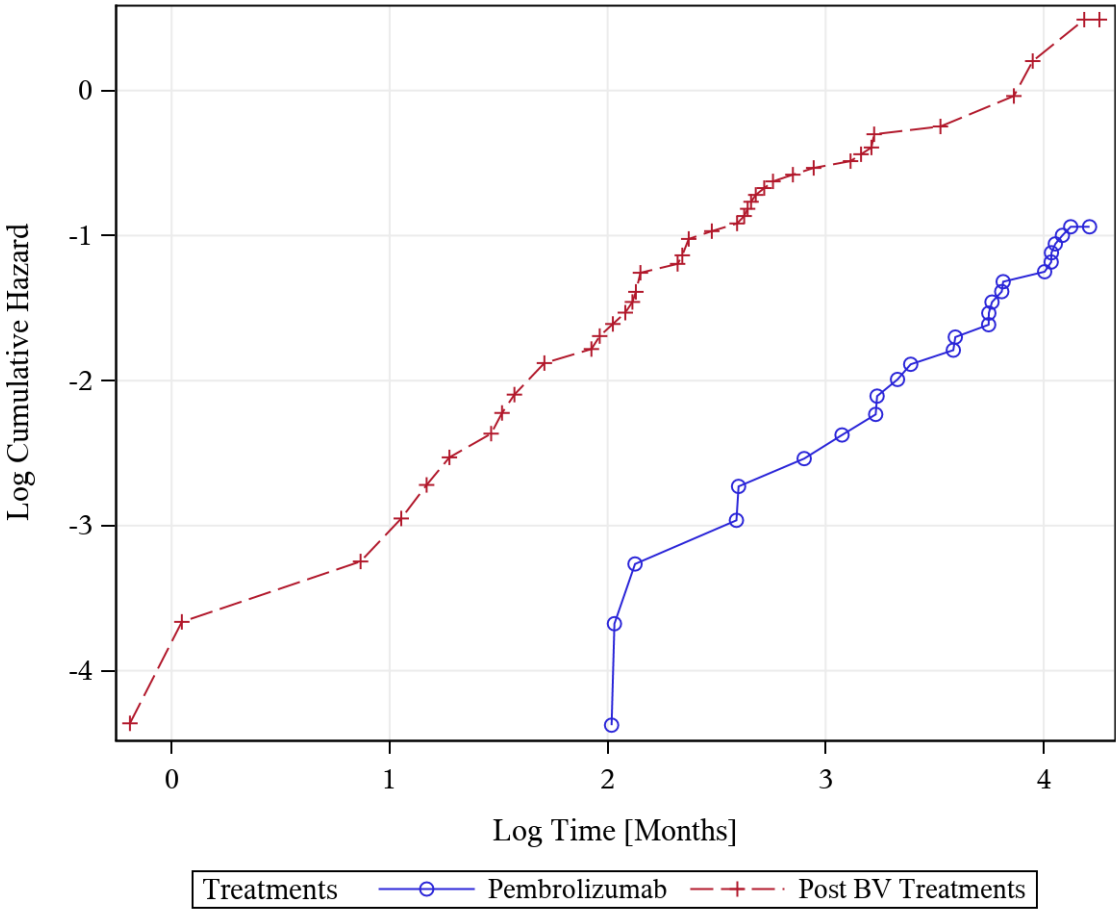
Post BV Treatments	79	66	50	39	33	27	23	20	15	9	6	2	0
Adjusted MAIC - Pembrolizumab	75.4	74.4	71.0	67.4	65.9	62.0	60.3	57.9	52.7	52.7	43.1	3.8	0.0
Unadjusted - Pembrolizumab	81	80	76	73	71	66	65	62	56	56	47	4	0

KEYNOTE-087, Database Cutoff Date: 15MAR2021, participants of Cohort 2
Comparator data based on Cheah et al. 2016

5.2.1.3 Diagnostics

Figure 14

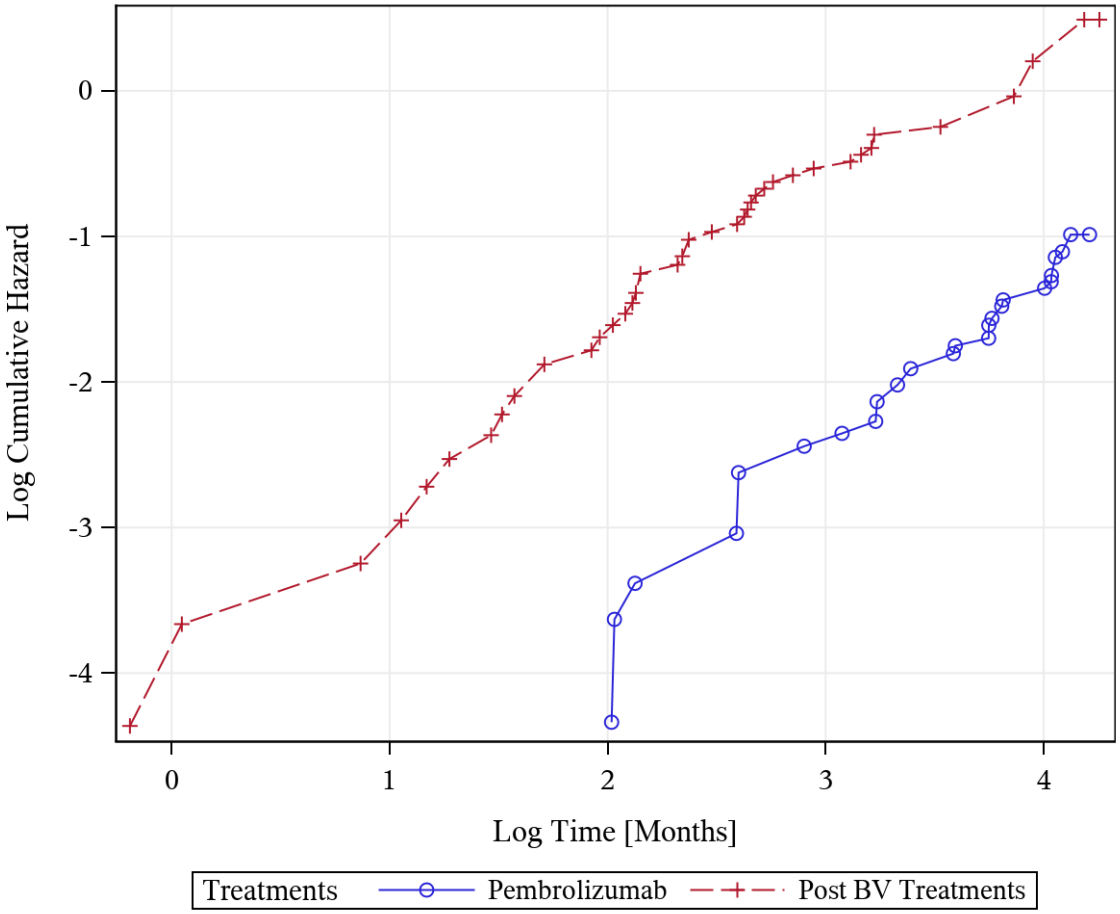
Diagnostic Plot
Log Cumulative Hazard vs. Log Time for Overall Survival
(Before Matching)
Cohort 2
(All-Participants-as-Treated Population)



KEYNOTE-087, Database Cutoff Date: 15MAR2021, participants of Cohort 2
Comparator data based on Cheah et al. 2016
BV: Brentuximab Vedotin;

Figure 15

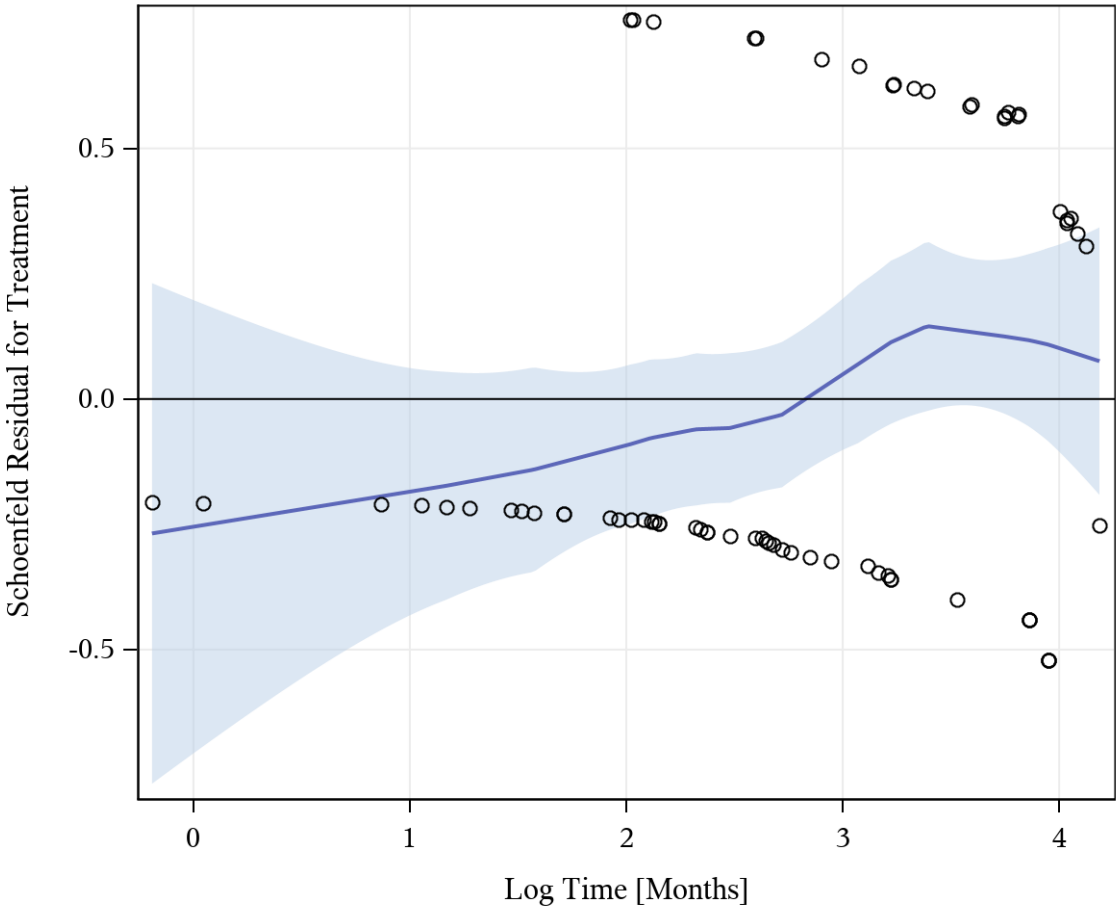
Diagnostic Plot
Log Cumulative Hazard vs. Log Time for Overall Survival
(After Matching)
Cohort 2
(All-Participants-as-Treated Population)



KEYNOTE-087, Database Cutoff Date: 15MAR2021, participants of Cohort 2
Comparator data based on Cheah et al. 2016
BV: Brentuximab Vedotin;

Figure 16

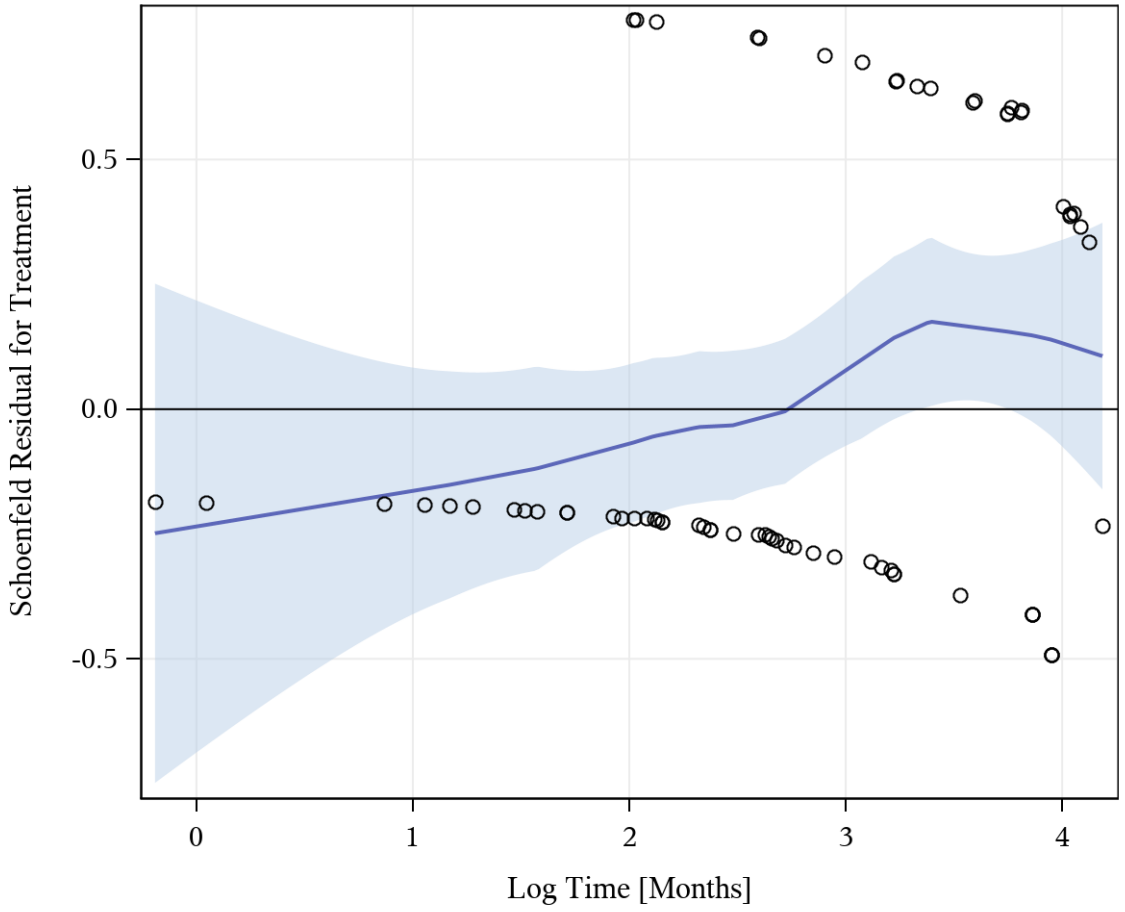
Diagnostic Plot
Schoenfeld residuals for treatment vs. Log Time for Overall Survival
(Before Matching)
Cohort 2
(All-Participants-as-Treated Population)



KEYNOTE-087, Database Cutoff Date: 15MAR2021, participants of Cohort 2
Comparator data based on Cheah et al. 2016
solid line represents LOESS (Locally Estimated Scatterplot Smoothing) curve. Area around the solid line shows 95% CI.

Figure 17

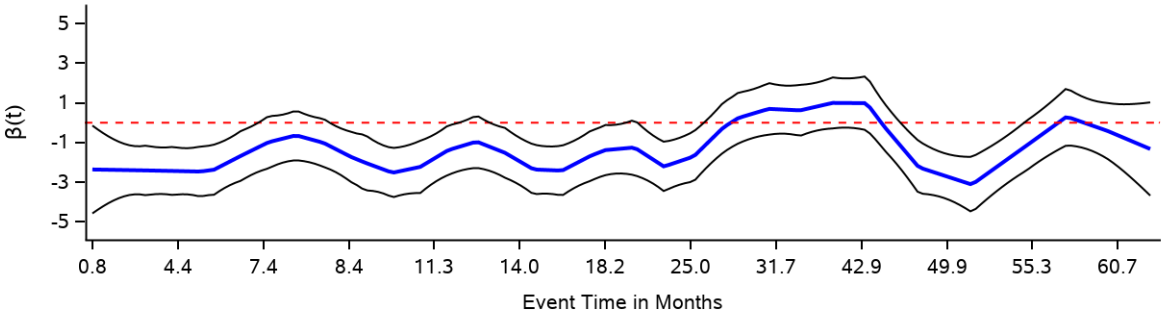
Diagnostic Plot
Schoenfeld residuals for treatment vs. Log Time for Overall Survival
(After Matching)
Cohort 2
(All-Participants-as-Treated Population)



KEYNOTE-087, Database Cutoff Date: 15MAR2021, participants of Cohort 2
Comparator data based on Cheah et al. 2016
solid line represents LOESS (Locally Estimated Scatterplot Smoothing) curve. Area around the solid line shows 95% CI.

Figure 18

Log Hazard Ratio β(t) Over Time
(Before Matching)
Cohort 2
(All-Participants-as-Treated Population)



ρ = 0.237

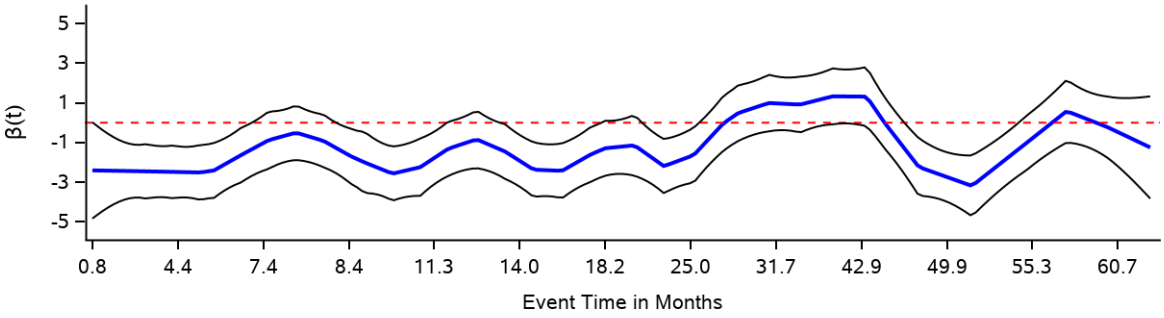
χ^2 = 3.424

p-value = 0.064

Database Cutoff Date: 15MAR2021
KEYNOTE-087, Database Cutoff Date: 15MAR2021, participants of Cohort 2
Comparator data based on Cheah et al. 2016
Estimation of hazard ratios over time is based on scaled Schoenfeld residuals

Figure 19

Log Hazard Ratio β(t) Over Time
(After Matching)
Cohort 2
(All-Participants-as-Treated Population)



ρ = 0.245

χ^2 = 4.013

p-value = 0.045

Database Cutoff Date: 15MAR2021
KEYNOTE-087, Database Cutoff Date: 15MAR2021, participants of Cohort 2
Comparator data based on Cheah et al. 2016
Estimation of hazard ratios over time is based on scaled Schoenfeld residuals

5.2.2 Pembrolizumab (SACT) vs. Post BV (Cheah et al. 2016)

5.2.2.1 Baseline Characteristics

Table 11

Relevant patient characteristics SACT (UK RWE data) and Cheah et al., 2016

Characteristics	Variable as measured in studies		SACT (n=215)	Cheah et al., 2016 (n=89)
Age	Age, n/N (%)	≥ 65 years		--
		< 65 years		--
		≥ 60 years	81 (38)	--
		< 60 years	134 (62)	--
	Age, median (range)		47	32 (18-84)
Sex	Male, n (%)	Yes	130 (60)	51/89 (53)
Disease status	Disease status, n/N (%)	Refractory	--	--
		Relapsed < 12 months	--	--
		Relapsed ≥ 12 months	--	--
Number of prior lines of therapy	Number of prior lines, n/N (%)	2	--	--
		3	--	--
		4		--
	Number of prior lines, median (range)		--	--
Prior auto-SCT	Prior auto-SCT, n (%)	Yes	--	--
Prior treatment	Radiation therapy, n (%)	Yes	--	--
Presence of B symptoms	Presence of B symptoms, n (%)	Yes	--	7/86(8)
Performance status	ECOG, n/N (%)	0	59 (27)	33/89 (37)
		>=1	104 (49)	56/89 (63)
		Missing	52(24)	
Presence of bulky disease	Bulky disease, n (%)	Bulky disease		--

Summary statistics are reported as described in both Real World Evidence SACT report (National Disease Registration Service, 2021) as well as publication of relevant comparator study (Cheah, et al., 2016).

5.2.2.2 Overall Survival

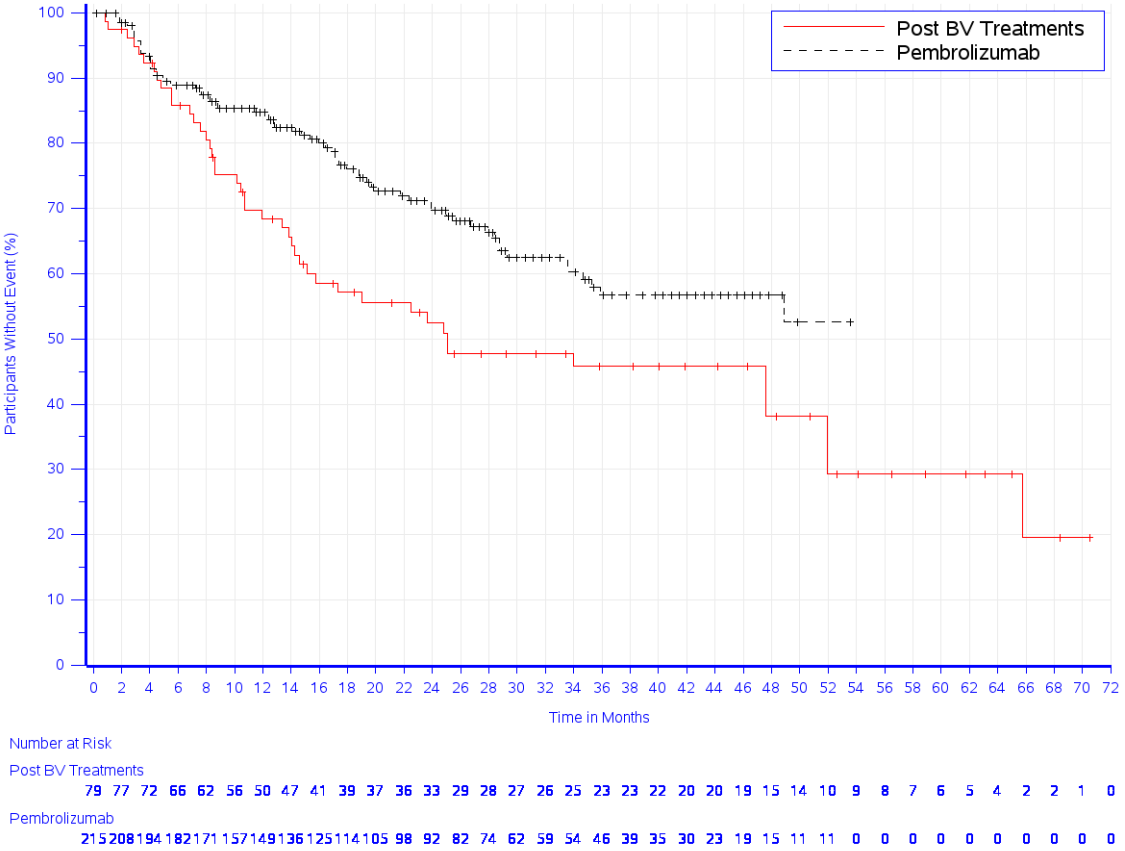
Table 12
Analysis of Overall Survival
Unadjusted Indirect Comparison Analysis of Pembrolizumab vs. Post BV Treatments
(modified All-Participants-as-Treated Population)

	Pembrolizumab			Post BV Treatments			Pembrolizumab vs. Post BV Treatments	
	N[a]	Participants with Event n (%)	Median Time[b] in Months [95 %-CI]	N[c]	Participants with Event n (%)	Median Time[b] in Months [95 %-CI]	Hazard Ratio [95 %-CI][d]	p-Value[d,e]
Overall survival	215	68 (31.6)	Not reached [35.27; -]	79	46 (58.2)	25.09 [14.58; 51.92]	0.59 [0.40; 0.86]	0.006

a: Number of participants: Based on Real World Evidence source (SACT database) in participants relapsing after treatment with or failing to respond to BV and ineligible for SCT, mAPaT
b: From product-limit (Kaplan-Meier) method
c: Number of participants: Based on Cheah et al. 2016, mAPaT population
d: Based on Cox regression model with treatment as a single covariate
e: Two-sided p-value using Wald test (Score test in case of zero event in one treatment group)
BV: Brentuximab Vedotin; CI: Confidence Interval; mAPaT: modified APaT population; SCT: Stem Cell Transplantation

Figure 20

Kaplan-Meier Curves of Overall Survival
Comparison of Pembrolizumab vs. Post BV Treatments
(modified All-Participants-as-Treated Population)

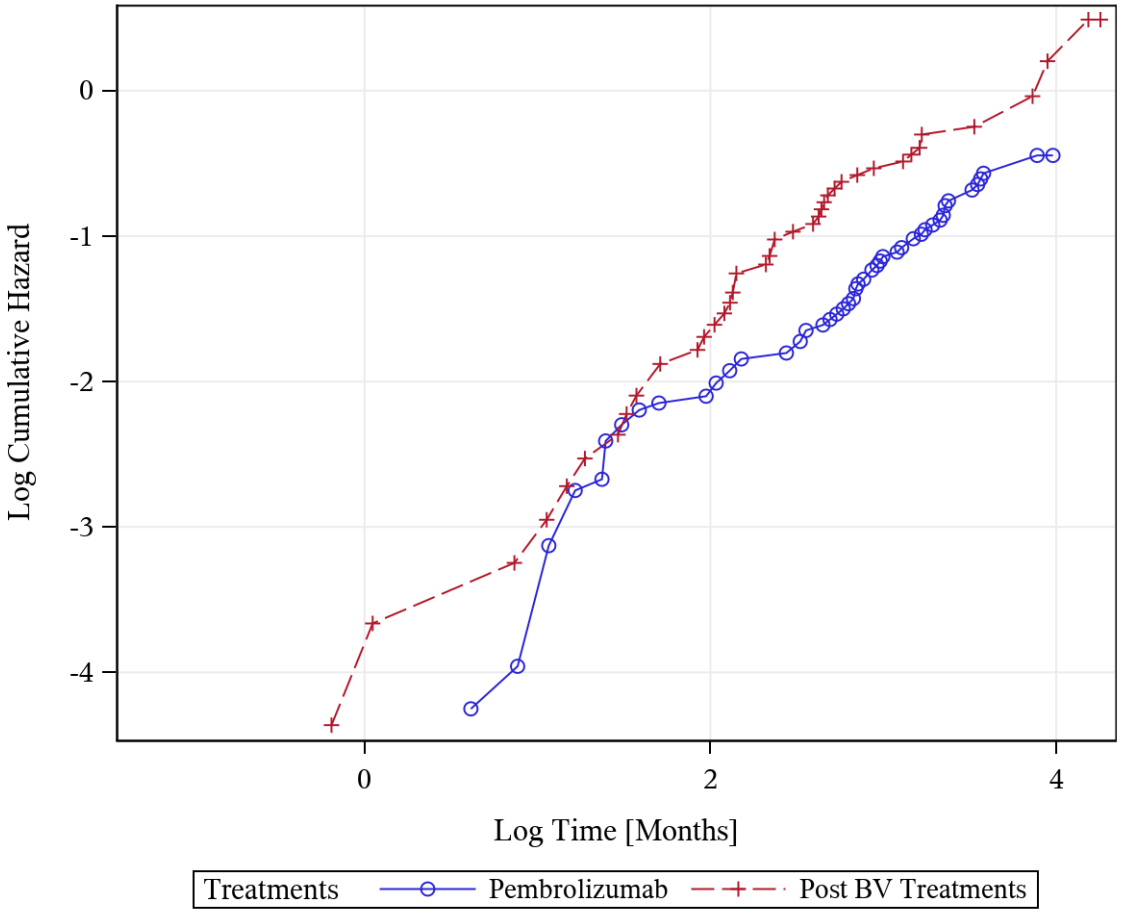


Pembrolizumab arm based on Real World Evidence source (SACT database) in participants relapsing after treatment with or failing to respond to BV and ineligible for SCT
Comparator data based on Cheah et al. 2016
BV: Brentuximab Vedotin; SCT: Stem Cell Transplantation;

5.2.2.3 Diagnostics

Figure 21

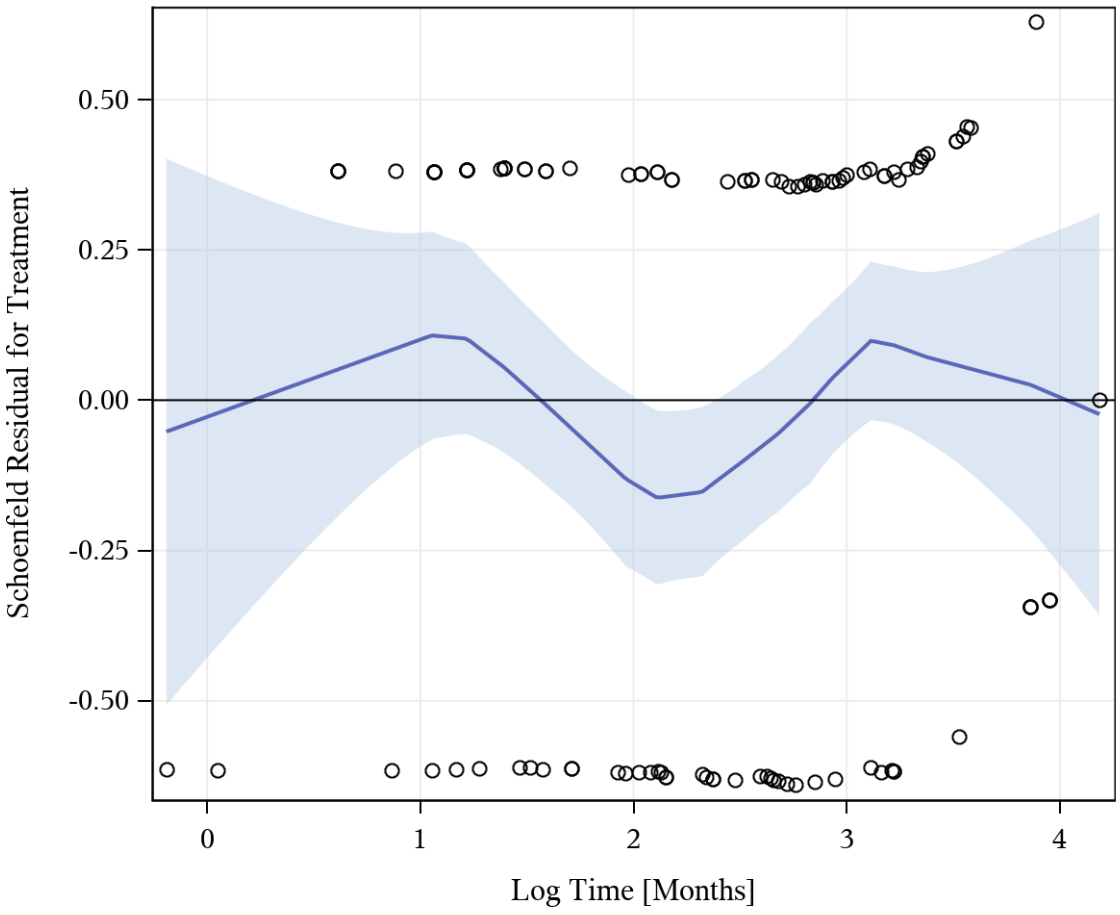
Diagnostic Plot
Log Cumulative Hazard vs. Log Time for Overall Survival
(modified All-Participants-as-Treated Population)



Pembrolizumab arm based on Real World Evidence source
(SACT database) in participants relapsing after treatment with or failing to respond to BV and ineligible for SCT
Comparator data based on Cheah et al. 2016
BV: Brentuximab Vedotin; SCT: Stem Cell Transplantation

Figure 22

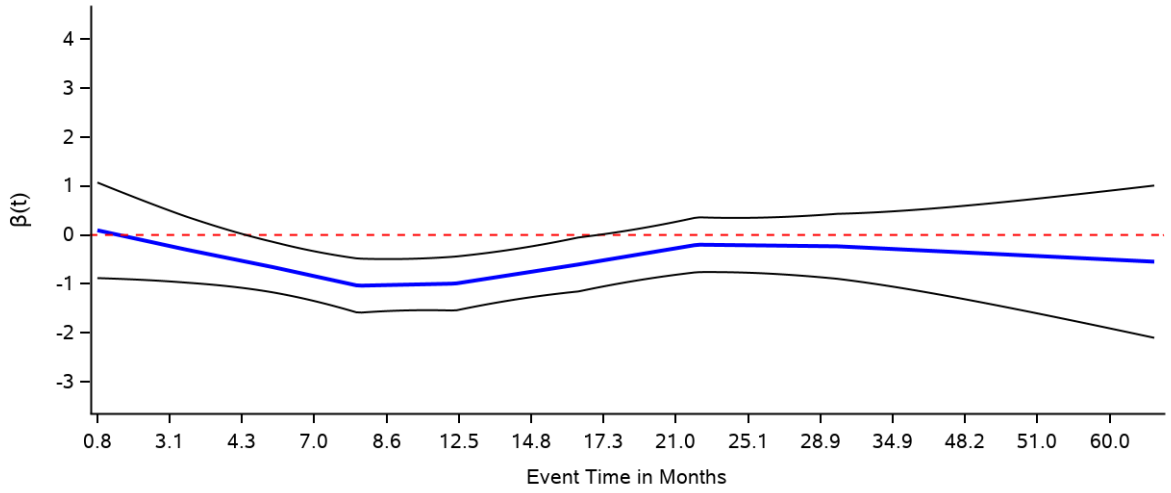
Diagnostic Plot
Schoenfeld residuals for treatment vs. Log Time for Overall Survival
(modified All-Participants-as-Treated Population)



Pembrolizumab arm based on Real World Evidence source
(SACT database) in participants relapsing after treatment with or failing to respond to BV and ineligible for SCT
Comparator data based on Cheah et al. 2016
solid line represents LOESS (Locally Estimated Scatterplot Smoothing) curve. Area around the solid line shows 95% CI.
BV: Brentuximab Vedotin ; SCT: Stem Cell Transplantation

Figure 23

Log Hazard Ratio β(t) Over Time
(modified All-Participants-as-Treated Population)



ρ = 0.037

χ^2 = 0.152

p-value = 0.697

Database Cutoff Date: 15MAR2021
Pembrolizumab arm based on Real World Evidence source (SACT database) in participants relapsing after treatment with or failing to respond to BV and ineligible for SCT
Estimation of hazard ratios over time is based on scaled Schoenfeld residuals. Comparator data based on Cheah et al. 2016

6 Appendix

6.1 UK RWE (SACT) data report

TA540_
Pembrolizumab_HL_

7 Bibliography

Cheah, C., Chihara, D., Horowitz, S., Sevin, A., Oki, Y., Zhou, S., . . . Romaguera, J. (2016). Patients with classical Hodgkin lymphoma experiencing disease progression after brentuximab vedotin have poor outcomes. *Annals of Oncology*, 1317-1323.

Eyre, T., Phillips, E., Linton, K., Kassam, S., Gibb, A., Allibone, S., . . . Stewart, G. (2017). Results of a multicentre UK-wide retrospective study evaluating the efficacy of brentuximab vedotin in relapsed, refractory classical Hodgkin lymphoma in the transplant naive setting. *British Journal of Haematology*, 471-479.

Grambsch, P. M., & Therneau, T. M. (1994). Proportional hazards tests and diagnostics based on weighted residuals. *Biometrika, 81*, 515-52.

Grambsch, P., & Therneau, T. (1994). Proportional hazards tests and diagnostics based on weighted residuals. *Biometrika*, 515-552.

Guyot, P., Ades, A., Ouwens, M., & Welton, N. (2012). Enhanced secondary analysis of survival data: reconstructing the data from published Kaplan-Meier survival curves. *BMC medical research methodology*, 1-13.

National Disease Registration Service. (2021). *Pembrolizumab for treating relapsed or refractory classical Hodgkin lymphoma – data review.* Leeds: NHS England.

Signorovitch, J. E., Sikirica, V., Haim Erder, M., & et al. (2012). Matching-Adjusted Indirect Comparisons: A New Tool for Timely Comparative Effectiveness Research. *Value in Health*, 940-947.

Signorovitch, J. E., Wu, E. Q., Yu, A. P., & et al. (2010). Comparative Effectiveness Without Head-to-Head Trials: A Method for Matching-Adjusted Indirect Comparisons Applied to Psoriasis Treatment with Adalimumab or Etanercept. *Pharmacoeconomics, 28*(10), 935-945.

KN204 Utility Analysis

Mitashri Chaudhuri <mitashri.chaudhuri@merck.com>, HEDS Oncology
Version 1.0, executed on Wed Oct 18 16:50:51 2023

Execution notes

1. Specification of the software system used to execute this package

System	Specification
platform	x86_64-w64-mingw32
arch	x86_64
os	mingw32
crt	ucrt
system	x86_64, mingw32
status	
major	4
minor	2.2
year	2022
month	10
day	31
svn rev	83211
language	R
version.string	R version 4.2.2 (2022-10-31 ucrt)
nickname	Innocent and Trusting

2. Specification of the R libraries used to execute this package

Package	Version
officer	0.3.19
haven	2.5.1
lmerTest	3.1.3
dplyr	1.0.10
ggplot2	3.4.0
bshazard	1.1
lsmeans	2.30.0
xlsx	0.6.5
stringr	1.4.1
flextable	0.6.10

1. Introduction

The purpose of this utility analysis, focused on the subgroup with >= 2 Lines of Prior Therapies and without Prior SCT, is to investigate how UK UTILITY VALUE is associated with patient characteristics at baseline, such as Age, Continuous Age, centralized at 40, Gender and Post-Treatment SCT Status. It is also of interest to understand how some dynamic parameters mediated potentially by the anti-tumor therapy interventions during the trial, such as Treatment and Grade 3-5 AE, are related to the utility score.

This report is organized as follows. First, linear-mixed effect model is conducted for each of the above factors using the longitudinally measured UK UTILITY VALUE as the

outcome and individual factors of interest as the single covariate. Followed by the reports of these univariate analyses, results from a multi-variate linear mixed effect model, with all of the above factors included as covariates, are presented. Based on the statistical significance and clinical interpretations, a final multi-variate linear mixed effect is chosen and reported.

2. Exploration of the Associations of Various Baseline Characteristics with UK UTILITY VALUE

2.1. Age

```
Type III Analysis of Variance Table with Satterthwaite's method
     Sum Sq Mean Sq NumDF  DenDF F value Pr(>F)
Age 0.05145 0.05145     1 120.49  1.6115 0.2067



2 x 2 Matrix of class "dpoMatrix"
               Age < 40     Age >= 40
Age < 40   0.0003946271 -0.0003946271
Age >= 40 -0.0003946271  0.0011035406



Linear mixed model fit by REML. t-tests use Satterthwaite's method ['lmerModLmerTest']
Formula: y ~ z + (1 | USUBJID)
   Data: thisData

REML criterion at convergence: -191.3

Scaled residuals:
    Min      1Q  Median      3Q     Max
-4.2590 -0.3030  0.1730  0.5818  2.4577

Random effects:
 Groups   Name        Variance Std.Dev.
 USUBJID  (Intercept) 0.02659  0.1631
 Residual             0.03193  0.1787
Number of obs: 684, groups:  USUBJID, 134

Fixed effects:
           Estimate Std. Error        df t value Pr(>|t|)
Age < 40    0.78891    0.01987 118.82531  39.713   <2e-16 ***
Age >= 40  -0.04217    0.03322 120.48779  -1.269    0.207
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Indicated from the ANOVA summary table above, there lacks of strong evidence that Age is associated with the UK UTILITY VALUE(P-value: 2.067e-01)

Note that, in the coefficient output table above, The 1st row refers to the estimate in the reference group (so-called Intercept) while the rows below the reference group should be interpreted as the difference between the group at that row and the reference group.

2.2. Continuous Age, centralized at 40

```
Type III Analysis of Variance Table with Satterthwaite's method
                                  Sum Sq  Mean Sq NumDF  DenDF F value Pr(>F)
Continuous Age, centralized at 40 0.061572 0.061572     1 120.94  1.9284 0.1675



2 x 2 Matrix of class "dpoMatrix"
                                    Intercept Continuous Age, centralized at 40
Intercept                        2.528412e-04                      1.318120e-07
Continuous Age, centralized at 40 1.318120e-07                     8.108751e-07



Linear mixed model fit by REML. t-tests use Satterthwaite's method ['lmerModLmerTest']
Formula: y ~ z + (1 | USUBJID)
   Data: thisData

REML criterion at convergence: -184.4

Scaled residuals:
    Min      1Q  Median      3Q     Max
-4.2583 -0.3094  0.1676  0.5859  2.4581

Random effects:
 Groups   Name        Variance Std.Dev.
 USUBJID  (Intercept) 0.02649  0.1628
 Residual             0.03193  0.1787
Number of obs: 684, groups:  USUBJID, 134

Fixed effects:
                                   Estimate Std. Error         df t value Pr(>|t|)
Intercept                         7.736e-01  1.590e-02  1.198e+02  48.654   <2e-16 ***
Continuous Age, centralized at 40 -1.251e-03  9.005e-04  1.209e+02  -1.389    0.167
---
Signif. codes:  0 â€˜***â€™ 0.001 â€˜**â€™ 0.01 â€˜*â€™ 0.05 â€˜.â€™ 0.1 â€˜ â€™ 1
```

Indicated from the ANOVA summary table above, there lacks of strong evidence that Continuous Age, centralized at 40 is associated with the UK UTILITY VALUE(P-value: 1.675e-01)

Note that, in the coefficient output table above, The 1st row refers to the estimate in the reference group (so-called Intercept) while the rows below the reference group should be interpreted as the difference between the group at that row and the reference group.

2.3. Gender

```
Type III Analysis of Variance Table with Satterthwaite's method
         Sum Sq  Mean Sq NumDF  DenDF F value Pr(>F)
Gender 0.060882 0.060882     1 119.15  1.9073 0.1698



2 x 2 Matrix of class "dpoMatrix"
                Gender = Female Gender = Male
Gender = Female    0.0005858166 -0.0005858166
Gender = Male     -0.0005858166  0.0010316631



Linear mixed model fit by REML. t-tests use Satterthwaite's method ['lmerModLmerTest']
Formula: y ~ z + (1 | USUBJID)
   Data: thisData

REML criterion at convergence: -191.5

Scaled residuals:
    Min      1Q  Median      3Q     Max
-4.2692 -0.2820  0.1699  0.5508  2.4872

Random effects:
 Groups   Name        Variance Std.Dev.
 USUBJID  (Intercept) 0.02654  0.1629
 Residual             0.03192  0.1787
Number of obs: 684, groups:  USUBJID, 134

Fixed effects:
                 Estimate Std. Error        df t value Pr(>|t|)
Gender = Female   0.74865    0.02420 116.78923  30.931   <2e-16 ***
Gender = Male     0.04436    0.03212 119.15111   1.381     0.17
---
Signif. codes:  0 â€˜***â€™ 0.001 â€˜**â€™ 0.01 â€˜*â€™ 0.05 â€˜.â€™ 0.1 â€˜ â€™ 1
```

Indicated from the ANOVA summary table above, there lacks of strong evidence that Gender is associated with the UK UTILITY VALUE(P-value: 1.698e-01)

Note that, in the coefficient output table above, The 1st row refers to the estimate in the reference group (so-called Intercept) while the rows below the reference group should be interpreted as the difference between the group at that row and the reference group.

2.4. Post-Treatment SCT Status

```
Type III Analysis of Variance Table with Satterthwaite's method
                          Sum Sq Mean Sq NumDF  DenDF F value Pr(>F)
Post-Treatment SCT Status 0.03013 0.03013     1 596.81  0.9434 0.3318



2 x 2 Matrix of class "dpoMatrix"
              Before SCT     After SCT
 Before SCT  2.54936e-04 -0.0000614147
 After SCT   -6.14147e-05  0.0076890362



Linear mixed model fit by REML. t-tests use Satterthwaite's method ['lmerModLmerTest']
Formula: y ~ z + (1 | USUBJID)
   Data: thisData

REML criterion at convergence: -192.6

Scaled residuals:
    Min      1Q  Median      3Q     Max
-4.2221 -0.3034  0.1852  0.5570  2.4695

Random effects:
 Groups   Name        Variance Std.Dev.
 USUBJID  (Intercept) 0.02671  0.1634
 Residual             0.03194  0.1787
Number of obs: 684, groups:  USUBJID, 134

Fixed effects:
             Estimate Std. Error        df t value Pr(>|t|)
 Before SCT   0.77313    0.01597 121.18613  48.421   <2e-16 ***
 After SCT    0.08517    0.08769 596.80843   0.971    0.332
---
Signif. codes:  0 â€˜***â€™ 0.001 â€˜**â€™ 0.01 â€˜*â€™ 0.05 â€˜.â€™ 0.1 â€˜ â€™ 1
```

Indicated from the ANOVA summary table above, there lacks of strong evidence that Post-Treatment SCT Status is associated with the UK UTILITY VALUE(P-value: 3.318e-01)

Note that, in the coefficient output table above, The 1st row refers to the estimate in the reference group (so-called Intercept) while the rows below the reference group should be interpreted as the difference between the group at that row and the reference group.

3. Exploration of the Associations of Several Health Statuses with UK UTILITY VALUE That May be Treatment-mediated

3.1. Treatment

```
Type III Analysis of Variance Table with Satterthwaite's method
           Sum Sq Mean Sq NumDF  DenDF F value   Pr(>F)
Treatment 0.24155 0.24155     1 118.66  7.5565 0.006915 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

```
2 x 2 Matrix of class "dpoMatrix"
                                Treatment = Brentuximab Vedotin Treatment = MK-3475 200 mg
Treatment = Brentuximab Vedotin                     0.000486685              -0.0004866850
Treatment = MK-3475 200 mg                         -0.000486685               0.0009655153
```

```
Linear mixed model fit by REML. t-tests use Satterthwaite's method ['lmerModLmerTest']
Formula: y ~ z + (1 | USUBJID)
   Data: thisData

REML criterion at convergence: -196.9

Scaled residuals:
    Min      1Q  Median      3Q     Max
-4.1785 -0.2733  0.1587  0.5653  2.4900

Random effects:
 Groups   Name        Variance Std.Dev.
 USUBJID  (Intercept) 0.02497  0.1580
 Residual             0.03197  0.1788
Number of obs: 684, groups:  USUBJID, 134

Fixed effects:
                               Estimate Std. Error        df t value Pr(>|t|)
Treatment = Brentuximab Vedotin  0.73103    0.02206 123.59641  33.137  < 2e-16 ***
Treatment = MK-3475 200 mg       0.08542    0.03107 118.66333   2.749  0.00691 **
---
```

```
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Indicated from the ANOVA summary table above, there exists evidence or strong evidence that Treatment is associated with the UK UTILITY VALUE(P-value: 6.915e-03)

Note that, in the coefficient output table above, The 1st row refers to the estimate in the reference group (so-called Intercept) while the rows below the reference group should be interpreted as the difference between the group at that row and the reference group.

3.2. Grade 3-5 AE

```
Type III Analysis of Variance Table with Satterthwaite's method
               Sum Sq  Mean Sq NumDF  DenDF F value Pr(>F)
Grade 3-5 AE 0.026529 0.026529     1 680.62  0.8298 0.3626



2 x 2 Matrix of class "dpoMatrix"
                     w/o Grade3+ AE  During Grade3+ AEs
 w/o Grade3+ AE        2.660699e-04       -9.630258e-05
 During Grade3+ AEs   -9.630258e-05        7.431893e-04



Linear mixed model fit by REML. t-tests use Satterthwaite's method ['lmerModLmerTest']
Formula: y ~ z + (1 | USUBJID)
   Data: thisData

REML criterion at convergence: -190.1

Scaled residuals:
    Min      1Q  Median      3Q     Max
-4.2322 -0.3074  0.1828  0.5849  2.5430

Random effects:
 Groups   Name        Variance Std.Dev.
 USUBJID  (Intercept) 0.02659  0.1631
 Residual             0.03197  0.1788
Number of obs: 684, groups:  USUBJID, 134

Fixed effects:
                    Estimate Std. Error        df t value Pr(>|t|)
 w/o Grade3+ AE      0.77705    0.01631 129.80089  47.638   <2e-16 ***
 During Grade3+ AEs -0.02483    0.02726 680.61720  -0.911    0.363
---
Signif. codes:  0 â€˜***â€™ 0.001 â€˜**â€™ 0.01 â€˜*â€™ 0.05 â€˜.â€™ 0.1 â€˜ â€™ 1
```

Indicated from the ANOVA summary table above, there lacks of strong evidence that Grade 3-5 AE is associated with the UK UTILITY VALUE(P-value: 3.626e-01)

Note that, in the coefficient output table above, The 1st row refers to the estimate in the reference group (so-called Intercept) while the rows below the reference group should be interpreted as the difference between the group at that row and the reference group.

4. Multivariate Modeling of UK UTILITY VALUE

4.1. Full Model

All of the above individual baseline and time-dependent factors are included in the full model below for a multivariate analysis. From the consideration of easier interpretation and comparison of the p-values for the 2 covariates for age effects, Age Group with cutoff at 40 years old and continuous age centralized at 40, in the association strength with the utility values, Age group is used to represent age effect in the following multi-variate models.

Type III Analysis of Variance Table with Satterthwaite's method

	Sum Sq	Mean Sq	NumDF	DenDF	F value	Pr(>F)	
Continuous Age, centralized at 40	0.101508	0.101508	1	121.59	3.1711	0.077451	.
Gender	0.097915	0.097915	1	117.87	3.0588	0.082902	.
Post-Treatment SCT Status	0.029516	0.029516	1	599.05	0.9221	0.337320	
Treatment	0.258040	0.258040	1	117.44	8.0611	0.005331	**
Grade 3-5 AE	0.015861	0.015861	1	677.25	0.4955	0.481731	

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

6 x 6 Matrix of class "dpoMatrix"

	Intercept	Continuous Age, centralized at 40	Gender
Intercept	8.217546e-04	4.134327e-06	-5.698487e-04
Continuous Age, centralized at 40	4.134327e-06	7.899490e-07	-5.158407e-06
Gender	-5.698487e-04	-5.158407e-06	9.902236e-04
Post-Treatment SCT Status	3.972406e-06	1.695116e-06	-5.877464e-05
Treatment = MK-3475 200 mg	-4.908094e-04	-1.518862e-06	1.482030e-05
Grade 3-5 AE = During Grade3+ AEs	-1.206378e-04	-2.515733e-06	1.268749e-06

	Post-Treatment SCT Status	Treatment = MK-3475 200 mg
Intercept	3.972406e-06	-4.908094e-04
Continuous Age, centralized at 40	1.695116e-06	-1.518862e-06
Gender	-5.877464e-05	1.482030e-05
Post-Treatment SCT Status	7.712267e-03	-2.187146e-05
Treatment = MK-3475 200 mg	-2.187146e-05	9.440323e-04
Grade 3-5 AE = During Grade3+ AEs	-1.614958e-04	4.780026e-05

	Grade 3-5 AE = During Grade3+ AEs
Intercept	-1.206378e-04
Continuous Age, centralized at 40	-2.515733e-06
Gender	1.268749e-06
Post-Treatment SCT Status	-1.614958e-04
Treatment = MK-3475 200 mg	4.780026e-05
Grade 3-5 AE = During Grade3+ AEs	7.469600e-04

Linear mixed model fit by REML. t-tests use Satterthwaite's method ['lmerModLmerTest']

```
Formula: theFormula
   Data: thisData

REML criterion at convergence: -178.1

Scaled residuals:
    Min      1Q  Median      3Q     Max
-4.2956 -0.2695  0.1345  0.5654  2.5548

Random effects:
 Groups   Name        Variance Std.Dev.
 USUBJID  (Intercept) 0.02407  0.1552
 Residual             0.03201  0.1789
Number of obs: 684, groups:  USUBJID, 134

Fixed effects:
                                    Estimate Std. Error         df t value Pr(>|t|)
Intercept                          7.006e-01  2.867e-02  1.204e+02  24.439  < 2e-16 ***
Continuous Age, centralized at 40 -1.583e-03  8.888e-04  1.216e+02  -1.781  0.07745 .
Gender                             5.504e-02  3.147e-02  1.179e+02   1.749  0.08290 .
Post-Treatment SCT Status          8.433e-02  8.782e-02  5.990e+02   0.960  0.33732
Treatment = MK-3475 200 mg         8.723e-02  3.073e-02  1.174e+02   2.839  0.00533 **
Grade 3-5 AE = During Grade3+ AEs -1.924e-02  2.733e-02  6.772e+02  -0.704  0.48173
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Indicated from the ANOVA summary table above, there exist evidences or strong evidences that factors Treatment (P-value: 5.331e-03) are associated with UK UTILITY VALUE

4.2. Final Model

Based on the significance of the results from the above full model, a final model is proposed to include continuous age (centralized at 40), Post-Treatment SCT Status, Treatment and Grade 3-5 AE status

```
Type III Analysis of Variance Table with Satterthwaite's method
                                  Sum Sq  Mean Sq NumDF  DenDF F value   Pr(>F)
Continuous Age, centralized at 40 0.069451 0.069451     1 120.54  2.1694 0.143390
Post-Treatment SCT Status         0.031767 0.031767     1 598.51  0.9923 0.319589
Treatment                         0.249435 0.249435     1 118.25  7.7913 0.006123 **
Grade 3-5 AE                      0.015866 0.015866     1 677.77  0.4956 0.481688
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

```
5 x 5 Matrix of class "dpoMatrix"
                                     Intercept Continuous Age, centralized at 40
Intercept                         5.010624e-04                      1.180527e-06
Continuous Age, centralized at 40 1.180527e-06                      7.747495e-07
Post-Treatment SCT Status         -2.988029e-05                     1.390850e-06
Treatment = MK-3475 200 mg        -4.894943e-04                    -1.465462e-06
Grade 3-5 AE = During Grade3+ AEs -1.202140e-04                    -2.518579e-06
                                  Post-Treatment SCT Status Treatment = MK-3475 200 mg
Intercept                                    -2.988029e-05               -4.894943e-04
Continuous Age, centralized at 40             1.390850e-06               -1.465462e-06
Post-Treatment SCT Status                     7.716062e-03               -2.098817e-05
Treatment = MK-3475 200 mg                   -2.098817e-05                9.583848e-04
Grade 3-5 AE = During Grade3+ AEs            -1.616712e-04                4.783003e-05
                                  Grade 3-5 AE = During Grade3+ AEs
Intercept                                          -1.202140e-04
Continuous Age, centralized at 40                  -2.518579e-06
Post-Treatment SCT Status                          -1.616712e-04
Treatment = MK-3475 200 mg                          4.783003e-05
Grade 3-5 AE = During Grade3+ AEs                   7.490438e-04
```

```
Linear mixed model fit by REML. t-tests use Satterthwaite's method ['lmerModLmerTest']
Formula: AVAL ~ AGEoffset + PFSONT2 + TRT01P + G35AECAT + (1 | USUBJID)
   Data: thisData

REML criterion at convergence: -180.1
```

```
Scaled residuals:
    Min      1Q  Median      3Q     Max
-4.2220 -0.2824  0.1498  0.5563  2.5349

Random effects:
 Groups   Name        Variance Std.Dev.
 USUBJID  (Intercept) 0.02455  0.1567
 Residual             0.03201  0.1789
Number of obs: 684, groups:  USUBJID, 134

Fixed effects:
                                    Estimate Std. Error         df t value Pr(>|t|)
Intercept                          7.322e-01  2.238e-02  1.313e+02  32.709  < 2e-16 ***
Continuous Age, centralized at 40 -1.296e-03  8.802e-04  1.205e+02  -1.473  0.14339
Post-Treatment SCT Status          8.750e-02  8.784e-02  5.985e+02   0.996  0.31959
Treatment = MK-3475 200 mg         8.641e-02  3.096e-02  1.182e+02   2.791  0.00612 **
Grade 3-5 AE = During Grade3+ AEs -1.927e-02  2.737e-02  6.778e+02  -0.704  0.48169
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Indicated from the ANOVA summary table above, all of the factors included in the final model, Treatment (P-value: 6.123e-03) are associated with UK UTILITY VALUE

4.3. Alternative Final Model without Age Effect

Based on the significance of the results from the above final model, an alternative final model is proposed to exclude age effect, with Post-Treatment SCT Status, Treatment and Grade 3-5 AE status remained in the model.

```
Type III Analysis of Variance Table with Satterthwaite's method
                          Sum Sq  Mean Sq NumDF  DenDF F value   Pr(>F)
Post-Treatment SCT Status 0.033420 0.033420     1 598.15  1.0439 0.307331
Treatment                 0.234052 0.234052     1 119.06  7.3108 0.007858 **
Grade 3-5 AE              0.023684 0.023684     1 679.44  0.7398 0.390036
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

```
4 x 4 Matrix of class "dpoMatrix"
                                      Intercept Post-Treatment SCT Status = After SCT
Intercept                          5.035275e-04                          -3.201272e-05
Post-Treatment SCT Status = After SCT -3.201272e-05                          7.717295e-03
Treatment = MK-3475 200 mg         -4.915024e-04                          -1.835006e-05
Grade 3-5 AE = During Grade3+ AEs  -1.165445e-04                          -1.573122e-04
                                   Treatment = MK-3475 200 mg
Intercept                                      -4.915024e-04
Post-Treatment SCT Status = After SCT          -1.835006e-05
Treatment = MK-3475 200 mg                      9.641921e-04
Grade 3-5 AE = During Grade3+ AEs               4.307750e-05
                                   Grade 3-5 AE = During Grade3+ AEs
Intercept                                           -0.0001165445
Post-Treatment SCT Status = After SCT               -0.0001573122
Treatment = MK-3475 200 mg                           0.0000430775
Grade 3-5 AE = During Grade3+ AEs                    0.0007420597
```

```
Linear mixed model fit by REML. t-tests use Satterthwaite's method ['lmerModLmerTest']
Formula: AVAL ~ PFSONT2 + TRT01P + G35AECAT + (1 | USUBJID)
   Data: thisData

REML criterion at convergence: -190.2

Scaled residuals:
    Min      1Q  Median      3Q     Max
-4.1869 -0.2715  0.1581  0.5641  2.5593
```

```
Random effects:
 Groups   Name        Variance Std.Dev.
 USUBJID  (Intercept) 0.02484  0.1576
 Residual             0.03201  0.1789
Number of obs: 684, groups:  USUBJID, 134

Fixed effects:
                                      Estimate Std. Error        df t value Pr(>|t|)
Intercept                              0.73410    0.02244 131.66799  32.715  < 2e-16 ***
Post-Treatment SCT Status = After SCT  0.08976    0.08785 598.14744   1.022  0.30733
Treatment = MK-3475 200 mg             0.08396    0.03105 119.06185   2.704  0.00786 **
Grade 3-5 AE = During Grade3+ AEs     -0.02343    0.02724 679.43912  -0.860  0.39004
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Indicated from the ANOVA summary table above, all of the factors included in the alternative final model, Treatment (P-value: 7.858e-03) are associated with UK UTILITY VALUE

4.4. Alternative Final Model with only Treatment and Grade 3-5 AE status

Based on the significance of the results from the above final model, an alternative final model is proposed to exclude Post-Treatment SCT Status, with only Treatment and Grade 3-5 AE status remained in the model.

```
Type III Analysis of Variance Table with Satterthwaite's method
               Sum Sq  Mean Sq NumDF  DenDF F value   Pr(>F)
Treatment    0.234794 0.234794     1 119.02  7.3362 0.007754 **
Grade 3-5 AE 0.020189 0.020189     1 680.43  0.6308 0.427337
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

```
3 x 3 Matrix of class "dpoMatrix"
                                    Intercept Treatment = MK-3475 200 mg
Intercept                          0.0005041740              -4.923565e-04
Treatment = MK-3475 200 mg        -0.0004923565               9.657318e-04
Grade 3-5 AE = During Grade3+ AEs -0.0001171975               4.269269e-05
                                  Grade 3-5 AE = During Grade3+ AEs
Intercept                                            -1.171975e-04
Treatment = MK-3475 200 mg                            4.269269e-05
Grade 3-5 AE = During Grade3+ AEs                     7.388851e-04
```

```
Linear mixed model fit by REML. t-tests use Satterthwaite's method ['lmerModLmerTest']
Formula: AVAL ~ TRT01P + G35AECAT + (1 | USUBJID)
   Data: thisData

REML criterion at convergence: -192.2

Scaled residuals:
    Min      1Q  Median      3Q     Max
-4.1866 -0.2727  0.1604  0.5636  2.5539

Random effects:
 Groups   Name        Variance Std.Dev.
 USUBJID  (Intercept) 0.02489  0.1578
 Residual             0.03200  0.1789
Number of obs: 684, groups:  USUBJID, 134

Fixed effects:
```

```
                                Estimate Std. Error        df t value Pr(>|t|)
Intercept                        0.73447    0.02245 131.54691  32.710  < 2e-16 ***
Treatment = MK-3475 200 mg       0.08417    0.03108 119.01520   2.709  0.00775 **
Grade 3-5 AE = During Grade3+ AEs -0.02159    0.02718 680.43030  -0.794  0.42734
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Indicated from the ANOVA summary table above, all of the factors included in the alternative final model, Treatment (P-value: 7.754e-03) are associated with UK UTILITY VALUE

4.5. Final Model with An Interaction Term between Treatment and Grade 3-5 AE status

Based on the significance of the results from the above full model, an alternative final model is proposed to include Treatment, Grade 3-5 AE status, Treatment-by-Grade 3-5 AE status interaction

```
 contrast                  estimate     SE  df t.ratio p.value
 TRTdiffINchngFROM0ATwk24   0.0226 0.0549 679   0.412  0.6807

Degrees-of-freedom method: kenward-roger
Type III Analysis of Variance Table with Satterthwaite's method
                            Sum Sq Mean Sq NumDF  DenDF F value  Pr(>F)
Treatment                  0.19828 0.19828     1 211.02  6.1893 0.01363 *
Grade 3-5 AE               0.01769 0.01769     1 678.40  0.5522 0.45768
(Treatment) X (Grade 3-5 AE) 0.00546 0.00546     1 678.40  0.1704 0.67986
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1



4 x 4 Matrix of class "dpoMatrix"
                                                                    Intercept
Intercept                                                          0.0005201922
Treatment = MK-3475 200 mg                                        -0.0005201922
Grade 3-5 AE = During Grade3+ AEs                                 -0.0002113567
Treatment = MK-3475 200 mg & Grade 3-5 AE = During Grade3+ AEs     0.0002113567
                                                                  Treatment = MK-3475 200 mg
Intercept                                                                      -0.0005201922
Treatment = MK-3475 200 mg                                                      0.0010158742
Grade 3-5 AE = During Grade3+ AEs                                               0.0002113567
Treatment = MK-3475 200 mg & Grade 3-5 AE = During Grade3+ AEs                 -0.0003790555
                                                                  Grade 3-5 AE = During Grade3+ AEs
Intercept                                                                            -0.0002113567
Treatment = MK-3475 200 mg                                                            0.0002113567
Grade 3-5 AE = During Grade3+ AEs                                                     0.0013325569
Treatment = MK-3475 200 mg & Grade 3-5 AE = During Grade3+ AEs                       -0.0013325569
                                                                  Treatment = MK-3475 200 mg & Grade 3-5 AE =

During Grade3+ AEs
Intercept
0.0002113567
Treatment = MK-3475 200 mg
-0.0003790555
Grade 3-5 AE = During Grade3+ AEs
-0.0013325569
```

```
Treatment = MK-3475 200 mg & Grade 3-5 AE = During Grade3+ AEs
0.0029955557



Linear mixed model fit by REML. t-tests use Satterthwaite's method ['lmerModLmerTest']
Formula: AVAL ~ TRT01P * G35AECAT + (1 | USUBJID)
   Data: thisData

REML criterion at convergence: -188.4

Scaled residuals:
    Min      1Q  Median      3Q     Max
-4.1869 -0.2755  0.1581  0.5642  2.5836

Random effects:
 Groups   Name        Variance Std.Dev.
 USUBJID  (Intercept) 0.02495  0.158
 Residual             0.03204  0.179
Number of obs: 684, groups:  USUBJID, 134

Fixed effects:
                                                                 Estimate Std. Error        df
Intercept                                                         0.73605    0.02281 137.68676
Treatment = MK-3475 200 mg                                        0.08131    0.03187 128.51173
Grade 3-5 AE = During Grade3+ AEs                                -0.03163    0.03650 678.46493
Treatment = MK-3475 200 mg & Grade 3-5 AE = During Grade3+ AEs    0.02259    0.05473 678.39776
                                                                t value Pr(>|t|)
Intercept                                                        32.272   <2e-16 ***
Treatment = MK-3475 200 mg                                        2.551   0.0119 *
Grade 3-5 AE = During Grade3+ AEs                                -0.867   0.3865
Treatment = MK-3475 200 mg & Grade 3-5 AE = During Grade3+ AEs    0.413   0.6799
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Indicated from the ANOVA summary table above, all of the factors included in the final model, Treatment (P-value: 1.363e-02) are associated with UK UTILITY VALUE

4.6. Alternative Final Model with only Treatment and Post-Treatment SCT Status

Based on the significance of the results from the above final model, an alternative final model is proposed to exclude Grade 3-5 AE status, with only Post-Treatment SCT Status and Treatment remained in the model.

```
Type III Analysis of Variance Table with Satterthwaite's method
                          Sum Sq  Mean Sq NumDF  DenDF F value   Pr(>F)
Post-Treatment SCT Status 0.029928 0.029928     1 598.99  0.9359 0.333728
Treatment                 0.241375 0.241375     1 118.72  7.5480 0.006945 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

```
3 x 3 Matrix of class "dpoMatrix"
                                      Intercept Post-Treatment SCT Status = After SCT
Intercept                          4.865185e-04                          -5.667047e-05
Post-Treatment SCT Status = After SCT -5.667047e-05                           7.676920e-03
Treatment = MK-3475 200 mg            -4.860322e-04                          -9.205707e-06
                                   Treatment = MK-3475 200 mg
Intercept                                        -4.860322e-04
Post-Treatment SCT Status = After SCT            -9.205707e-06
Treatment = MK-3475 200 mg                        9.643349e-04
```

```
Linear mixed model fit by REML. t-tests use Satterthwaite's method ['lmerModLmerTest']
Formula: AVAL ~ PFSONT2 + TRT01P + (1 | USUBJID)
   Data: thisData

REML criterion at convergence: -194.8

Scaled residuals:
    Min      1Q  Median      3Q     Max
-4.1778 -0.2677  0.1596  0.5659  2.4897

Random effects:
 Groups   Name        Variance Std.Dev.
 USUBJID  (Intercept) 0.02493  0.1579
 Residual             0.03198  0.1788
Number of obs: 684, groups:  USUBJID, 134

Fixed effects:
```

	Estimate	Std. Error	df	t value	Pr(>|t|)	
Intercept	0.73041	0.02206	123.85757	33.114	< 2e-16	***
Post-Treatment SCT Status = After SCT	0.08476	0.08762	598.99002	0.967	0.33373	
Treatment = MK-3475 200 mg	0.08532	0.03105	118.72100	2.747	0.00694	**

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Indicated from the ANOVA summary table above, all of the factors included in the alternative final model, Treatment (P-value: 6.945e-03) are associated with UK UTILITY VALUE

4.7. Final Model with An Interaction Term between Treatment and Post-Treatment SCT Status

Based on the significance of the results from the above full model, an alternative final model is proposed to include Treatment, Post-Treatment SCT Status and Treatment-by-Post Treatment SCT Status interaction

```
 contrast                estimate   SE  df t.ratio p.value
 TRTdiffINchngFROM0ATwk24   0.0462 0.18 612   0.256  0.7978

Degrees-of-freedom method: kenward-roger
Type III Analysis of Variance Table with Satterthwaite's method
                                        Sum Sq  Mean Sq NumDF  DenDF F value Pr(>F)
Treatment                             0.042453 0.042453     1 676.02  1.3256 0.2500
Post-Treatment SCT Status             0.025089 0.025089     1 604.46  0.7834 0.3765
(Treatment) X (Post-Treatment SCT Status) 0.002108 0.002108     1 604.46  0.0658 0.7976



4 x 4 Matrix of class "dpoMatrix"
                                                                    Intercept
Intercept                                                          0.0004876106
Treatment = MK-3475 200 mg                                        -0.0004876106
Post-Treatment SCT Status = After SCT                             -0.0001461248
Treatment = MK-3475 200 mg & Post-Treatment SCT Status = After SCT  0.0001461248
                                                                  Treatment = MK-3475 200 mg
Intercept                                                                -0.0004876106
Treatment = MK-3475 200 mg                                                0.0009671728
Post-Treatment SCT Status = After SCT                                     0.0001461248
Treatment = MK-3475 200 mg & Post-Treatment SCT Status = After SCT       -0.0002539907
                                                                  Post-Treatment SCT Status = After SCT
Intercept                                                                       -0.0001461248
Treatment = MK-3475 200 mg                                                       0.0001461248
Post-Treatment SCT Status = After SCT                                            0.0197950981
Treatment = MK-3475 200 mg & Post-Treatment SCT Status = After SCT              -0.0197950981
                                                                  Treatment = MK-3475 200 mg & Post-Treatment

SCT Status = After SCT
Intercept
0.0001461248
Treatment = MK-3475 200 mg
-0.0002539907
Post-Treatment SCT Status = After SCT
-0.0197950981
Treatment = MK-3475 200 mg & Post-Treatment SCT Status = After SCT
0.0323652827
```

```
Linear mixed model fit by REML. t-tests use Satterthwaite's method ['lmerModLmerTest']
Formula: AVAL ~ TRT01P * PFSONT2 + (1 | USUBJID)
   Data: thisData

REML criterion at convergence: -193.3

Scaled residuals:
    Min      1Q  Median      3Q     Max
-4.1756 -0.2686  0.1597  0.5656  2.4875

Random effects:
 Groups   Name        Variance Std.Dev.
 USUBJID  (Intercept) 0.02495  0.1579
 Residual             0.03203  0.1790
Number of obs: 684, groups:  USUBJID, 134

Fixed effects:
                                                                        Estimate Std. Error
Intercept                                                               0.73062    0.02208
Treatment = MK-3475 200 mg                                              0.08495    0.03110
Post-Treatment SCT Status = After SCT                                   0.05654    0.14070
Treatment = MK-3475 200 mg & Post-Treatment SCT Status = After SCT      0.04616    0.17990
                                                                              df t value Pr(>|t|)
Intercept                                                               124.13416  33.087  < 2e-16
Treatment = MK-3475 200 mg                                              119.15849   2.732  0.00726
Post-Treatment SCT Status = After SCT                                   615.26948   0.402  0.68794
Treatment = MK-3475 200 mg & Post-Treatment SCT Status = After SCT      604.46210   0.257  0.79761

Intercept                                                               ***
Treatment = MK-3475 200 mg                                              **
Post-Treatment SCT Status = After SCT
Treatment = MK-3475 200 mg & Post-Treatment SCT Status = After SCT
---
Signif. codes:  0 â€˜***â€™ 0.001 â€˜**â€™ 0.01 â€˜*â€™ 0.05 â€˜.â€™ 0.1 â€˜ â€™ 1
```

Indicated from the ANOVA summary table above, all of the factors included in the final model, (P-value:) are associated with UK UTILITY VALUE

Test description (Please document how the test is conducted, as well)	Expected result of the test	Result of test in the present model
Pre-analysis calculations		
Does the technology (drug/device, etc.) acquisition costs increase with higher prices?	Yes	Yes, increasing the unit cost of pembrolizumab in the References tab increases treatment costs as expected.
Does the drug acquisition cost increase for higher weight or body surface area?	Yes	Yes, for the few drugs in the model that have weight-based rather than fixed dosing.
Does the probability of an event, derived from an odds ratio (OR)/ relative risk (RR) / hazard ratio (HR) and baseline probability, increase with higher OR/RR/HR?	Yes	Yes, for OS curves that are calculated using HRs (i.e., in the SOC arm), probabilities of death increase when increasing the associated HRs within the Efficacy Parameters tab.
In a partitioned survival model, does the progression free survival curve or the time on treatment curve cross the overall survival curve?	No	This partitioned survival model does not use PFS curves. However, the model does enforce that ToT is always below modeled OS in each arm in each cycle (see column G of each trace_treatment tab).
If survival parametric distributions are used in the extrapolations or time-to-event calculations, can the formulae used for the Weibull (generalized gamma) distribution generate the values obtained from the exponential (the Weibull or Gamma) distribution(s) after replacing/transforming some of the parameters?	Yes	Yes, the formula for the Weibull distribution gave the same hazard rates as the formula for the exponential distribution when the Weibull ln(scale) parameter was set to -ln(exponential rate) and the ln(shape) parameter was set to 0. The formula for the generalized gamma distribution gave the same hazard rates as the formula for the Weibull when the mu parameter was set to ln(scale), the sigma parameter was set to -ln(shape), and the Q parameter was set to 1.
For the treatment effect inputs, if the model uses outputs from WINBUGs, are the OR, HR and RR values all within plausible ranges? (should be all non-negative and the average of these WINBUGs outputs should give the mean treatment effect)	Yes	Not applicable.
Event-state calculations		
Calculate the sum of the number of patients at each health state	Should add up to the cohort size	The health state residencies sum to 1 in each model cycle, as demonstrated in the trace_treatment tabs of the model.
Check if all probabilities and number of patients in a state are greater than or equal to zero	Yes	Yes, all health state residencies are ≥0, demonstrated in the trace_treatment tabs
Check if all probabilities are smaller than or equal to one	Yes	Yes, all probabilities of death ≤1, demonstrated in the effectiveness_treatment tabs (columns DO:DQ).
Compare the number of dead (or any absorbing state) patients in a period with the number of dead (or any absorbing state) patients in the previous periods?	Should be larger	Yes, the modeled OS curve in each arm is strictly decreasing from cycle to cycle, as illustrated by the OS figures presented on the Efficacy_Selection tab.
In case of lifetime horizon, check if all patients are dead at the end of the time horizon	Yes	Yes
Discrete event simulation specific: sample one of the "time to event" types used in the simulation from the specified distribution. Plot the samples and compare the mean and the variance from the sample	Sample mean and variance & the simulation outputs should reflect the distribution it is sampled from.	Not applicable.
Set all utilities to one Set all utilities to zero	The QALYs accumulated at a given time would be the same as the life years accumulated at that time No utilities will be accumulated in the model	Yes, this test was performed excluding AE-related, age-related, and SCT-related disutility, setting all health state utilities to 1 (or 0) (including the utility in cell AM6

		of each trace_treatment tab), and setting discounting to 0% (since LYs are not discounted in the model).
Decrease all state utilities simultaneously (but keep event-based utility decrements constant)	Lower utilities will be accumulated each time	Yes
Set all costs to zero	No costs will be accumulated in the model at any time	Yes, setting all cost inputs in the References tab to zero yields zero costs in the All Results tabs
Put mortality rates to 0	Patients never die	Yes, this test was performed by changing the following cells: ▪ Column C in the Life Tables tab set to 0 ▪ Column DP of each effectiveness_treatment tab to 0 ▪ Column DM of each effectiveness_treatment tab to 1
Put mortality rate extremely high	Patients die in the first few cycles	Yes, this test was performed by setting annual national mortality rates (column C in the Life Tables tab) to 100 (*note:* these are rates, not probabilities, so they can take on any number >=0), and setting column DM of each effectiveness_treatment tab to 0 after cycle 1.
Set the effectiveness, utility and safety related model inputs for all treatment options equal	Same life years and QALYs should be accumulated for all treatment at any time	Yes, this test and the next test below were performed by changing the input values summarized in the following cells of trace_treatment2 to instead summarize the same inputs as in trace_treatment1. The two trace_treatment tabs then produced identical costs, QALYs, and LYs in all cycles: ▪ C5, B7, columns E, F, and I:K, and AD6:BN8
In addition to the inputs above, set cost related model inputs for all treatment options equal	Same costs, life years and QALYs should be accumulated for all treatment at any time	Yes, see description above.
Change around the effectiveness, utility and safety related model inputs between two treatment options	Accumulated life years and QALYs in the model at any time should be also reversed	The tests above encapsulate this test.
Check if the number of alive patients estimate at any cycle is in line with general population life table statistics	At any given age, the % alive should be lower or equal in comparison to the general population estimate	Yes, OS in both model arms (column F of each trace_treatment tab) is always lower than OS in a general population cohort with the same starting age and gender distribution (column K of the Life Table tab).
Check if the QALY estimate at any cycle is in line with general population utility estimates	At any given age, the utility assigned in the model should be lower or equal in comparison to the general population estimate	Yes, age-related disutility is applied based on the model cohort's age in each cycle.
Set the inflation rate of the previous year higher	The costs (which are based on a reference from previous years) assigned at each time will be higher	Not applicable, inflation adjustment of cost inputs (where needed) was performed outside of the model.
Calculate the sum of all ingoing and outgoing transition probabilities	Both should be one	Not applicable, as the present model is a partitioned survival model rather than a Markov model.
Calculate the number of patients entering and leaving a tunnel state throughout the time horizon	Numbers entering = Numbers leaving	Not applicable, there are no tunnel states in the present model.
Check if the time conversions for probabilities were conducted correctly.	Yes	Yes. Time was measured in weeks throughout all survival analyses performed to estimate parameter estimates and Kaplan-Meier curves in the model, which

		also aligned with the 1-week cycle length used in the Excel model.
Decision tree specific: calculate the sum of the expected probabilities of the terminal nodes	Should sum up to one	Not applicable.
Patient-level model specific: check if common random numbers are maintained for sampling for the treatment arms?	Yes	Not applicable.
Patient-level model specific: check if correlation in patient characteristics is taken into account when determining starting population?	Yes	Not applicable.
Increase the treatment acquisition cost	Costs accumulated at a given time will increase during the period when the treatment is administered	Yes, increasing the unit drug price of pembrolizumab in the References tab increases treatment acquisition costs.
Population model specific: set the mortality and incidence rates to zero	Prevalence should be constant in time	Not applicable, prevalence is not an output of this model.
Result calculations		
Check the incremental life years and QALYs gained results. Are they in line with the comparative clinical effectiveness evidence of the treatments involved?	If a treatment is more effective, it generally results in positive incremental LYs and QALYs in comparison with the less effective treatments	Yes
Check the incremental cost results. Are they in line with the treatment costs?	If a treatment is more expensive, and if it does not have much effect on other costs, it generally results in positive incremental costs.	Yes, the incremental cost results have face validity based on the higher drug costs and the cost offsets (e.g., lower terminal care and subsequent treatment costs) associated with pembrolizumab relative to the SOC arm. The higher health state costs for pembrolizumab vs. SOC is expected due to longer OS and higher stem cell transplantation rate in the pembrolizumab arm.
Total life years > total quality adjusted life years	Yes	Yes
Undiscounted results > discounted results	Yes	Yes
Divide undiscounted total QALYs by undiscounted life years.	This value should be within the outer ranges (maximum and minimum) of the all utility value inputs.	Yes, for both model arms, this ratio is within the range of all health state utilities used in each arm.
Subgroup analysis results: How do the outcomes change if the characteristics of the baseline change?	Better outcomes for better baseline health conditions and worse outcomes for worse health conditions are expected.	Not applicable.
Could you generate all the results in the report from the model (including the uncertainty analysis results)?	Yes	Yes, all submitted results are replicable using the model.
Does the total life years, QALYs and costs decrease if a shorter time horizon is selected?	Yes	Yes
Is the reporting and contextualization of the incremental results correct?	The use of the terms such as: "dominant"/ "dominated"/ "extendedly dominated"/ "cost-effective" etc. should be in line with the results. In the incremental analysis table involving multiple treatments, ICERs should be calculated against the next non-dominated treatment.	Yes, ICERs are reported correctly based on which CE quadrant the new intervention falls in relative to the comparator. Extended dominance is not applicable in the present model, as there are only 2 model arms.
Are the reported ICERs in the fully incremental analysis non-decreasing?	Yes	Not applicable, there are only two treatment arms in this model.

If disentangled results are presented, do they sum up to the total results? (e.g. different cost types sum up to the total costs estimate)	Yes	Yes, category-specific costs correctly aggregate to total costs, as presented on the All Results tab.
Check if half cycle correction is implemented correctly (total life years with half cycle correction should be lower than without)	The half cycle correction implementation should be error free. Also check if it should be applied for all costs, for instance if a treatment is administered at the start of a cycle, half cycle correction might be unnecessary.	Yes, this can be seen by changing the dropdown menu on the Model Settings tab for "within cycle correction" from Yes to No. It was not necessary to apply half-cycle correction to drug costs (which are incurred at scheduled dosing intervals starting at time 0) or AE costs, AE disutility, and SCT procedure disutility (which are applied in the first cycle as lump sums).
Check the discounted value of costs/QALYs after 2 years	Discounted value=undiscounted/$(1+r)^2$	Yes, this was checked by comparing the undiscounted vs. discounted cost and QALY columns in cycle 119 (104 weeks).
Set discount rates to zero	The discounted and undiscounted results should be the same	Yes, changing the discounting dropdown menu on the Model Settings tab to "no" gives the same results as setting this dropdown to "yes" and setting both discount rates to 0%.
Set mortality rate to zero	The undiscounted total life years per patient should be equal to the length of the time horizon	Yes, this test was performed by changing the following cells: ▪ Column C in the Life Tables tab set to 0 ▪ Column DP of each effectiveness_treatment tab to 0 ▪ Column DM of each effectiveness_treatment tab to 1
Put the consequence of adverse event/discontinuation to zero. (zero costs and zero mortality/utility decrements)	The results would be the same as the results when AE rate is set to zero.	Yes, this was tested via the Safety tab.
Divide total undiscounted treatment acquisition costs by the average duration on treatment.	This should be similar to treatment related unit acquisition costs	Yes, the drug acquisition costs in D14 of the All Results tab roughly equals the mean number of Q3W administrations in KEYNOTE-087 (approximated as the sum of weekly time points in the ToT curve divided by 3) multiplied by the unit cost of a 200 mg dosage.
Set discount rates to a higher value	Total discounted results should decrease	Yes, total costs and total QALYs increase when increasing the discount rates.
Set discount rates of costs/effects to an extremely high value	Total discounted results should be more or less the same as the discounted results accrued in the first cycles	Yes, when discounting is set to an extremely high value, total costs and total QALYs show little change when the time horizon is reduced from 40 years to 1 year.
Put adverse event/discontinuation rates to zero and then to extremely high level.	Less costs higher QALYS/LYs when adverse event rates are 0, higher costs and lower QALYS/LYs when AE rates are extreme	Not applicable, time on treatment for pembrolizumab is based on the observed Kaplan-Meier curve for ToT
Double the difference in efficacy and safety between new intervention and comparator and report the incremental results.	Approximately twice of the incremental effect results of the base case. If this is not the case : report and explain the underlying reason/ mechanism	Not applicable, the efficacy-related parameter estimates in this model are not expressed in terms of difference vs. the comparator.
Do the same for a scenario in which the difference in efficacy and safety is halved.	Approximately halve of the incremental effect results of the base case. If this is not the case : report and explain the underlying reason/ mechanism	Not applicable, the efficacy-related parameter estimates in this model are not expressed in terms of difference vs. the comparator.

Uncertainty analysis calculations		
Are all parameters subject to uncertainty included in the one-way sensitivity analysis (OWSA)? Check if the OWSA includes any parameters associated with joint uncertainty (e.g. parts of a utility regression equation, survival curves with multiple parameters).	Yes No	Yes, the one-way, high/low deterministic sensitivity analysis includes all major parameters that are subject to uncertainty, except the OS parameters that are varied according to multivariate normal distributions in the PSA.
Are the upper and lower bounds used in the one-way sensitivity analysis used confidence intervals based on the statistical distribution assumed for that parameter? Are the resulting ICER, incremental costs/QALYs with upper and lower bound of a parameter plausible and in line with a priori expectations?	Yes Yes	In the one-way, high/low deterministic sensitivity analyses, parameters are varied between the upper and lower limits of their 95% CIs (based on the distributions used in the PSA).
Check that all parameters used in the sensitivity analysis have an appropriate associated distributions - upper and lower bounds should surround the deterministic value (i.e. Upper bound ≥ mean ≥ Lower bound) - standard error and not standard deviation used in sampling - Lognormal / gamma distribution for hazard ratios and costs/ resource use - Beta for utilities and proportions/probabilities - Dirichlet for multinomial - Multivariate normal for correlated inputs (e.g. survival curve or regression parameters) - Normal for other variables as long as samples don't violate requirement to remain positive when appropriate	Yes	Yes, the distributions specified for all parameters in the PSA Setup tab align with their allowable ranges, and SEs (not SDs) are used. Parameter estimates from multi-parameter OS models are varied according to multivariate normal distributions.
Check PSA output mean costs, QALYs and ICER compared to the deterministic results. Is there a large discrepancy?	No (in general)	No. The probabilistic ICER (based on the average of incremental costs / average of incremental QALYs across all PSA iterations) is similar to the deterministic ICER.
If you take new PSA runs from the excel model do you get similar results?	Yes	Yes, the PSA results are similar when the PSA is rerun following a change to the random seed. (The seed can be changed via the PSA Results tab.)
Is(are) the CEAC line(s) in line with the CE scatter plots and the efficient frontier?	Yes	Yes, the CEAC lines are consistent with the scatterplot (both are shown side-by-side on the PSA Results tab)
Does the PSA cloud demonstrate an unexpected behavior or has an unusual shape?	No	No. As shown on the PSA Results tab, the PSA cloud demonstrates a typical and logical shape.
Is the sum of all CEAC lines equal to 1 for all WTP values?	Yes	Yes, the two lines sum to 1 at all WTP values.
Are the explored scenario analyses provide a balanced view on the structural uncertainty? (i.e. not always looking at more optimistic scenarios)	Yes	Yes, conservative scenarios are explored, including a scenario that combines multiple pessimistic assumptions on efficacy, treatment costs, and utility.
Are the scenario analysis results plausible and in line with a priori expectations?	Yes	Yes, all scenario analysis results were checked for face validity.
Check the correlation between 2 PSA results (i.e. costs/QALYs under the SoC and costs/QALYs under the comparator)	Should be very low (very high) if different (same) random streams are used for different arms	Not applicable, the PSA uses different random numbers for all parameters included in the PSA.
If a certain seed is used for random number generation (or previously generated random numbers are used), check if they are they scattered evenly between 0-1 when they are plotted?	Yes	Yes, in the PSA Setup tab, cells I5:I102 display the random numbers from the last iteration of the last PSA run, based on the user-specified random seed in the same tab. The numbers are evenly distributed between

		0 and 1; for example, the formula =PERCENTILE.EXC(I5:I102,0.95) will resolve to a number that is close to 0.95.
Compare the mean of the parameter samples generated by the model against the point estimate for that parameter, use graphical methods to examine distributions, functions	The sample means and the point estimates will overlap, the graphs will be similar to the corresponding distribution functions (e.g. Normal, Gamma, etc.)	The submitted model does not store parameter values in each PSA iteration for efficiency purposes. However, this check was performed when programming the PSA for this model.
Check if sensitivity analyses include any parameters associated with methodological/ structural uncertainty (e.g. annual discount rates, time horizon).	No	These parameters are varied in scenario analyses.
Value of information analysis if applicable: Was this implemented correctly? Which types of analysis? Were aggregated parameters used? Which parameters are grouped together? Does it match the write-up's suggestions? Is EVPI larger than all individual EVPPI? Is EVPPI for a (group of) parameters larger than the EVSI of that (group) of parameter(s)? Are the results from EVPPI in line with OWSA or other parameter importance analysis (e.g. ANCOVA)?	Yes	Not applicable.
Did the electronic model pass the black-box tests of the previous verification stages in all PSA iterations and in all scenario analysis settings? (additional macro can be embedded to PSA code, which stops the PSA when an error such as negative transition probability, is detected)	Yes	Yes, the PSA Results tab will display errors if any of the PSA iterations led to non-calculating cost and QALY results.
OWSA=one-way sensitivity analysis; ICER = incremental cost-effectiveness ratio; PSA = probabilistic sensitivity analysis; WTP = willingness to pay; CE = cost-effectiveness; CEAC = cost-effectiveness acceptability curve; LY = life years; QALYs = Quality adjusted life years; OR = odds ratio; RR= relative risk; HR = hazard ratio		

NICE National Institute for Health and Care Excellence

Single Technology Appraisal

Guidance review following a period of managed access - Patient organisation submission

Pembrolizumab for treating relapsed or refractory classical Hodgkin lymphoma [Review of TA540] [ID5084]

Thank you for agreeing to give us your organisation's views on this treatment following a period of managed access. You can provide a unique perspective on conditions and their treatment that is not typically available from other sources.

PLEASE NOTE: You do not have to answer every question. Your organisations involvement in the managed access agreement for this treatment is likely to determine which questions you can answer.

To help you give your views, please use this questionnaire with **NICE's guide for patient organisations "completing an organisation submission following a period of Managed Access for Technology Appraisals or Highly Specialised Technologies".** Please contact pip@nice.org.uk if you have not received a copy with your invitation to participate.

Information on completing this submission

- Please do not embed documents (such as a PDF) in a submission because this may lead to the information being mislaid or make the submission unreadable
- We are committed to meeting the requirements of copyright legislation. If you intend to include **journal articles** in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs.
- Your response should not be longer than 20 pages.

NICE National Institute for Health and Care Excellence

This form has 8 sections

NICE National Institute for Health and Care Excellence

Section 1. About you

Table 1 Name, job, organisation

1. Your name	[redacted]
2. Name of organisation	Lymphoma Action
3. Job title or position	[redacted]
4a. Provide a brief description of the organisation. How many members does it have?	Lymphoma Action is a national charity, established in 1986, registered in England and Wales and in Scotland. We provide high quality information, advice and support to people affected by lymphoma – the 5th most common cancer in the UK. We also provide education, training and support to healthcare practitioners caring for lymphoma patients. In addition, we engage in policy and lobbying work at government level and within the National Health Service with the aim of improving the patient journey and experience of people affected by lymphoma. We are the only charity in the UK dedicated to lymphoma. Our mission is to make sure no one faces lymphoma alone. Lymphoma Action is not a membership organisation. We are funded from a variety of sources predominantly fundraising activity with some limited sponsorship and commercial activity. We have a policy for working with healthcare and pharmaceutical companies – those that provide products, drugs or services to patients on a commercial or profit-making basis. The total amount of financial support from healthcare companies will not exceed 20% of our total budgeted income for the financial year (this includes donations, gifts in kind, sponsorship etc) and a financial cap of £50,000 of support from individual healthcare companies per annum (excluding employee fundraising), unless approval to accept a higher amount is granted by the Board of Trustees.

	The policy and approach ensures that under no circumstances will these companies influence our strategic direction, activities or the content of the information we provide to people affected by lymphoma. https://lymphoma-action.org.uk/about-us-how-we-work-policies-and-terms-use/working-healthcare-and-pharmaceutical-companies
4b. Has the organisation received any funding from the company/companies of the treatment and/or comparator products in the last 12 months? [Relevant companies are listed in the appraisal stakeholder list which was provided to you when the appraisal started] **If so, please state the name of company, amount, and purpose of funding.**	Funding received in 2022: Merck & Co None 1. Beigene – none 2. AbbVie - £10,000 3. AstraZeneca - £11,000 4. Gilead - £10,000 5. Janssen - £12,500 6. Pfizer – 300 7. Roche – £26,000 (as of June 2022)
4c. Do you have any direct or indirect links with, or funding from, the tobacco industry?	No
5. How did you gather information about the experiences of patients and carers to include in your submission?	We spoke to members of our community to understand their experience of living with Hodgkin lymphoma. We received no responses from anyone who had, or was aware that they had, experienced this treatment via the MAA process

Section 2 Living with the condition and current treatment

Table 2 What it's like for patients, carers and families to live with the condition and current NHS treatment

6. What is it like to live with the condition? Consider the experience of living with the condition and the impact on daily life (physical and emotional health, ability to work, adaptations to your home, financial impact, relationships, and social life). For children, consider their ability to go to school, develop emotionally, form friendships and participate in school and social life. Is there any impact on their siblings?	Around 2,000 people develop classical Hodgkin lymphoma each year in the UK. It can develop at any age, but most people diagnosed are between the ages of 15 and 34, or over 60. Classical Hodgkin lymphoma affects slightly more males than females. It can affect people of any ethnic background. The most common symptoms are swollen lymph nodes, often in the neck, armpit or groin but they can be in the chest, causing breathlessness. Symptoms can vary depending on where the lymphoma is growing. Systemic symptoms are common, including fevers, night sweats, unexplained weight loss, fatigue, loss of appetite and severe itching. Classical Hodgkin lymphoma generally responds very well to treatment. Most people are cured, even if their lymphoma is advanced when it is diagnosed. Treatment for classical Hodgkin lymphoma usually involves chemotherapy, sometimes followed by radiotherapy. The exact treatment depends on the stage of disease and how it is affecting the patient. For those with relapsed / refractory classical Hodgkin lymphoma further treatment options are available, generally chemotherapy, often referred to as salvage chemotherapy. A good response to this treatment may result in the opportunity to have a stem cell transplant. Hodgkin lymphoma and its treatment significantly affect patients' quality of life. Just over half of patients report that symptoms and side effects of treatment negatively impact their social lives and the everyday activities they are able to do. Fatigue is the most commonly reported symptom, affecting around 3 in 4 people, and it can persist for many years. Patients report that this affects their work, physical activity and social activities. Fatigue, nausea and vomiting and infections are considered to be the most troublesome side effects. One patient who had had chemotherapy, radiotherapy, a stem cell transplant and nivolumab for Hodgkin lymphoma said, "The fatigue is the most difficult to manage over the long term – it may be from the lymphoma or the treatment. The fatigue and stress have often made it very difficult to contribute normally at work. I have no energy to do anything in the evening – my fatigue then can be overwhelming."

	Hodgkin lymphoma can also have a financial impact on patients and their families. One patient, who had been treated with ABVD and bretuximab vedotin (available second-line during the coronavirus pandemic), said, “I have not worked for almost 12 months, though fortunately the financial impact has been mitigated by insurance and a good company sick pay scheme.” The emotional impact of lymphoma is also considerable. Around a third of patients experience depression, anxiety, isolation and loss of self-esteem, with even more (>40%) reporting fear of lymphoma progression or relapse. Over a quarter of patients say they feel overwhelmed by managing their lymphoma and many fee they do not get enough emotional or financial support to help them. About half of patients with Hodgkin lymphoma report needing information on psychological support and counselling, with around 1 in 3 listing that access to support for their families would be beneficial. One patient with relapsed Hodgkin lymphoma said, “There were periods earlier when I felt death was very near, and that was a very difficult time emotionally. I have also had to deal with a lot of uncertainty, and although my condition is now stable, fear can overtake me when I experience even mild symptoms.” From a practical viewpoint, patients with Hodgkin lymphoma find the treatments and associated blood tests and waiting times a huge time commitment. Travel costs and transport logistics can also be an issue for patients who live some distance away from their treatment centre.
7. What do carers experience when caring for someone with the condition?	The impact of Hodgkin lymphoma extends beyond the patient to their carers and families. One patient said, “Having two small children, the impact on myself and my family has been huge.” Carers provide emotional support, practical support with transport, help with personal care, errands and household chores, and many also take responsibility for managing finances and healthcare appointments. They provide an essential role in supporting people affected by lymphoma, but this is a huge psychological and emotional burden. Almost all caregivers report feeling worried or anxious, and scared by the prospect of their loved ones’ lymphoma relapsing. One patient with relapsed Hodgkin lymphoma told us how stressful it was for their partner trying to manage their work around treatment and increased childcare responsibilities, and how their partner had really suffered emotionally.
8. What do patients and carers think of current treatments and care available on the NHS	Most people with Hodgkin lymphoma are treated with chemotherapy, sometimes followed by radiotherapy. High-dose chemotherapy regimens might be used. For relapsed or refractory Hodgkin lymphoma, salvage chemotherapy followed by stem cell transplant is the most common treatment option. Treatment is very intense and some people are not able to tolerate it. People who experience a subsequent relapse might be treated with more chemotherapy or targeted treatments such as brentuximab vedotin, nivolumab or

Please state how they help and what the limitations are.	pebrolizumab. At present, these less toxic options are only available for people who have either relapsed after a stem cell transplant or who are not able to have a stem cell transplant. One patient who had received multiple lines of treatment for Hodgkin lymphoma said, "I am grateful for the treatment I have received on the NHS, but I have found it inadequate on multiple occasions." In particular, the patient felt that more effective, better tolerated – and less risky – treatment options should be available earlier in the treatment pathway and that at many points in their pathway, the options available on the NHS were very limited. When they experienced a relapse after an autologous stem cell transplant, the patient resorted to private treatment to enable them to access a combination of brentuximab vedotin and nivolumab rather than undergo an allogeneic stem cell transplant on the NHS. Patients feel that current treatment options for relapsed or refractory Hodgkin lymphoma are difficult to cope with. Most patients experience significant side effects, such as fatigue, nausea, pain and hair loss, and many go on to develop late effects. One patient told us how treatment left them unable to care for their children – with emotional as well as physical consequences. Treatment has a long-lasting impact on physical and mental wellbeing. However, patients are grateful that treatment has given them another chance. One patient described how daunted they feel at the prospect of a stem cell transplant, which will be an inevitable part of their treatment once they achieve a remission. Patients feel that the high response rate to pembrolizumab, combined with its tolerability profile, offer a significant advantage over many other treatments. Patients feel that pembrolizumab has a more favourable side effect profile than most other treatments for relapsed and refractory Hodgkin lymphoma, which would have a significant impact on their quality of life. They also feel that, as an outpatient treatment with minimal pre-meds required, it is more convenient and less time consuming than many other options. It is also likely to have a much lower impact on family life, since it does not require prolonged hospital stays and the less troublesome side effects allow patients to carry on with day-to-day activities. Two patients who had been treated with a similar checkpoint inhibitor experienced far less onerous side effects with the checkpoint inhibitor than with the radiotherapy, chemotherapy or stem cell transplant they had previously had. The targeted treatment allowed them to carry on with a more 'normal' family life. One commented, "I don't know how I would have managed my son's school years on those other treatments."
9. Considering all treatments available to patients are	Patients feel there is a definite unmet need for an effective, less demanding treatment with fewer side effects and will therefore allow a better quality of life. One patient commented, "Many of the options after failure of initial treatment do not have especially high success rates. This is not very reassuring."

there any unmet needs for patients with this condition? If yes please state what these are	The three most important factors patients with lymphoma rate in a treatment are, in order: effectiveness (in terms of improved survival or response rates); quality of life; and tolerability.

Section 3 Experience during the managed access agreement (MAA)

Table 3 Experience, advantages and disadvantages during the MAA

10. What are patients' and carers' experience of accessing and having the treatment? • Please refer to the MAA re-evaluation patient submission guide	We received no feedback which enables us to answer these specific questions.
11. What do patients and carers think are the advantages of the treatment? Please refer to the MAA re-evaluation patient submission guide	
12. What do patients or carers think are the disadvantages of the treatment? Please refer to the MAA re-evaluation patient submission guide	

13. What place do you think this treatment has in future NHS treatment and care for the condition? Consider how this treatment has impacted patients and how it fits alongside other treatments and care pathway.	

Section 4 Patients views on assessments used during the MAA

Table 4 Measurements, tests and assessments

14. Results from tests and assessments are used to help reduce uncertainty about the effectiveness of treatment. **How well do you think these tests and assessments worked in measuring the effectiveness of the treatment?**	
15. Were there any tests or assessments that were difficult or unhelpful from a patient's or carer's perspective?	

16. Do patients and carers consider that their experiences (clinical, physical, emotional and psychological) were captured adequately in the MAA tests and assessments? If not please explain what was missing.	
17. What outcomes do you think have not been assessed or captured in the MAA data? Please tell us why	

Section 5 Patient population

Table 5 Groups who may benefit and those who declined treatment

18. Are there any groups of patients who might benefit more or less from the treatment than others? If so, please describe them and explain why.	One patient felt that people who found it hard to tolerate chemotherapy side effects might in particular benefit from Pembrolizumab.

19. Were there people who met the MAA eligibility criteria who decided not to start treatment? Please state if known the proportion of eligible patients who did not start the treatment and any reasons for this.	No response.

Section 6 Equality

20. Are there any potential equality issues that that should be taken into account when considering this condition and the treatment? See NICE's equality scheme for more details.

Section 7 Other issues

21. Are there any other issues that you would like the committee to consider?

Section 8 Key messages

In up to 5 sentences, please summarise the key messages of your statement:

- Click or tap here to enter text.
- Click or tap here to enter text.
- Click or tap here to enter text.
- Click or tap here to enter text.
- Click or tap here to enter text.

Thank you for your time.

Please log in to your NICE Docs account to upload your completed statement, declaration of interest form and consent form.

Your privacy

The information that you provide on this form will be used to contact you about the topic above.

☐ **Please tick this box** if you would like to receive information about other NICE topics.

For more information about how we process your personal data please see NICE's privacy notice.

NICE National Institute for Health and Care Excellence

Single Technology Appraisal

Pembrolizumab for treating relapsed or refractory classical Hodgkin's lymphoma after brentuximab vedotin [review of TA540] ID5084

Professional organisation submission

Thank you for agreeing to give us your organisation's views on this technology and its possible use in the NHS.

You can provide a unique perspective on the technology in the context of current clinical practice that is not typically available from the published literature.

To help you give your views, please use this questionnaire. You do not have to answer every question – they are prompts to guide you. The text boxes will expand as you type.

Information on completing this submission

- Please do not embed documents (such as a PDF) in a submission because this may lead to the information being mislaid or make the submission unreadable
- We are committed to meeting the requirements of copyright legislation. If you intend to include **journal articles** in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs.
- Your response should not be longer than 13 pages.

About you

1. Your name	[redacted]
2. Name of organisation	Royal College of Pathologists
3. Job title or position	Consultant Haematologist [redacted]
4. Are you (please select Yes or No):	An employee or representative of a healthcare professional organisation that represents clinicians? Yes A specialist in the treatment of people with this condition? Yes A specialist in the clinical evidence base for this condition or technology? No Other (please specify):
5a. Brief description of the organisation (including who funds it).	The Royal College of Pathologists is a professional membership organisation with charitable status, concerned with all matters relating to the science and practice of pathology. It is a body of its Fellows, Affiliates, and trainees, supported by the staff who are based at the College's London offices.
5b. Has the organisation received any funding from the manufacturer(s) of the technology and/or comparator products in the last 12 months? [Relevant manufacturers are listed in the appraisal matrix.] **If so, please state the name of manufacturer, amount, and purpose of funding.**	No
5c. Do you have any direct or indirect links with, or funding from, the tobacco industry?	No

The aim of treatment for this condition

6. What is the main aim of treatment? (For example, to stop progression, to improve mobility, to cure the condition, or prevent progression or disability.)	To cure the condition.
7. What do you consider a clinically significant treatment response? (For example, a reduction in tumour size by x cm, or a reduction in disease activity by a certain amount.)	Eradication of the disease. Anything less than this is a treatment failure.
8. In your view, is there an unmet need for patients and healthcare professionals in this condition?	1) Patients who have failed first line therapy and subsequent salvage therapy/autologous transplant/brentuximab vedotin currently have available nivolumab (NICE TA462 2017). Pembrolizumab works in the same way but may have a different side-effect profile in individual patients. This unmet need is therefore to provide an alternative in the uncommon patient unable to tolerate nivolumab. 2) NICE TA462 recommends the use of nivolumab only in patients who have relapsed after autologous stem cell transplant. There is an unmet need for anti-PD1 therapy (nivolumab or pembrolizumab) in patients who are not suitable for stem cell transplant because of disease progression despite salvage chemotherapy or brentuximab vedotin. The use of an anti-PD1 therapy in this situation would be as a bridge to transplant.

What is the expected place of the technology in current practice?

9. How is the condition currently treated in the NHS?	At presentation: initial chemotherapy +/- radiotherapy is curative in the majority of patients. If primary refractory or later relapse: alternative (salvage) chemotherapy followed by autologous transplant in responding patients. If unresponsive to salvage chemotherapy: brentuximab vedotin (alone or in combination with bendamustine) as a bridge to transplantation. If unresponsive to brentuximab (alone or in combination with bendamustine): nivolumab as a bridge to allogeneic transplantation.
9a. Are any clinical guidelines used in the treatment of the condition, and if so, which?	A commonly used guideline: Guideline for the first-line management of Classical Hodgkin Lymphoma — A British Society for Haematology guideline. British Journal of Haematology 2022; 197: 558-572.
9b. 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.)	It is well-defined in the great majority of cases.
9c. What impact would the technology have on the current pathway of care?	There is an unmet need for anti-PD1 therapy in primary refractory patients who have progressive disease despite salvage therapy and are unsuitable for stem cell transplant, as a bridge to transplant. It would also provide an alternative to nivolumab in the uncommon patient who suffers debilitating side-effects and who could receive pembrolizumab as an alternative.
10. Will the technology be used (or is it already used) in the same way as current care in NHS clinical practice?	Yes.

10a. How does healthcare resource use differ between the technology and current care?	No difference.
10b. In what clinical setting should the technology be used? (For example, primary or secondary care, specialist clinics.)	Specialist clinics.
10c. What investment is needed to introduce the technology? (For example, for facilities, equipment, or training.)	None.
11. Do you expect the technology to provide clinically meaningful benefits compared with current care?	Yes, in a) patients who are unsuitable for stem cell transplant and need anti-PD1 therapy as a bridge to transplant and b) in patients who suffer debilitating side-effects with nivolumab and who may tolerate pembrolizumab.
11a. Do you expect the technology to increase length of life more than current care?	No.
11b. Do you expect the technology to increase health-related quality of life more than current care?	No.
12. Are there any groups of people for whom the technology would be more or less effective (or	No.

appropriate) than the general population?	

The use of the technology

13. Will the technology be easier or more difficult to use for patients or healthcare professionals than current care? Are there any practical implications for its use (for example, any concomitant treatments needed, additional clinical requirements, factors affecting patient acceptability or ease of use or additional tests or monitoring needed.)	No.
14. Will any rules (informal or formal) be used to start or stop treatment with the technology? Do these include any additional testing?	No.
15. Do you consider that the use of the technology will result in any	No.

substantial health-related benefits that are unlikely to be included in the quality-adjusted life year (QALY) calculation?	
16. Do you consider the technology to be innovative in its potential to make a significant and substantial impact on health-related benefits and how might it improve the way that current need is met?	No.
16a. Is the technology a 'step-change' in the management of the condition?	No.
16b. Does the use of the technology address any particular unmet need of the patient population?	Only in a) patients who require anti-PD1 therapy as a bridge to transplant and b) the uncommon patient who is unable to tolerate nivolumab due to side-effects.
17. How do any side effects or adverse effects of the technology affect the management of the condition and the patient's quality of life?	The use of pembrolizumab as a bridge to transplantation in a patient who has been unable to tolerate nivolumab may clearly affect the management of that patient's condition.

Sources of evidence

18. Do the clinical trials on the technology reflect current UK clinical practice?	Yes.
18a. If not, how could the results be extrapolated to the UK setting?	N/A
18b. What, in your view, are the most important outcomes, and were they measured in the trials?	Overall response rate, duration of response, and adverse events causing treatment discontinuation. The trials have shown an overall response rate to pembrolizumab of approximately 70%, a median duration of response of greater than 6 months, and significant adverse events sufficient to discontinue therapy of approximately 15%.
18c. If surrogate outcome measures were used, do they adequately predict long-term clinical outcomes?	N/A
18d. Are there any adverse effects that were not apparent in clinical trials but have come to light subsequently?	No.
19. Are you aware of any relevant evidence that might not be found by a systematic review of the trial evidence?	No.

20. Are you aware of any new evidence for the comparator treatment(s) since the publication of NICE technology appraisal guidance [TA462]?	No.
21. How do data on real-world experience compare with the trial data?	Very similar outcomes from both settings (trial and non-trial data).

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.	N/A.

Key messages

24. In up to 5 bullet points, please summarise the key messages of your submission.	• Anti-PD1 therapy is an important treatment to have available in the management of Hodgkin Lymphoma patients who have failed first-line therapy, salvage therapy, and brentuximab vedotin. • Nivolumab has been appraised previously (NICE TA462 2017). It is currently restricted to patients who have failed stem cell transplant. • Some patients, due to progressive chemo-refractory disease, need anti-PD1 therapy as a bridge to transplant. This can be with either nivolumab or pembrolizumab. • The side-effect profile of nivolumab and pembrolizumab are somewhat different, and so patients experiencing debilitating complications with either may be treated with the other to potentially overcome this.

Thank you for your time.

Please log in to your NICE Docs account to upload your completed submission.

Your privacy

The information that you provide on this form will be used to contact you about the topic above.

Please select YES if you would like to receive information about other NICE topics - YES or NO

For more information about how we process your personal data please see our privacy notice.

NDRS
NATIONAL DISEASE REGISTRATION SERVICE

NHS England

Pembrolizumab for treating relapsed or refractory classical Hodgkin lymphoma – data review

About the NDRS

The National Disease Registration Service (NDRS) is part of NHS England. Its purpose is to collect, collate and analyse data on patients with cancer, congenital anomalies, and rare diseases. It provides robust surveillance to monitor and detect changes in health and disease in the population. NDRS is a vital resource that helps researchers, healthcare professionals and policy makers make decisions about NHS services and the treatments people receive.

The NDRS includes:

- the National Cancer Registration and Analysis Service (NCRAS) and
- the National Congenital Anomaly and Rare Disease Registration Service (NCARDRS)

Healthcare professionals, researchers and policy makers use data to better understand population health and disease. The data is provided by patients and collected by the NHS as part of their care and support. The NDRS uses the data to help:

- understand cancer, rare diseases, and congenital anomalies
- improve diagnosis
- plan NHS services
- improve treatment
- evaluate policy
- improve genetic counselling

NDRS
NATIONAL DISEASE REGISTRATION SERVICE

National Disease Registration Service
The Leeds Government Hub
7&8 Wellington Place
Leeds
LS1 4AP

For queries relating to this document, please contact:
NDRSenquiries@nhs.net

NHS England

Contents

1. Executive summary

Introduction

The National Institute for Health and Care Excellence (NICE) appraised the clinical and cost effectiveness of pembrolizumab for treating relapsed or refractory classical Hodgkin lymphoma. The appraisal committee highlighted clinical uncertainty around estimates of overall survival (OS) in the evidence submission. As a result, they recommended the commissioning of pembrolizumab through the Cancer Drugs Fund (CDF) to allow a period of managed access, supported by additional data collection to answer the clinical uncertainty.

NHS England have evaluated the real-world treatment effectiveness of pembrolizumab in the CDF population, during the managed access period. This report presents the results of the use of pembrolizumab in clinical practice in England, using the routinely collected Systemic Anti-Cancer Therapy (SACT) dataset.

This report, and the data presented, demonstrate the potential within the English health system to collect real-world data to inform decision-making about patient access to cancer treatments via the CDF. The opportunity to collect real-world data enables patients to access promising new treatments much earlier than might otherwise be the case, whilst further evidence is collected to address clinical uncertainty.

The collection and follow up of real-world SACT data for patients treated through the CDF in England has resulted in analysis being carried out on 97% of patients and 89% of patient outcomes reported in the SACT dataset. NHS England are committed to providing world first, high-quality real-world data on CDF cancer treatments to be appraised alongside the outcome data from the relevant clinical trials.

Methods

The NHS England Blueteq® system was used to provide a reference list of all patients with an application for pembrolizumab for treating relapsed or refractory classical Hodgkin lymphoma in the CDF. Patient NHS numbers were used to link Blueteq applications to NDRS' routinely collected SACT data to provide SACT treatment history.

Between 25 Jul 2018 and 30 September 2022, 242 applications for pembrolizumab were identified in the Blueteq system. Following appropriate exclusions (see Figures 1 and 2), 215 unique patients who received treatment were included in these analyses. All patients were traced to obtain their vital status using the personal demographics service (PDS)[1].

Results

215 /220 (98%) unique patients with CDF applications were reported in the SACT dataset and were included in the final cohort.

Median treatment duration was 5.0 months [95% CI: 4.3, 6.2] (152 days). 45% of patients were still receiving treatment at 6 months [95% CI: 38%, 51%], 24% of patients were still receiving treatment at 12 months [95% CI: 18%, 31%], 16% of patients were still receiving treatment at 18 months [95% CI: 11%, 22%] and 10% of patients were still receiving treatment at 24 months [95% CI: 6%, 16%].

At data cut off, 83% (N=178) of patients were identified as no longer being on treatment. Of these 178 patients:

- 38% (N=68) of patients stopped treatment due to disease progression
- 30% (N=54) of patients completed treatment as prescribed
- 12% (N=21) of patients died not on treatment
- 6% (N=11) of patients did not have a treatment record in SACT in at least three months and are assumed to have completed treatment
- 4% (N=8) of patients stopped treatment due to acute toxicity
- 3% (N=6) of patients were treated palliatively and did benefit from the treatment they received
- 3% (N=6) of patients chose to end their treatment
- 1% (N=2) of patients were treated palliatively and did not benefit from the treatment they received
- 1% (N=2) of patients died on treatment

The median OS was not reached. OS at 6 months was 88% [95% CI: 83%, 92%], 12 months OS was 82% [95% CI: 76%, 87%], OS at 18 months was 75% [95% CI: 68%, 80%], OS at 24 months was 68% [95% CI: 61%, 75%], OS at 36 months was 56% [95% CI: 47%, 64%] and OS at 48 months was 55% [95% CI: 46%, 63%].

A treatment duration sensitivity analysis was conducted for a cohort with at least 6 months' data follow-up in the SACT dataset. Results were consistent with the full analysis cohort.

Conclusion

This report analysed SACT real-world data for patients treated with pembrolizumab for treating relapsed or refractory classical Hodgkin lymphoma in the CDF. It evaluated treatment duration, OS and treatment outcomes for all patients treated with pembrolizumab for this indication.

Introduction

Hodgkin lymphoma (ICD-10: C81) accounts for 1% of all cancer diagnoses in England. In 2020, 1,721 patients were diagnosed with Hodgkin lymphoma (males 976, females 745)[2].

- pembrolizumab is not recommended for treating relapsed or refractory classical Hodgkin lymphoma in adults who have had autologous stem cell transplant and brentuximab vedotin.
- pembrolizumab is recommended for use within the Cancer Drugs Fund as an option for treating relapsed or refractory classical Hodgkin lymphoma in adults who have had brentuximab vedotin and cannot have autologous stem cell transplant, only if:
    - pembrolizumab is stopped after 2 years of treatment or earlier if the person has a stem cell transplant or the disease progresses and
    - the conditions in the managed access agreement for pembrolizumab are followed[3]

2. Background to this report

Using routinely collected data to support effective patient care

High quality and timely cancer data underpin NHS England's ambitions of monitoring cancer care and outcomes across the patient pathway. NHS England produces routine outcome reports on patients receiving treatments funded through the Cancer Drugs Fund (CDF) during a period of managed access using the Systemic Anti-Cancer Therapy (SACT) data collected by the National Cancer Registration and Analysis Service (NDRS).

The CDF is a source of funding for cancer drugs in England[4]. From 29 July 2016 NHS England implemented a new approach to the appraisal of drugs funded by the CDF. The new CDF operates as a managed access scheme that provides patients with earlier access to new and promising treatments where there is uncertainty as to their clinical effectiveness. During this period of managed access, ongoing data collection is used to answer the clinical uncertainties raised by the NICE committee and inform drug reappraisal at the end of the CDF funding period[5].

NHS England analyse data derived from patient-level information collected in the NHS, as part of the care and support of cancer patients. The data is collated, maintained, quality-assured and analysed by the NDRS.

NICE Appraisal Committee review of pembrolizumab for treating Hodgkin lymphoma [TA540]

The NICE Appraisal Committee reviewed the clinical and cost effectiveness of pembrolizumab (Merck Sharp & Dohme Ltd) in treating relapsed or refractory classical Hodgkin lymphoma [TA540] and published guidance for this indication in September 2018[6].

Due to the clinical uncertainties identified by the committee and outlined below, the committee recommended the commissioning of pembrolizumab for the treatment of relapsed or refractory classical Hodgkin lymphoma through the CDF for a period of 50 months, from July 2018 to September 2022. The drug will be funded through the CDF until NICE publish their final guidance.

During the CDF funding period, results from SACT are likely to answer the main clinical uncertainties raised by the NICE committee.

As part of the guidance review, Merck Sharp & Dohme Ltd will provide supportive data from the KEYNOTE-087 clinical trial, the phase II multi-centre, single-arm, multi-cohort, non-randomised clinical trial[4].

Analysis of the SACT dataset provides information on real-world treatment patterns and outcomes for pembrolizumab for the treatment of relapsed or refractory classical Hodgkin lymphoma in England, during the CDF funding period. This acts as a primary source of information alongside the results of the KEYNOTE-087 clinical trial[7].

The committee identified the key areas of uncertainty below for re-appraisal at the end of the CDF data collection;

- time at which point a stem cell transplant (SCT) occurs in patients with Hodgkin lymphoma (time from first pembrolizumab treatment to SCT); and
- proportion of patients who receive a SCT; and
- overall survival

Treatment duration was not an area of clinical uncertainty but has been included in this report.

Approach

Upon entry to the CDF, representatives from NHS England, NICE and the company (Marck Sharp & Dohme Ltd) formed a working group to agree the Data Collection Agreement (DCA)[6]. The DCA set out the real-world data to be collected and analysed to support the NICE re-appraisal of pembrolizumab. It also detailed the eligibility criteria for patient access to pembrolizumab through the CDF, and CDF entry and exit dates.

This report includes patients with approved CDF applications for pembrolizumab, approved through Blueteq® and followed up in the SACT dataset collected by NDRS in NHS England.

3. Methods

CDF applications – identification of the cohort of interest

NHS England collects applications for CDF treatments through their online prior approval system (Blueteq®). The Blueteq application form captures essential baseline demographic and clinical characteristics of patients needed for CDF evaluation purposes. Where appropriate, Blueteq data are included in this report.

Consultants must complete a Blueteq application form for every patient receiving a CDF funded treatment. As part of the application form, consultants must confirm that a patient satisfies all clinical eligibility criteria to commence treatment. NDRS has access to the Blueteq database and key data items such as NHS number, primary diagnosis and drug information of all patients with an approved CDF application (which therefore met the treatment eligibility criteria).

The lawfulness of this processing is covered under Article 6(1)(e) of the United Kingdom (UK) General Data Protection Regulations (GDPR) (processing is necessary for the performance of a task carried out in the public interest or in the exercise of official authority vested in the controller). NHS England, through the National Disease Registration Service (NDRS), does have statutory authority to process confidential patient information (without prior patient consent) afforded through the National Disease Registries (NDRS) Directions 2021 issued to it by the Secretary of State for Health and Social Care, and has issued the NDRS Data Provision Notice under section 259 of the Health and Social Care Act 2012 regarding collection of the Blueteq data from NHS England.

NDRS in NHS England collates data on all SACT prescribed drugs by NHS organisations in England, irrespective of the funding mechanism. The Blueteq extract is therefore essential to identify the cohort of patients whose treatment was funded by the CDF.

Pembrolizumab clinical treatment criteria

- the patient is an adult and has histologically documented classical Hodgkin lymphoma
- the patient has failed at least 2 lines of chemotherapy and also treatment with brentuximab vedotin
- the patient has not received stem cell transplantation of any kind
- the patient is currently ineligible for stem cell transplantation
- the patient is either a candidate for future stem cell transplantation if there is sufficient benefit of treatment with pembrolizumab or is not a candidate for stem cell transplantation however good the response to pembrolizumab may be
- the patient has an Eastern Cooperative Oncology Group (ECOG) performance status (PS) of 0 or 1
- the patient has not received prior treatment with an anti-PD-1, anti-PD-L1, anti-PD-L2, anti-CD137, or anti-cytotoxic T-lymphocyte-associated antigen-4 (CTLA-4) antibody
- pembrolizumab is being given as monotherapy and will commence at a fixed dose of 200mg per infusion
- a formal medical review as to whether treatment with pembrolizumab should continue or not will be scheduled to occur at least by the end of the third cycle of treatment
- the patient will be treated until stem cell transplantation occurs or loss of clinical benefit or excessive toxicity or patient choice to discontinue treatment, whichever is the sooner
- the patient will receive a maximum treatment duration with pembrolizumab of 2 years (or 35 x 3-weekly cycles of pembrolizumab or its equivalent if 6-weekly pembrolizumab monotherapy dosing is used)
- treatment breaks of up to 12 weeks beyond the expected cycle length are allowed but solely to allow immune toxicities to settle
- pembrolizumab will otherwise be used as set out in its Summary of Product Characteristics (SPC)

CDF applications - de-duplication criteria

Before conducting any analysis on CDF treatments, the Blueteq data is examined to identify duplicate applications. The following de-duplication rules are applied:

1. If two trusts apply for pembrolizumab for the treatment of relapsed or refractory classical Hodgkin lymphoma for the same patient (identified using the patient's NHS number), and both applications have the same approval date, then the record where the CDF trust (the trust applying for CDF treatment) matches the SACT treating trust is selected.

2. If two trusts apply for pembrolizumab for the treatment of relapsed or refractory classical Hodgkin lymphoma for the same patient, and the application dates are different, then the record where the approval date in the CDF is closest to the regimen start date in SACT is selected, even if the CDF trust did not match the SACT treating trust.

3. If two applications are submitted for pembrolizumab for the treatment of relapsed or refractory classical Hodgkin lymphoma and the patient has no regimen start date in SACT capturing when the specific drug was delivered, then the earliest application in the CDF is selected.

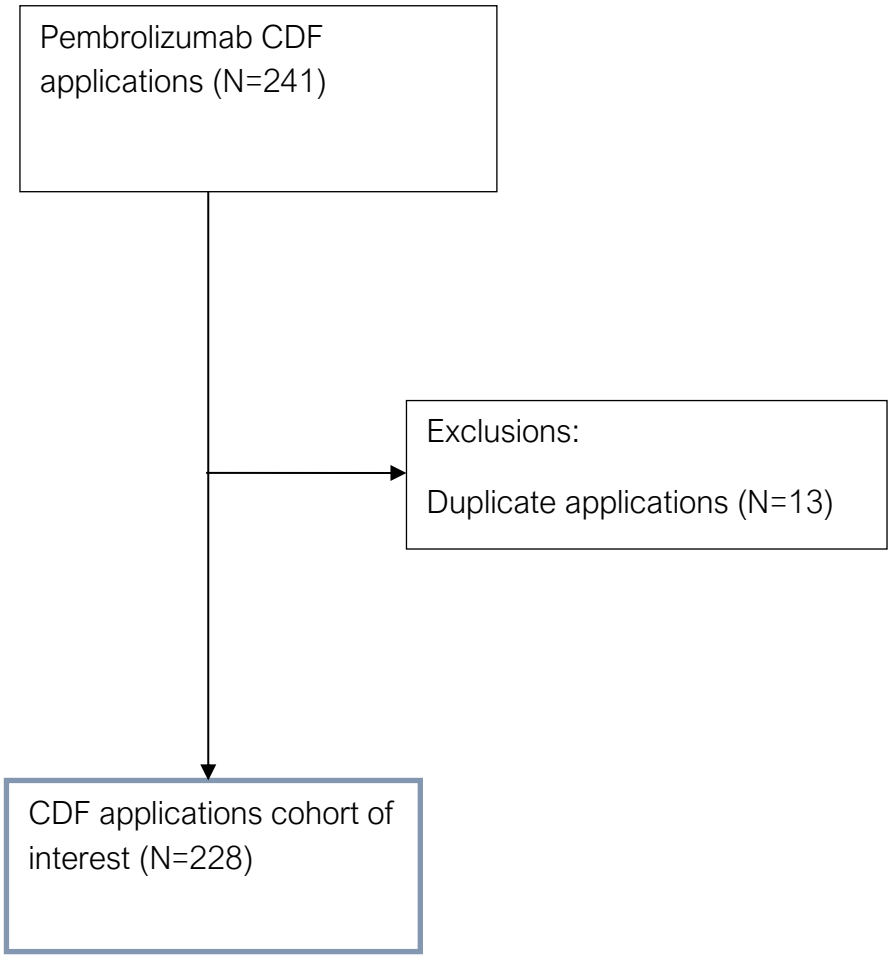
Initial CDF cohorts

The analysis cohort is limited to the date pembrolizumab entered the CDF for this indication, onwards. Any treatments delivered before the CDF entry date are excluded as they are likely to be patients receiving treatment via an Early Access to Medicines Scheme (EAMS) or a compassionate access scheme run by the company. These schemes may have different eligibility criteria compared to the clinical treatment criteria detailed in the CDF managed access agreement for this indication.

The CDF applications included in these analyses are from 25 July 2018 to 30 September 2022. A snapshot of SACT data was taken on 7 January 2023 and made available for analysis on 16 January 2023 and includes SACT activity up to 30 September 2022. Tracing the patients' vital status was carried out on 18 January 2023 using the Personal Demographics Service (PDS)[1].

There were 241 applications for CDF funding for pembrolizumab for the treatment of relapsed or refractory classical Hodgkin lymphoma between 25 July 2018 and 30 September 2022 in the NHS England Blueteq database. Following de-duplication this relates to 228 unique patients. No patients received pembrolizumab prior to the drug being available through the CDF.

Figure 1: Derivation of the cohort of interest from all CDF (Blueteq) applications made for pembrolizumab for the treatment relapsed or refractory classical Hodgkin lymphoma between 25 July 2018 and 30 September 2022



Linking CDF cohort to SACT

NHS numbers were used to link SACT records to CDF applications for pembrolizumab in the Blueteq system. Information on treatments in SACT were examined to ensure the correct SACT treatment records were matched to the CDF application; this includes information on treatment dates (regimen, cycle and administration dates) and primary diagnosis codes in SACT.

Addressing clinical uncertainties

Treatment duration

Treatment duration is calculated from the start of a patient's treatment to their last known treatment date in SACT.

Treatment start date is defined as the date the patient started their CDF treatment. This date is identified as the patient's earliest treatment date in the SACT dataset for the treatment of interest. Data items[8] used to determine a patient's earliest treatment date are:

- Start date of regimen – SACT data item #22
- Start date of cycle – SACT data item #27
- Administration date – SACT data item #34

The earliest of these dates is used as the treatment start date.

The same SACT data items (#22, #27, #34)[8] are used to identify a patient's final treatment date. The latest of these three dates is used as the patient's final treatment date.

Additional explanation of these dates is provided below:

Start date of regimen

A regimen defines the drugs used, their dosage and frequency of treatment. A regimen may contain many cycles. This date is generally only used if cycle or administration dates are missing.

Start date of cycle

A cycle is a period of time over which treatment is delivered. A cycle may contain several administrations of treatment, after each treatment administration, separated by an appropriate time delay. For example; a patient may be on a 3-weekly cycle with treatment being administered on the 1st and 8th day, but nothing on days 2 to 7 and days 9 to 20. The 1st day would be recorded as the "start day of cycle". The patient's next cycle would start on the 21st day.

Administration date

An administration is the date a patient is administered the treatment, which should coincide with when they receive treatment. Using the above example, the administrations for a single 3-week cycle would be on the 1st and 8th day. The next administration would be on the 21st day, which would be the start of their next cycle.

The interval between treatment start date and final treatment date is the patient's time on treatment.

All patients are then allocated a 'prescription length', which is a set number of days added to the final treatment date to allow for the fact that they are effectively still 'on treatment' between administrations. The prescription length should correspond to the typical interval between treatment administrations.

If a patient dies between administrations, then their censor date is their date of death and these patients are deemed to have died on treatment unless an outcome summary is submitted to the SACT database confirming that the patient ended treatment due to disease progression or toxicity before death.

Pembrolizumab is administered intravenously. As such, treatment is generally administered in a healthcare facility and healthcare professionals can confirm that treatment administration has taken place on a specified date. A duration of 20 days has been added to the final treatment date for all patients; this represents the duration from a patient's last cycle to their next[9].

Treatment duration is calculated for each patient as:

Treatment duration (days) = (Final treatment date – Treatment start date) + prescription length (days). This date would be the patient's censored date, unless a patient dies in between their last treatment and the prescription length added, in this case, the censored date would be the patients date of death.

Once a patient's treatment duration has been calculated, the patient's treatment status is identified as one of the following:

No longer receiving treatment (event), if:

- the patient has died.
- the outcome summary, detailing the reason for stopping treatment has been completed:
    - SACT v2.0 data item #41
    - SACT v3.0 data item #58 - #61.
- there is no further SACT records for the patient following a three-month period.

If none of the above apply, the patient is assumed to still be on treatment and is censored.

Overall survival (OS)

OS is calculated from the CDF treatment start date, not the date of a patient's cancer diagnosis. Survival from the treatment start date is calculated using the patient's earliest treatment date, as described above, and the patient's date of death or the date the patient was traced for their vital status.

All patients in the cohort of interest are submitted to the PDS to check their vital status (dead or alive). Patients are traced before any analysis takes place. The date of tracing is used as the date of follow-up (censoring) for patients who have not died.

OS is calculated for each patient as the interval between the earliest treatment date where a specific drug was given to the date of death or date of follow-up (censoring).

OS (days) = Date of death (or follow up) - treatment start date

The patient is flagged as either:

Dead (event):

At the date of death recorded on the PDS.

Alive (censored):

At the date patients were traced for their vital status as patients are confirmed as alive on this date.

Lost to follow-up:

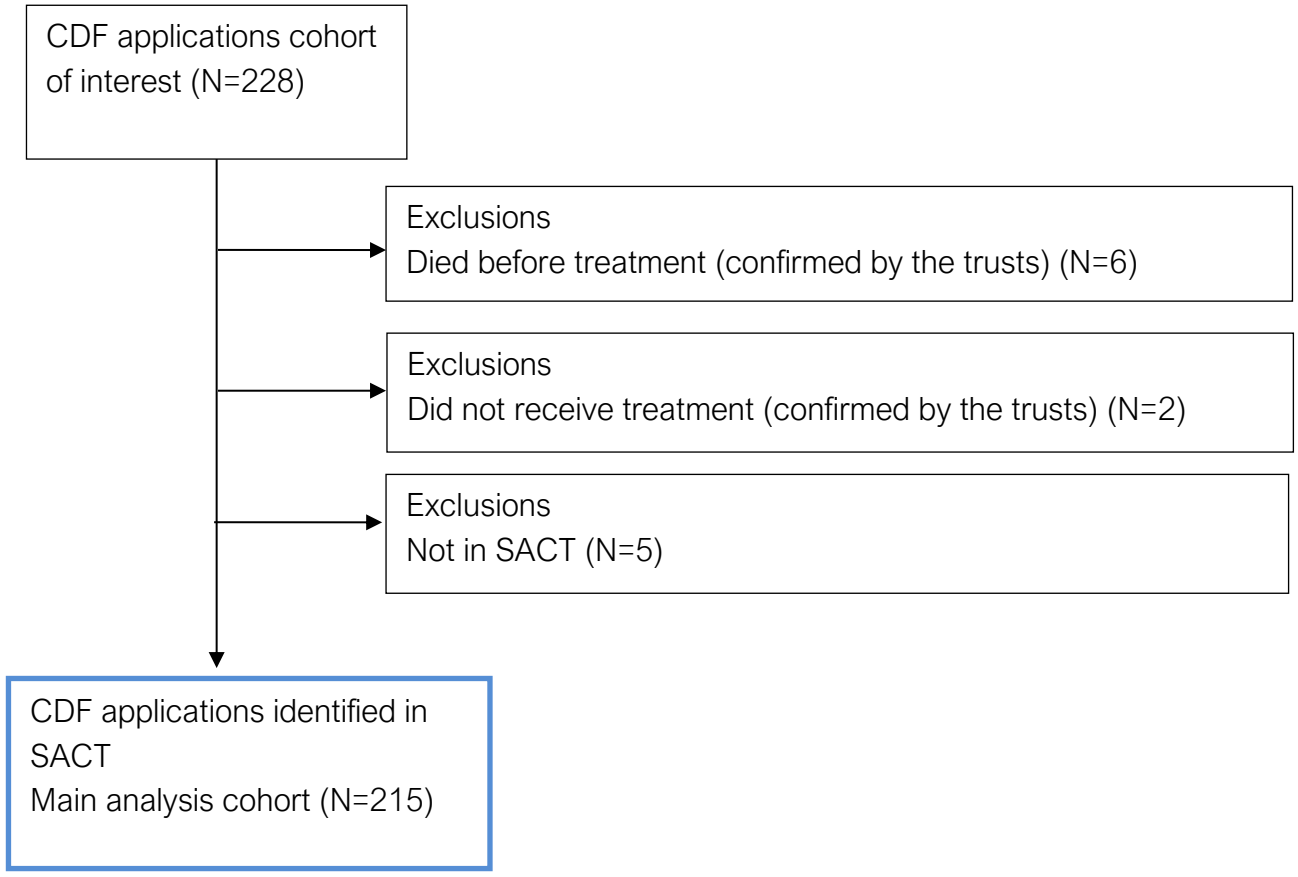
Where we cannot determine whether a patient is alive or not on the censor date; this happens when a patient cannot be successfully traced, for example, because they have emigrated or because important identifiers such as NHS number or date of birth contain errors, the patient's record will be censored at their last known treatment date in SACT. This is the date the patient was last known to be alive.

4. Results

Cohort of interest

Of the 228 applications for CDF funding for pembrolizumab for the treatment of relapsed or refractory classical Hodgkin lymphoma, two patients did not receive treatment, six patients died before treatment and five patients were missing from SACT[a] (see Figure 2).

Figure 2: Matched cohort - SACT data to CDF (Blueteq®) applications for pembrolizumab for the treatment of relapsed or refractory classical Hodgkin lymphoma between 25 July 2018 and 30 September 2022



[a] Of the two patients that did not receive treatment, all were confirmed by the relevant trust. Of the six patients who died before treatment, all were confirmed by the relevant trust by the SACT data liaison team.

A maximum of 220 pembrolizumab records are expected in SACT for patients who were alive, eligible and confirmed to have commenced treatment (Figure 2). 98% (215/220) of these applicants for CDF funding have a treatment record in SACT.

Completeness of SACT key variables

Table 1 presents the completeness of key data items required from SACT. Completeness is 100% for primary diagnosis, date of birth, gender and treatment dates. Performance status at the start of regimen is 76% complete.

Table 1: Completeness of key SACT data items for the pembrolizumab cohort (N=215)

Variable	Completeness (%)
Primary diagnosis	100%
Date of birth (used to calculate age)	100%
Gender	100%
Start date of regimen	100%
Start date of cycle	100%
Administration date	100%
Performance status at start of regimen	76%

Table 2 presents the completeness of regimen outcome summary. A patient's outcome summary, detailing the reason why treatment was stopped, is only captured once a patient has completed their treatment. Therefore, the percentage completeness provided for outcome summary is for records where we assume treatment has stopped and an outcome is expected. Outcomes are expected if a patient has died, has an outcome in SACT stating why treatment has ended or has not received treatment with pembrolizumab in at least three months[9]. These criteria are designed to identify all cases where a patient is likely to have finished treatment. Based on these criteria, outcomes are expected for 178 patients. Of these, 158 (89%) have an outcome summary recorded in the SACT dataset.

Table 2: Completeness of outcome summary for patients that have ended treatment (N=178)

Variable	Completeness (%)
Outcome summary of why treatment was stopped	89%

Completeness of Blueteq key variables

Table 3 presents the completeness of key data items required from Blueteq. Previous treatments information is 100% complete (215/215).

Table 3: Stem cell transplant suitability (N=215)

Variable	Completeness (%)
Stem cell transplant suitability	100%

Patient characteristics

The median age of the 215 patients receiving pembrolizumab for the treatment relapsed or refractory classical Hodgkin lymphoma was 54 years. The median age in males and females was 57 and 47 years respectively.

Table 4: Patient characteristics (N=215)

Patient characteristics[b]			
		N	%
Gender	Male	130	60%
	Female	85	40%
Age	<40	75	35%
	40 to 49	22	10%
	50 to 59	37	17%
	60 to 69	32	15%
	70 to 79	43	20%
	80+	6	3%

[b] Figures may not sum to 100% due to rounding.

Patient characteristics[c]			
		N	%
Performance status at the start of regimen	0	59	27%
	1	86	40%
	2	16	7%
	3	2	1%
	4	0	0%
	Missing	52	24%

[c] Figures may not sum to 100% due to rounding.

Stem cell transplant suitability and procedures

Of the 215 patients in SACT, 132 (61%) patients were identified in Blueteq as being suitable candidates for a stem cell transplant (SCT). Table 5 shows of those who were suitable candidates, the number that were found in the Hospital Episodes Statistics (HES) dataset. HES includes procedures carried out up to 30 September 2022. SACT includes treatments prescribed up to 30 September 2022.

Table 5: Matched cohort - SACT data to CDF (Blueteq) applications made for pembrolizumab for treating Hodgkin lymphoma between 25 July 2018 and 30 September 2022, SCT suitability in Blueteq and SCT procedures in HES[10]

SCT suitability[d]	Blueteq SCT suitability (30 September 2022) (N)	HES Allogenic transplants (N)	HES Autologous transplants (N)	HES SCT (N)	HES SCT (%)
Candidate for future SCT	132	42	23	65	49%
Not a candidate for SCT	83	3	1		
Total	215	45	24	65	

Of the 215 patients with a pembrolizumab treatment record in SACT, who are candidates for a SCT, 65 patients were identified in HES[7] as having received a SCT. Both allogeneic and autologous were included in the search criteria, all 65 patients were recorded as receiving a SCT after their last pembrolizumab treatment date in SACT.

[d] Allogeneic and autologous are the two main types of stem cell transplant.

Of the 65 patients who received a SCT, all were identified as ending treatment with pembrolizumab, outcomes shown in table 6.

Table 6: SCT suitability in Blueteq and SCT procedures in HES

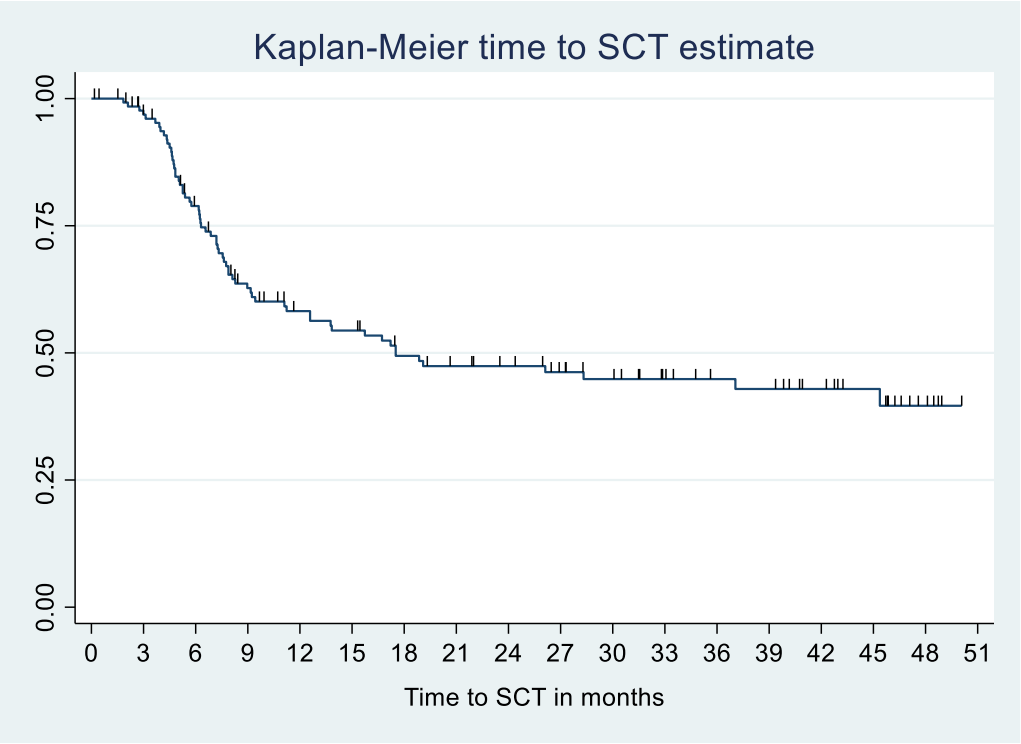
Outcome in SACT – reason for ending treatment	Received a SCT in HES
Completed treatment as prescribed	36
Progressive disease during chemotherapy	21
Acute chemotherapy toxicity	3
No treatment in at least three months	2
Patient choice (stopped or interrupted treatment)	2
Palliative, patient did benefit	1
Total	65

Of the 132 patients with CDF applications and who are suitable candidates for a SCT, 65 (49%) were identified as receiving a SCT in HES by 30 September 2022. Of the 65 patients, the time from their first pembrolizumab treatment date in SACT to receiving a SCT ranges from 1.8 months to 45.4 months with the median amongst only those who had a SCT being 6.9 months, this is calculated for the 65 patients from their earliest pembrolizumab treatment date in SACT to their SCT procedure date in HES.

The Kaplan-Meier curve for time to SCT, shown in Figure 3, presents the median time from a patient's first pembrolizumab treatment to a SCT, calculated from a patient's first pembrolizumab treatment date in SACT to their SCT procedure date in HES, or 30 September 2022 if a SCT was not carried out at the time this report was produced. This uses all patients to calculate the time between pembrolizumab and SCT, rather than those who only had a SACT (as in the previous paragraph).

Figure 3 shows, of all patients eligible for SCT, the median time after which 50% of patients have gone on to receive a SCT, is 17.5 months[e] (532 days),

Figure 3: Kaplan-Meier SCT suitability in Blueteq and SCT procedures in HES (N=132)



[e] Confidence intervals could not be produced as there was an insufficient number of events at the time this report was produced.

Table 7 and Table 8 show the number of patients at risk, the number of patients that were censored and the number of patients that received a stem cell transplant (events) from the time patients started treatment to the end of the follow-up period.

Table 7: Includes the number of patients at risk, by quarterly breakpoints.

Time intervals (months)	0	3	6	9	12	15	18	21	24	27	30	33	36	39	42	45	48
Number at risk	132	120	94	71	61	57	49	45	42	37	32	26	23	22	17	13	4

Table 8 shows that for all patients who received treatment and were a candidate for SCT, 67 were yet to receive a stem cell transplant (censored) at the date of follow-up and 65 had received a stem cell transplant (events).

Table 8: Number of patients at risk, by quarterly breakpoints split between patients that have ended treatment (events) and patients that are still on treatment (censored).

Time intervals (months)	0	3	6	9	12	15	18	21	24	27	30	33	36	39	42	45	48
Censored	67	59	55	51	46	46	43	41	38	34	30	24	21	21	16	12	4
Events	65	61	39	20	15	11	6	4	4	3	2	2	2	1	1	1	0

Table 9 reports, that of the 69 patients who received a SCT, suitable and not suitable candidates, 63 (91%) patients were prescribed a regimen after pembrolizumab and before receiving a SCT, as recorded in the SACT dataset, some patients have more than one intermediary therapy.

Table 9: Distribution of treatments received after pembrolizumab but before SCT (N(Patients)=63)[f,g,h]

Regimen	Number of subsequent treatments
Carmustine + Cytarabine + Etoposide + Melphalan	10
Cytarabine + Etoposide + Lomustine + Melphalan	9
Bendamustine	7
Transplant Alemtuzumab + Carmustine + Cytarabine + Etoposide + Melphalan	7
Transplant Alemtuzumab + Cytarabine + Etoposide + Lomustine + Melphalan	6
Trial Unspecified	6
Alemtuzumab + Cytarabine + Etoposide + Lomustine + Melphalan	5
Transplant Alemtuzumab + Fludarabine + Melphalan	5
Bendamustine + Gemcitabine + Vinorelbine	4
Transplant Carmustine + Cytarabine + Etoposide + Melphalan	4
Transplant Cyclophosphamide + Fludarabine (+/- Tbi)	4
Cisplatin + Gemcitabine	3
Cyclophosphamide	3
Alemtuzumab + Fludarabine + Melphalan	2
Bleomycin + Cyclophosphamide + Doxorubicin + Etoposide + Procarbazine + Vincristine	2
Transplant Alemtuzumab + Cytarabine + Etoposide + Fludarabine + Lomustine + Melphalan	2
Transplant Fludarabine + Treosulfan	2
Bendamustine + Rituximab	1
Carboplatin + Etoposide + Ifosfamide	1
Cisplatin + Cytarabine + Etoposide	1
Cisplatin + Gemcitabine + Rituximab	1
Gemcitabine	1

[f] Table 9 lists all therapies prescribed between a patient's last pembrolizumab treatment in SACT and their first SCT.
[g] These data have not been validated/confirmed with trusts or by the SACT DLO team.
[h] Some patients will have received more than one therapy between their last pembrolizumab treatment in SACT and their first SCT.

Regimen	Number of subsequent treatments
Transplant Alemtuzumab + Cyclophosphamide + Cytarabine + Etoposide + Lomustine	1
Vinblastine	1
Vinorelbine	1
Total	89

Treatment duration

Of the 215 patients with CDF applications, 178 (83%) were identified as having completed treatment by 30 September 2022 (latest follow up in SACT dataset). Patients are assumed to have completed treatment if they have died, have an outcome summary recorded in the SACT dataset or they have not received treatment with pembrolizumab in at least three months (see Table 14). The median follow-up time in SACT was 4.8 months (146 days). The median follow-up time in SACT is the patients' median observed time from the start of their treatment to their last treatment date in SACT plus the prescription length.

Presently, 94% (N=132) of trusts submit their SACT return to the submission portal two months after the month's treatment activity has ended; this provides a maximum follow-up period of 50 months. 6% (N=9) of trusts submit their SACT return to the submission portal one month after the month's treatment activity has ended; this provides a maximum follow-up period of 51 months. SACT follow-up ends 30 September 2022.

Table 10: Breakdown by patients' treatment status [i,j,k]

Patient status	Frequency (N)	Percentage (%)
Patient died – not on treatment	71	33%
Patient died – on treatment	2	1%
Treatment stopped	105	49%
Treatment ongoing	37	17%
Total	215	100%

[i] Figures may not sum to 100% due to rounding.
[j] Table 10 presents the outcome summary data reported by trusts. This includes patients from Table 6 who 'died on treatment', 'died not on treatment' and 'stopped treatment'.
[k] 'Deaths on treatment' and 'deaths not on treatment' are explained in the methodology paper available on the SACT website: hiip://www.chemodataset.nhs.uk/nhse_partnership/ .

Table 11: Treatment duration at 6, 12, 18, 24-month intervals

Time period	Treatment duration (%)
6 months	45% [95% CI: 38%, 51%]
12 months	24% [95% CI: 18%, 31%]
18 months	16% [95% CI: 11%, 22%]
24 months	10% [95% CI: 6%, 16%]

The Kaplan-Meier curve for treatment duration is shown in Figure 4. The median treatment duration for all patients was 5.0 months [95% CI: 4.3, 6.2] (152 days) (N=215).

Figure 4: Kaplan-Meier treatment duration (N=215)

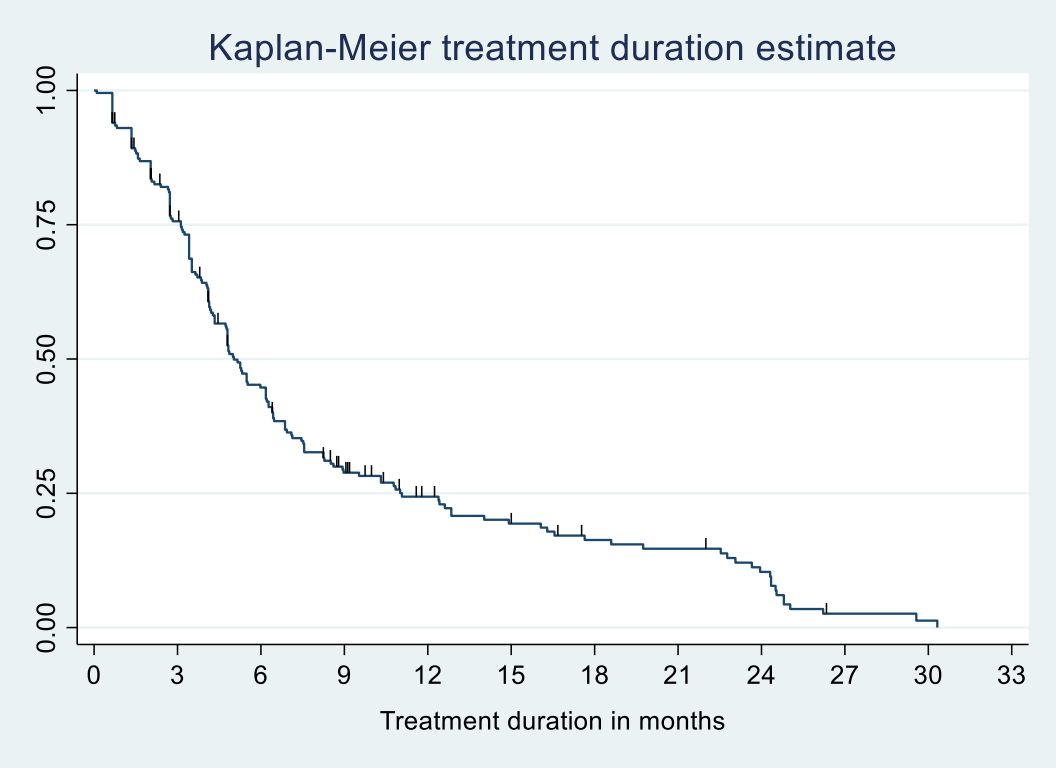


Table 12 and Table 13 show the number of patients at risk, the number of patients that were censored and the number of patients that ended treatment (events) from the time patients started treatment to the end of the follow-up period. The maximum follow-up period for all patients for treatment duration was 50 months[l] (1,521 days). SACT contains more follow-up for some patients.

Table 12: Number of patients at risk, by quarterly breakpoints

Time intervals (months)	0-33	3-33	6-33	9-33	12-33	15-33	18-33	21-33	24-33	27-33	30-33
Number at risk	215	152	86	50	35	26	20	18	12	2	1

Table 13 shows that for all patients who received treatment, 37 were still on treatment (censored) at the date of follow-up and 178 had ended treatment (events).

Table 13: Number of patients at risk, by quarterly breakpoints split between patients that have ended treatment (events) and patients that are still on treatment (censored)

Time intervals (months)	0-33	3-33	6-33	9-33	12-33	15-33	18-33	21-33	24-33	27-33	30-33
Censored	37	25	20	14	6	4	2	2	1	0	0
Events	178	127	66	36	29	22	18	16	11	2	1

[l] Pembrolizumab monotherapy will be continued for a maximum of 12 months (or a maximum of 18 cycles if given 3-weekly).

Table 14 gives a breakdown of a patient's treatment outcome recorded in SACT when a patient's treatment has come to an end. 83% (N=178) of patients had ended treatment at 30 September 2022.

Table 14: Treatment outcomes for patients that have ended treatment (N=178)[m,n]

Outcome	Frequency (N)	Percentage (%)
Stopped treatment – progression of disease	68	38%
Stopped treatment – completed as prescribed	54	30%
Stopped treatment – died not on treatment[o]	21	12%
Stopped treatment – no treatment in at least 3 months	11	6%
Stopped treatment – acute toxicity	8	4%
Stopped treatment – palliative, patient did benefit	6	3%
Stopped treatment – patient choice	6	3%
Stopped treatment – palliative, patient did not benefit	2	1%
Stopped treatment – died on treatment	2	1%
Total	**178**	**100%**

[m] Figures may not sum to 100% due to rounding.
[n] Table 14 presents the outcome summary data reported by trusts. This includes patients from Table 10 who 'died on treatment', 'died not on treatment' and 'stopped treatment'.
[o] 'Deaths on treatment' and 'deaths not on treatment are explained in the methodology paper available on the SACT website.

Table 15: Treatment outcomes and treatment status for patients that have ended treatment (N=178)

Outcome[p]	Patient died [q] not on treatment	Treatment stopped	Patient died on treatment
Stopped treatment – progression of disease	29	39	
Stopped treatment – completed as prescribed	9	45	
Stopped treatment – died not on treatment	21		
Stopped treatment – no treatment in at least 3 months		11	
Stopped treatment – acute toxicity	4	4	
Stopped treatment – palliative, patient did benefit	3	3	
Stopped treatment – patient choice	3	3	
Stopped treatment – palliative, patient did not benefit	2		
Stopped treatment – died on treatment			2
Total	**71**	**105**	**2**

[p] Relates to outcomes submitted by the trust in Table 14.
[q] Relates to treatment status in Table 10 for those that have ended treatment.

Overall survival (OS)

Of the 215 patients with a treatment record in SACT, the minimum follow-up was 3.6 months (109 days) from the last CDF application. Patients were traced for their vital status on 18 January 2023. This date was used as the follow-up date (censored date) if a patient is still alive. The median follow-up time was 19.2 months (584 days). The median follow-up is the patients' median observed time from the start of their treatment to death or censored date.

Table 16: OS at 6, 12, 18, 24, 36 and 48-month intervals

Time period	OS (%)
6 months	88% [95% CI: 83%, 92%]
12 months	82% [95% CI: 76%, 87%]
18 months	75% [95% CI: 68%, 80%]
24 months	68% [95% CI: 61%, 75%]
36 months	56% [95% CI: 47%, 64%]
48 months	55% [95% CI: 46%, 63%]

Figure 5 provides the Kaplan-Meier curve for OS, censored at 18 January 2023. The median OS was not reached.

Figure 5: Kaplan-Meier survival plot (N=215)

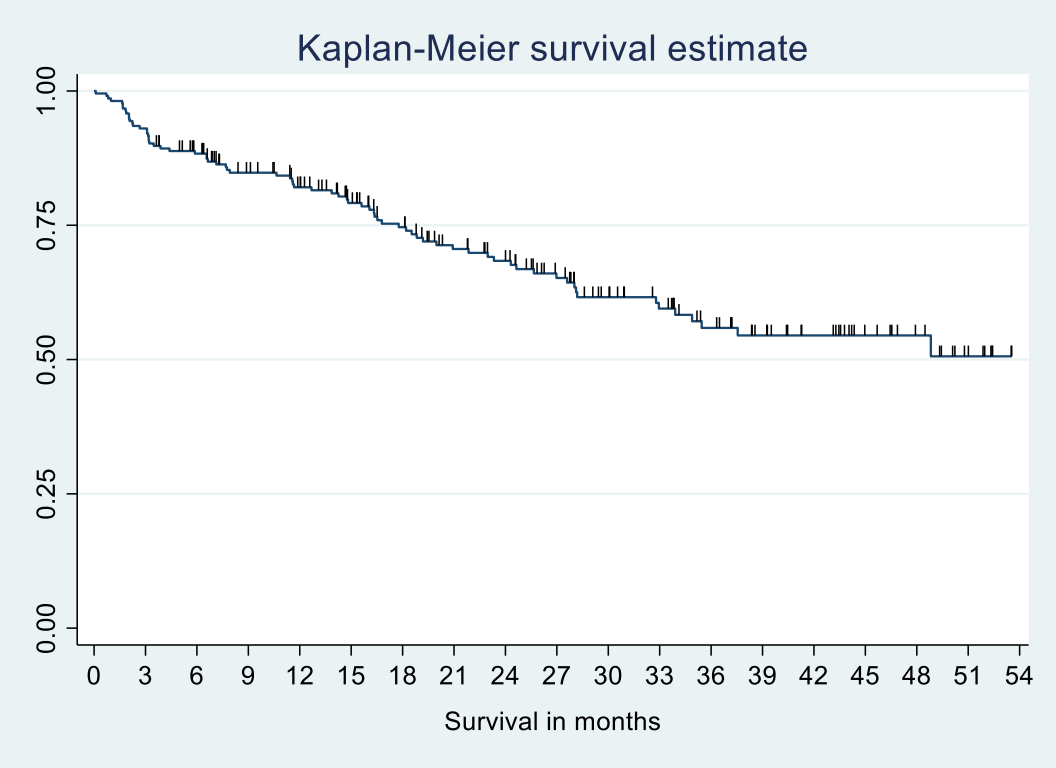


Table 17 and Table 18 show the number of patients at risk, the number of patients that were censored and the number of patients that died (events) from the time patients started treatment to the end of the follow-up period. The maximum follow-up period for survival was 53.8 months (1,637 days), all patients were traced on 18 January 2023.

Table 17: Includes the number of patients at risk, by quarterly breakpoints

Time intervals (months)	0	3	6	9	12	15	18	21	24	27	30	33	36	39	42	45	48	51
Number at risk	215	200	182	161	148	129	114	100	91	77	62	56	45	36	29	20	15	7

Table 18 shows that for all patients who received treatment, 142 were still alive (censored) at the date of follow-up and 73 had died (events).

Table 18: Number of patients at risk, those that have died (events) and those that are still alive (censored) by quarterly breakpoints

Time intervals (months)	0	3	6	9	12	15	18	21	24	27	30	33	36	39	42	45	48	51
Censored	142	142	134	120	112	98	90	82	76	66	55	51	43	35	28	19	14	7
Events	73	58	48	41	36	31	24	18	15	11	7	5	2	1	1	1	1	0

5. Sensitivity analyses

6-months follow up

Treatment duration

Sensitivity analyses were carried out on a cohort with at least six months follow-up in SACT. To identify the treatment duration cohort, CDF applications were limited from 25 July 2018 to 31 March 2022 and SACT activity was followed up to the 30 September 2022.

Following the exclusions above, 187 patients (87%) were identified for inclusion. The median follow-up time in SACT was 5.2 months (159 days). The median follow-up time in SACT is the patients' median observed time from the start of their treatment to their last treatment date in SACT plus prescription length.

The Kaplan-Meier curve for treatment duration is shown in Figure 6. The median treatment duration for patients in this cohort was 5.2 months [95% CI: 4.3, 6.2] (158 days) (N=187).

Figure 6: Kaplan-Meier treatment duration plot (N=187)

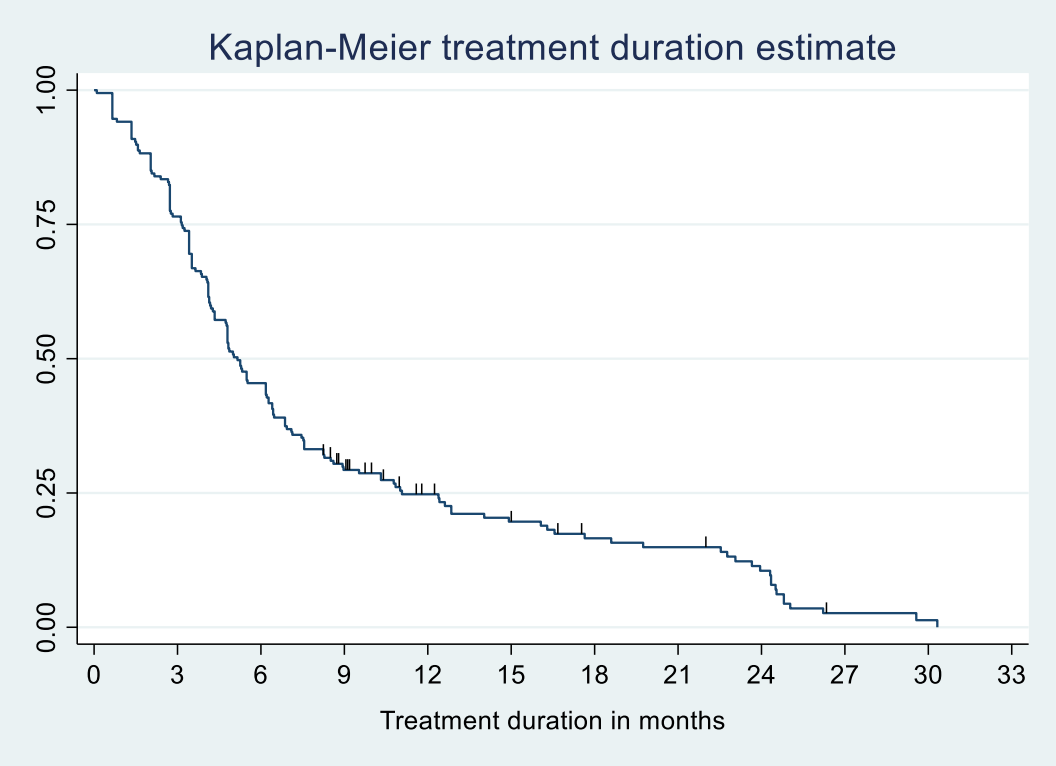


Table 19 and Table 20 show the number of patients at risk, the number of patients that were censored and the number of patients that ended treatment (events) from the time patients started treatment to the end of the follow-up period. The maximum follow-up period for all patients for treatment duration was 50 months[r] (1,521 days). SACT contains more follow-up for some patients.

Table 19: Includes the number of patients at risk, by quarterly breakpoints

Time intervals (months)	0-33	3-33	6-33	9-33	12-33	15-33	18-33	21-33	24-33	27-33	30-33
Number at risk	187	143	85	50	35	26	20	18	12	2	1

Table 20 shows that for all patients who received treatment, 19 were still on treatment (censored) at the date of follow-up and 168 had ended treatment (events).

Table 20: Number of patients at risk, by quarterly breakpoints split between patients that have ended treatment (events) and patients that are still on treatment (censored)

Time intervals (months)	0-33	3-33	6-33	9-33	12-33	15-33	18-33	21-33	24-33	27-33	30-33
Censored	19	19	19	14	6	4	2	2	1	0	0
Events	168	124	66	36	29	22	18	16	11	2	1

[r] Pembrolizumab monotherapy will be continued for a maximum of 12 months (or a maximum of 18 cycles if given 3-weekly.

Overall survival (OS)

Sensitivity analyses was also carried out for OS on a cohort with at least six months follow-up. To identify the cohort, CDF applications were limited from 25 July 2018 to 18 July 2022 and patients were traced for their vital status on 18 January 2023.

Following the exclusions above, 206 patients (96%) were identified for inclusion. The median follow-up time was 20 months (608 days).

The median follow-up is the patients' median observed time from the start of their treatment to death or censored date.

The Kaplan-Meier curve for OS is shown in Figure 7. The median OS for patients in this cohort was not reached.

Figure 7: Kaplan-Meier survival plot (N=206)

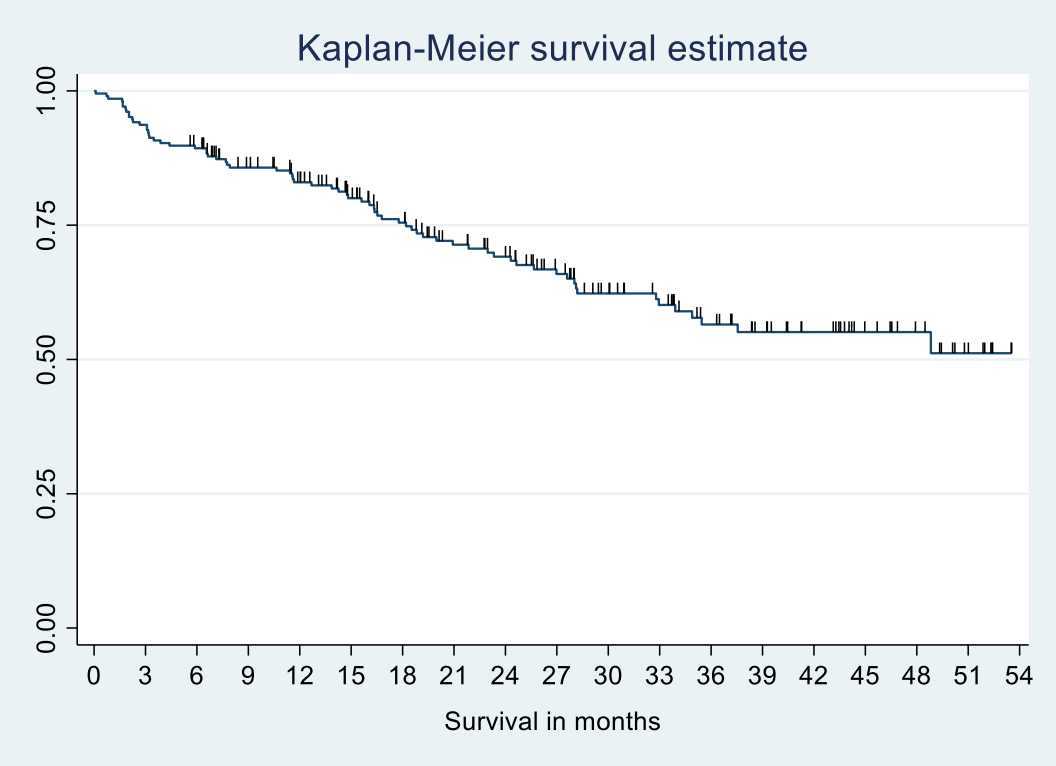


Table 21 and Table 22 show the number of patients at risk, the number of patients that were censored and the number of patients that died (events) from the time patients started treatment to the end of the follow-up period. The maximum follow-up period for survival was 53.8 months (1,637 days), all patients were traced on 18 January 2023.

Table 21: Includes the number of patients at risk, by quarterly breakpoints

Time intervals (months)	0	3	6	9	12	15	18	21	24	27	30	33	36	39	42	45	48	51
Number at risk	206	193	182	161	148	129	114	100	91	77	62	56	45	36	29	20	15	7

Table 22 shows that for all patients who received treatment, 136 were still alive (censored) at the date of follow-up and 70 had died (events).

Table 22: Number of patients at risk, those that have died (events) and those that are still alive (censored) by quarterly breakpoints

Time intervals (months)	0	3	6	9	12	15	18	21	24	27	30	33	36	39	42	45	48	51
Censored	136	136	134	120	112	98	90	82	76	66	55	51	43	35	28	19	14	7
Events	70	57	48	41	36	31	24	18	15	11	7	5	2	1	1	1	1	0

Table 23: Median treatment duration and OS, full cohort and sensitivity analysis

Metric	Main CDF cohort Standard analysis: Full cohort	Sensitivity analysis: 6 months follow-up cohort: treatment duration	Sensitivity analysis: 6 months follow-up cohort: OS
N	215	187	206
Median treatment duration	5.0 months [95% CI: 4.3, 6.2] (152 days)	5.2 months [95% CI: 4.3, 6.2] (158 days)	
OS	Not reached		Not reached

6. Conclusions

220 patients received pembrolizumab for the treatment of relapsed or refractory classical Hodgkin lymphoma [TA540] through the CDF in the reporting period (25 July 2018 and 30 September 2022). 215 patients were reported to the SACT dataset, giving a SACT dataset ascertainment of 98%. An additional two patients with a CDF application did not receive treatment and six patients died before treatment. Both patients who did not receive treatment and the six patients identified as death before treatment were confirmed by the trust responsible for the CDF application by the team at NHS England.

Patient characteristics from the SACT dataset show that 60% (N=130) of patients who received pembrolizumab for the treatment of relapsed or refractory classical Hodgkin lymphoma were male and 40% (N=85) of patients were female. Most of the cohort were aged <40 36%, (N=75) and 66% (N=145) of patients had a performance status between 0 and 1 at the start of their regimen.

At data cut off, 83% (N=178) of patients were identified as no longer being on treatment. Of these 178 patients:

- 38% (N=68) of patients stopped treatment due to disease progression
- 30% (N=54) of patients completed treatment as prescribed
- 12% (N=21) of patients died not on treatment
- 6% (N=11) of patients did not have a treatment record in SACT in at least three months and are assumed to have completed treatment
- 4% (N=8) of patients stopped treatment due to acute toxicity
- 3% (N=6) of patients were treated palliatively and did benefit from the treatment they received
- 3% (N=6) of patients chose to end their treatment
- 1% (N=2) of patients were treated palliatively and did not benefit from the treatment they received
- 1% (N=2) of patients died on treatment

Median treatment duration was 5.0 months [95% CI: 4.3, 6.2] (152 days). 45% of patients were still receiving treatment at 6 months [95% CI: 38%, 51%], 24% of patients were still receiving treatment at 12 months [95% CI: 18%, 31%], 16% of patients were still receiving treatment at 18 months [95% CI: 11%, 22%] and 10% of patients were still receiving treatment at 24 months [95% CI: 6%, 16%].

The median OS was not reached. OS at 6 months was 88% [95% CI: 83%, 92%], 12 months OS was 82% [95% CI: 76%, 87%], OS at 18 months was 75% [95% CI: 68%, 80%], OS at 24 months was 68% [95% CI: 61%, 75%], OS at 36 months was 56% [95% CI: 47%, 64%] and OS at 48 months was 55% [95% CI: 46%, 63%].

Sensitivity analysis was carried out on treatment duration and OS to evaluate a cohort for which all patients had a minimum follow-up of six months. Results for treatment duration showed a very slight difference that was not statistically significant (full cohort = 5.0 months; sensitivity analysis cohort = 5.2 months). OS was the same when comparing the full cohort to the limited cohort, the median OS was not reached.

7. References

1. The Personal Demographics Service (PDS). NHS Digital: 2023 [cited 2023 Feb]. Available from: hiips://digital.nhs.uk/Demographics

2. National Statistics. Cancer Registration Statistics, England: 2020. 2022 [cited 2023 Feb]. Available from: hiips://digital.nhs.uk/data -and-information/publications/statistical/cancer-registration-statistics/england-2020

3. National Institute for Health and Care Excellence: 2018 [cited 2023 Feb] Available from: hiips://www.nice.org.uk/guidance/ta540/chapter/1 -Recommendations

4. Cancer Drugs Fund. [Internet]. NHS England: 2017 [cited 2023 Feb]. Available from: hiips://www.england.nhs.uk/cancer/cdf/

5. Appraisal and funding of Cancer Drugs. NHS England: 2016 [cited 2023 Feb]. Available from: hiips://www.england.nhs.uk/wp -content/uploads/2013/04/cdf-sop.pdf

6. National Institute for Health and Care Excellence: 2018 [cited 2023 Feb]. Available from: hiips://www.nice.org.uk/guidance/ta540/resources

7. Phase II clinical study (KEYNOTE-087) clinical trial: 2015 [cited 2023 Feb] Available from: hiips://clinicaltrials.gov/ct2/show/NCT02453594

8. Systemic Anti-Cancer Therapy [Internet]: SACT: 2023 [cited 2023 Feb]. Available from: hiips://digital.nhs.uk/ndrs/data/data -sets/sact

9. CDF analytical methods. [Internet]. NHSD: 2019 [cited 2023 Feb]. Available from: hiip://www.chemodataset.nhs.uk/nhse_partnership/

10. Hospital Episode Statistics (HES). [Internet]. NHS Digital: 2022 [cited 2022 Feb]. Available from: hiips://digital.nhs.uk/data -and-information/data-tools-and-services/data-services/hospital-episode-statistics

8. Appendix

OPCS codes used in HES to identify a stem cell transplant procedure can be found in Table A-1. The procedure time frame used in HES was 2018 to 2022.

Table A-1: OPCS codes used to identify a stem cell transplant procedure in HES

OPCS code	Stem cell transplant
X334	Autologous peripheral blood stem cell transplant
X336	Allogeneic peripheral blood stem cell transplant
X335	Syngenetic peripheral blood stem cell transplant

9. Addendum

Pembrolizumab for treating relapsed or refractory classical Hodgkin lymphoma (TA540)

Subsequent to provision of the initial draft of this report to NICE and Marck Sharp & Dohme Ltd, NHS England were requested to supply some additional information relating to Kaplan-Meier curves for two cohorts defined as the following:

- overall survival amongst patients who did not receive a SCT[s]
- Kaplan-Meier curve where the event is either a SCT or death.

Overall survival amongst patients who did not receive a SCT

Of the 215 patients included in this report, 150 (70%) were identified as not receiving a SCT at the time this report was produced.

The median follow-up time amongst the 150 patients was 14.6 months (444 days).

Table 1: OS at 6, 12, 18, 24, 36 and 48-month intervals

Time period	OS (%) – patients who did not receive a SCT
6 months	83% [95% CI: 76%, 88%]
12 months	76% [95% CI: 68%, 82%]
18 months	65% [95% CI: 56%, 73%]
24 months	54% [95% CI: 44%, 63%]
36 months	37% [95% CI: 26%, 48%]
48 months	34% [95% CI: 24%, 46%]

[s] Patients were not identified in HES at the time this report was produced as receiving a SCT.

Figure 1 provides the Kaplan-Meier curve for OS, censored at 18 January 2023. The median OS was 28 months [95% CI: 20.0, 34.9] (852 days)

Figure 1: Kaplan-Meier survival plot (N=150)

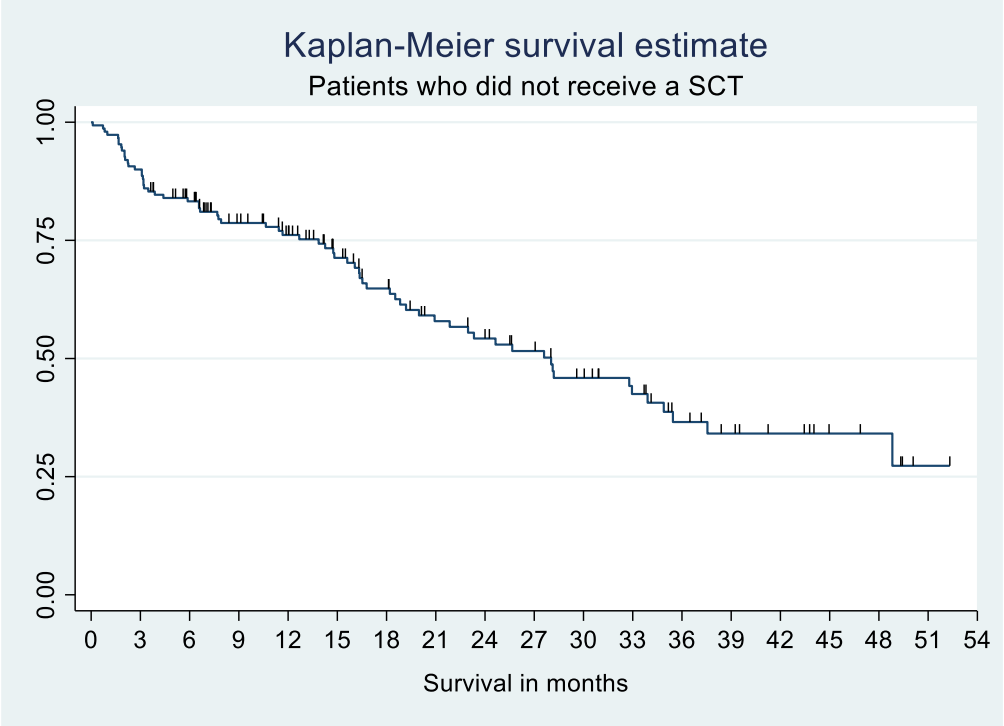


Table 2 and Table 3 show the number of patients at risk, the number of patients that were censored and the number of patients that died (events) from the time patients started treatment to the end of the follow-up period. The maximum follow-up period for survival was 53.8 months (1,637 days), all patients were traced on 18 January 2023.

Table 2: Includes the number of patients at risk, by quarterly breakpoints

Time intervals (months)	0	3	6	9	12	15	18	21	24	27	30	33	36	39	42	45	48	51
Number at risk	150	135	117	97	86	70	58	48	43	37	30	25	17	13	10	6	5	1

Table 3 shows that for all patients who received treatment, 83 were still alive (censored) at the date of follow-up and 67 had died (events).

Table 3: Number of patients at risk, those that have died (events) and those that are still alive (censored) by quarterly breakpoints

Time intervals (months)	0	3	6	9	12	15	18	21	24	27	30	33	36	39	42	45	48	51
Censored	83	83	75	61	53	42	36	32	30	26	23	20	15	12	9	5	4	1
Events	67	52	42	36	33	28	22	16	13	11	7	5	2	1	1	1	1	0

Kaplan-Meier curve where the event is either a SCT or death

Of the 215 patients included in this report, 65 (30%) patients who were identified as candidates for receiving a stem cell transplant (SCT) in the NHS England Blueteq database, were recorded as receiving a SCT in the hospital episode statistics (HES) dataset, and 73 (34%) patients had died at the time this report was produced. The event date used was the date a SCT was carried out or a patient's date of death.

- patients who received a SCT, who subsequently died had their SCT procedure date allocated.
- patients who received a SCT and are still alive had their SCT procedure date allocated.
- patients who have died and did not receive a SCT had their date of death allocated.
- patients who are still alive and have not received a SCT had the death trace date allocated, death trace was carried out on 18 January 2023.

Table 4: Includes the number of patients who received a SCT and those who have died, or both.

	Deaths		
SCT	no	yes	Total
no	83	67	150
yes	59	6	65
Total	142	73	215

The median follow-up time amongst the 215 patients was 11.3 months (343 days).

Table 5: Events at 6, 12, 18, 24, 36 and 48-month intervals

Time period	Kaplan-Meier curve where the event is either a SCT or death
6 months	76% [95% CI: 70%, 81%]
12 months	59% [95% CI: 51%, 65%]
18 months	46% [95% CI: 38%, 53%]
24 months	37% [95% CI: 30%, 45%]
36 months	25% [95% CI: 18%, 33%]
48 months	19% [95% CI: 11%, 28%]

Figure 2 provides the Kaplan-Meier curve for events, censored at a patients SCT procedure date, date of death or trace date, patients were traced for their vital status on 18 January 2023. The median time to event was 16.4 months [95% CI: 12.7, 18.9] (499 days)

Figure 2: Kaplan-Meier time to event plot (N=215)

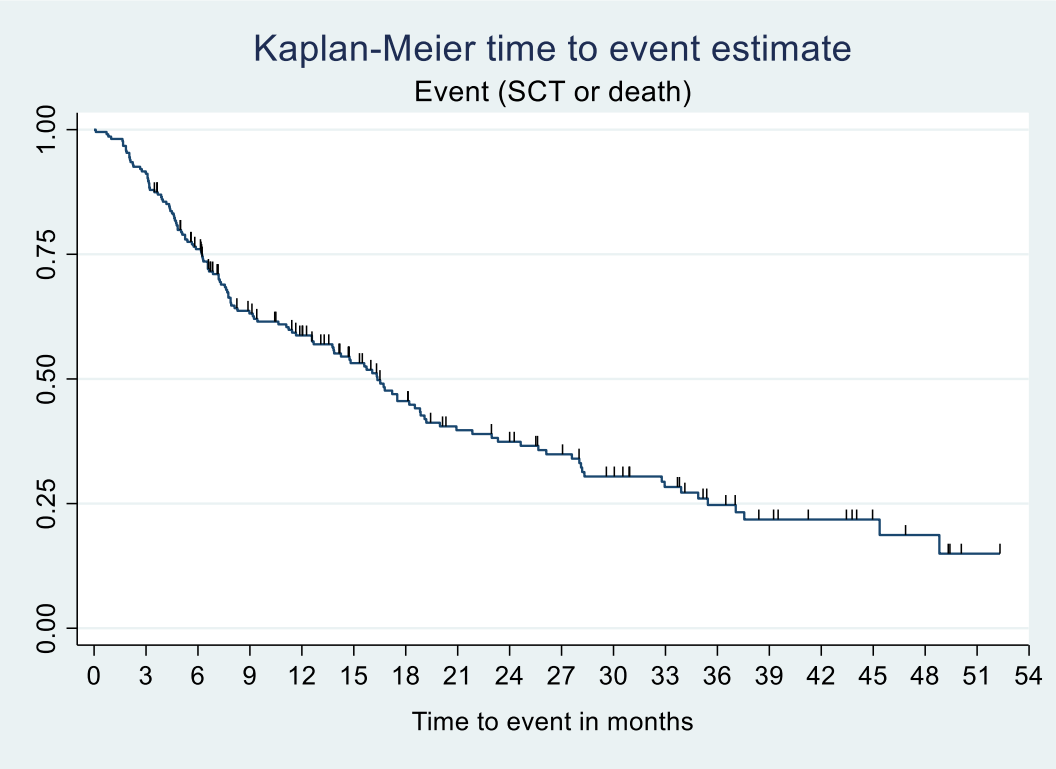


Table 6 and Table 7 show the number of patients at risk, the number of patients that were censored and the number of patients that had an event (SCT or death) from the time patients started treatment to the end of the follow-up period.

Table 6: Includes the number of patients at risk, by quarterly breakpoints

Time intervals (months)	0	3	6	9	12	15	18	21	24	27	30	33	36	39	42	45	48	51
Number at risk	215	196	156	117	101	81	64	52	47	40	32	27	19	14	11	7	5	1

Table 7 shows that for all patients who received treatment, 83 patients had no event (censored) at the date of follow-up and 132 patients had an event (SCT or death).

Table 7: Number of patients at risk, those who have had an event (SCT or death) and those who have not had an event (censored) by quarterly breakpoints

Time intervals (months)	0	3	6	9	12	15	18	21	24	27	30	33	36	39	42	45	48	51
Censored	83	83	75	61	53	42	36	32	30	26	23	20	15	12	9	5	4	1
Events	132	113	81	56	48	39	28	20	17	14	9	7	4	2	2	2	1	0

NICE National Institute for Health and Care Excellence

Single Technology Appraisal

Pembrolizumab for treating relapsed or refractory classical Hodgkin's lymphoma after brentuximab vedotin [review of TA540] ID5084

Clinical expert statement

Information on completing this form

In part 1 we are asking for your views on this technology. The text boxes will expand as you type.

In part 2 we are asking you to provide 5 summary sentences on the main points contained in this document.

Please do not embed documents (such as a PDF) in a submission because this may lead to the information being mislaid or make the submission unreadable. Please type information directly into the form.

Do not include medical information about yourself or another person that could identify you or the other person.

We are committed to meeting the requirements of copyright legislation. If you want to include **journal articles** in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs. For copyright reasons, we will have to return forms that have attachments without reading them. You can resubmit your form without attachments, but it must be sent by the deadline.

Combine all comments from your organisation (if applicable) into 1 response. We cannot accept more than 1 set of comments from each organisation.

Please underline all confidential information, and separately highlight information that is submitted as 'confidential [CON]' in turquoise, and all information submitted as 'depersonalised data [DPD]' in pink. If confidential information is submitted, please also send a second version of your comments with that information redacted. See Health technology evaluations: interim methods and process guide for the proportionate approach to technology appraisals (section 3.2) for more information.

The deadline for your response is **5pm** on **Monday 5 February 2024** Please log in to your NICE Docs account to upload your completed form, as a Word document (not a PDF).

Thank you for your time.

We reserve the right to summarise and edit comments received, or not to publish them at all, if we consider the comments are too long, or publication would be unlawful or otherwise inappropriate.

Comments received are published in the interests of openness and transparency, and to promote understanding of how recommendations are developed. The comments are published as a record of the comments we received, and are not endorsed by NICE, its officers or advisory committees.

Part 1: Pembrolizumab for treating relapsed or refractory classical Hodgkin's lymphoma after brentuximab vedotin [review of TA540] ID5084

Table 1 About you, aim of treatment, place and use of technology, sources of evidence and equality

NICE National Institute for Health and Care Excellence

1. Your name	Patrick Medd
2. Name of organisation	University Hospitals Plymouth NHS Trust
3. Job title or position	Consultant Haematologist
4. Are you (please tick all that apply)	☐ An employee or representative of a healthcare professional organisation that represents clinicians? ☒ A specialist in the treatment of people with <<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 did not submit one, I do not know if they submitted one etc.)
6. If you wrote the organisation submission and/or do not have anything to add, tick here. (If you tick this box, the rest of this form will be deleted after submission)	☐ Yes
7. Please disclose any past or current, direct or indirect links to, or funding from, the tobacco industry.	None

8. What is the main aim of treatment for **treating relapsed or refractory classical Hodgkin's lymphoma after brentuximab vedotin[review of TA540] ID5084**? (For example, to stop progression, to improve mobility, to cure the condition, or prevent progression or disability)	To achieve a disease response, ideally a complete response which will prolong survival and will cure the condition in a proportion of patients. Patients achieving a partial or complete response may be able to proceed to autologous or allogeneic stem cell transplant with curative intent, although not all patients will be suitable for this treatment.
9. What do you consider a clinically significant treatment response? (For example, a reduction in tumour size by x cm, or a reduction in disease activity by a certain amount)	A partial response or better as defined by revised response criteria for malignant lymphoma 2007 (Revised Response Criteria for Malignant Lymphoma | Journal of Clinical Oncology (ascopubs.org)).
10. In your view, is there an unmet need for patients and health ltcare professionals in **treating relapsed or refractory classical Hodgkin's lymphoma after brentuximab Vedotin [review of TA540] ID5084 ?**	The need for treatment of relapsed or refractory (R/R) classical Hodgkin lymphoma (cHL) following brentuximab vedotin (BV) is currently met using pembrolizumab accessed via the Cancer Drugs Fund (CDF). In the absence of the availability of this treatment there is no standard of care and there would be a clear unmet need. In general, the treatment of R/R cHL is unsatisfactory and better treatments are required.

11. How is treatment for treating relapsed or refractory classical Hodgkin's lymphoma after brentuximab vedotin[review of TA540] ID5084 ? **currently treated in the NHS?** • Are any clinical guidelines used in the treatment of the condition, and if so, which? • Is the pathway of care well defined? Does it vary or are there differences of opinion between professionals across the NHS? (Please state if your experience is from outside England.) • What impact would the technology have on the current pathway of care?	For the majority of patient with R/R cHL after BV current treatment would be with pembrolizumab accessed via the CDF. Outcomes with single agent or combination chemotherapy in this setting are poor. The British Society for Haematology (BSH) Guideline on the management of R/R cHL is currently in draft form and there is no current national UK guideline in this area. Local network guidelines are available which recommend the use of pembrolizumab accessed via the CDF, for example the Thames Valley Cancer Alliance Network Site Specific Group (NSSG) guideline (PowerPoint Presentation (oxford-haematology.org.uk)) and the Northern Cancer Alliance guidance (Haematology-Cancer-Clinical-Guidelines-S12-Management-of-Classical-Hodgkin-Lymphoma.pdf (northerncanceralliance.nhs.uk)). In my experience most clinicians and lymphoma multi-disciplinary teams (MDTs) would recommend pembrolizumab via the CDF in this situation and the pathway is well defined (see the local network guidelines given as examples above). If approved the technology would continue to be used as it currently is but access would be via NICE TAG rather than the CDF.

12. Will the technology be used (or is it already used) in the same way as current care in NHS clinical practice? • How does healthcare resource use differ between the technology and current care? • In what clinical setting should the technology be used? (for example, primary or secondary care, specialist clinic) • What investment is needed to introduce the technology? (for example, for facilities, equipment, or training)	Pembrolizumab is already used in current care in this setting. It should be given in secondary care under the care of a specialist in the management of lymphoma (consultant haematologist or oncologist). Pembrolizumab is given by intravenous infusion in the setting of a haematology or oncology day care unit. No additional special facilities are required and as it is already in use via the CDF no additional training should be required.
13. Do you expect the technology to provide clinically meaningful benefits compared with current care? • Do you expect the technology to increase length of life more than current care? • Do you expect the technology to increase health-related quality of life more than current care?	Pembrolizumab via the CDF is the current standard of care. In the absence of the availability of pembrolizumab there is no current standard of care in this situation. Long term follow-up of pembrolizumab use in this situation indicates that approximately a quarter of patients can expect a complete response to pembrolizumab and of these half will remain with a complete response beyond four years, some of these patients may be cured. Achieving a complete or partial response to pembrolizumab in this situation will prolong survival and improve health-related quality of life.
14. Are there any groups of people for whom the technology would be more or less effective (or appropriate) than the general population?	No.

15. Will the technology be easier or more difficult to use for patients or healthcare professionals than current care? Are there any practical implications for its use? (For example, any concomitant treatments needed, additional clinical requirements, factors affecting patient acceptability or ease of use or additional tests or monitoring needed)	Pembrolizumab via the CDF is the current standard of care in this situation. In the absence of the availability of pembrolizumab then alternative treatments would involve one of a variety of single agent or combination chemotherapy regimens. In comparison to the adverse event profile of combination chemotherapy, pembrolizumab is better tolerated and requires less day unit attendance time and fewer monitoring tests. However, the duration of pembrolizumab is longer than treatment with combination chemotherapy regimens as up to two years of pembrolizumab treatment is available via the CDF whereas most combination chemotherapy regimens are given for four to six months.
16. Will any rules (informal or formal) be used to start or stop treatment with the technology? Do these include any additional testing?	Treatment would be started by demonstration of disease progression following BV treatment. This is usually assessed by CT or CT-PET scanning and these assessments form part of the current standard of care. Treatment is continued for two years or less if there is disease progression or relapse in that time. Most centres would restage the disease with CT or PET-CT scanning at three and six months after initiation of pembrolizumab therapy to assess response and stop treatment in the event of disease progression. Reassessment scans form a part of the management of any patient with cHL and would be used in alternative treatments so are not additional testing in that sense.

17. Do you consider that the use of the technology will result in any substantial health-related benefits that are unlikely to be included in the quality-adjusted life year (QALY) calculation? • Do the instruments that measure quality of life fully capture all the benefits of the technology or have some been missed? For example, the treatment regimen may be more easily administered (such as an oral tablet or home treatment) than current standard of care	Compared to combination chemotherapy pembrolizumab can be given less frequently (every six weeks) as a single infusion, reducing the number of hospital attendances and associated monitoring tests. Pembrolizumab has a generally favourable side-effect profile and, unlike combination chemotherapy, it is not associated with hair loss and has a lower risk of neutropenic infection as a complication of treatment.
18. Do you consider the technology to be innovative in its potential to make a significant and substantial impact on health-related benefits and how might it improve the way that current need is met? • Is the technology a 'step-change' in the management of the condition? • Does the use of the technology address any particular unmet need of the patient population?	Outside of the availability of pembrolizumab via the CDF there is no current standard of care, so this drug meets a need for the population under consideration. As immunotherapy pembrolizumab represents an alternative approach to cHL treatment when compared to combination chemotherapy and so at the time of its introduction could certainly have been regarded as a step-change in treatment.
19. How do any side effects or adverse effects of the technology affect the management of the condition and the patient's quality of life?	The principal adverse effects of pembrolizumab relate to an increased risk of auto-immune conditions. The most frequent adverse effects reported in the long-term follow up of pembrolizumab for cHL were hypothyroidism, pyrexia, fatigue and rash. The most frequent grade 3-4 adverse effects were neutropenia, pericarditis and diarrhoea. The rate of grade 3-4 AEs in the KEYNOTE-087 trial was 12.9%. In real world practice this treatment is generally well tolerated.

20. Do the clinical trials on the technology reflect current UK clinical practice? • If not, how could the results be extrapolated to the UK setting? • What, in your view, are the most important outcomes, and were they measured in the trials? • If surrogate outcome measures were used, do they adequately predict long-term clinical outcomes? • Are there any adverse effects that were not apparent in clinical trials but have come to light subsequently?	The principal clinical trial in question is KEYNOTE-087. This included three cohorts, one of which (cohort 2) reflects current UK practice for pembrolizumab use via the CDF. Cohort 2 represented 81 of 210 patients enrolled in KEYNOTE-087 (38.6%). The most important outcomes are: complete response rate, partial response rate, duration of response, progression-free survival, overall survival and rate of adverse events. These outcomes were captured in the trial. The adverse effect profile of immune checkpoint inhibitors is now better understood than at the time of the original clinical trials but remains broadly similar with lower rates of cytotoxic complications than are seen with combination chemotherapy, but higher rates of auto-immune/auto-inflammatory complications.
21. Are you aware of any relevant evidence that might not be found by a systematic review of the trial evidence?	No.
22. Are you aware of any new evidence for the comparator treatment(s) since the publication of NICE technology appraisal guidance [TA857]	The comparator referenced in TA540 was the paper by Cheah *et al*. (2016) reporting outcomes in patients relapsing after brentuximab vedotin, there was no standard approach in this patient cohort with patients receiving a variety of experimental agents and conventional chemotherapy drugs. I am not aware of further systematic studies along these lines since the publication by Cheah *et al*.

23. How do data on real-world experience compare with the trial data?	Some small retrospective studies on immune checkpoint inhibitors including pembrolizumab in this setting have been published (for instance Bair *et al.* (2019) and Gaudio *et al.* (2023)). These report broadly comparable outcomes to the KEYNOTE-087 data with overall response rates between 70 and 80% and slightly higher CR rates of approximately 45% but with small numbers of patients and both pembrolizumab and nivolumab being analysed together in these studies.

24. NICE considers whether there are any equalities issues at each stage of an evaluation. Are there any potential equality issues that should be taken into account when considering this condition and this treatment? Please explain if you think any groups of people with this condition are particularly disadvantaged. Equality legislation includes people of a particular age, disability, gender reassignment, marriage and civil partnership, pregnancy and maternity, race, religion or belief, sex, and sexual orientation or people with any other shared characteristics. Please state if you think this evaluation could • exclude any people for which this treatment is or will be licensed but who are protected by the equality legislation • lead to recommendations that have a different impact on people protected by the equality legislation than on the wider population • lead to recommendations that have an adverse impact on disabled people. Please consider whether these issues are different from issues with current care and why. More information on how NICE deals with equalities issues can be found in the NICE equality scheme. Find more general information about the Equality Act and equalities issues here.	I don't think there are any potential equality issues that need to be taken into account.

Part 2: Key messages

In up to 5 sentences, please summarise the key messages of your statement:

Pembrolizumab accessed via the CDF is the current UK standard of care for R/R cHL after brentuximab.

Pembrolizumab produces meaningful clinical responses of complete and partial remissions.

Pembrolizumab responses are associated with prolonged survival.

Pembrolizumab has a favourable resource usage profile in terms of day unit time and associated tests compared to combination chemotherapy.

Pembrolizumab has a manageable adverse effect profile with few patients needing to discontinue treatment for adverse effects.

Thank you for your time.

Your privacy

The information that you provide on this form will be used to contact you about the topic above.

☐ **Please tick this box** if you would like to receive information about other NICE topics.

For more information about how we process your personal data please see our privacy notice.

NICE National Institute for Health and Care Excellence

Single Technology Appraisal

Pembrolizumab for treating relapsed or refractory classical Hodgkin lymphoma [Review of TA540] [ID5084]

Clinical expert statement

Information on completing this form

In part 1 we are asking for your views on this technology. The text boxes will expand as you type.

In part 2 we are asking you to provide 5 summary sentences on the main points contained in this document.

Please do not embed documents (such as a PDF) in a submission because this may lead to the information being mislaid or make the submission unreadable. Please type information directly into the form.

Do not include medical information about yourself or another person that could identify you or the other person.

We are committed to meeting the requirements of copyright legislation. If you want to include **journal articles** in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs. For copyright reasons, we will have to return forms that have attachments without reading them. You can resubmit your form without attachments, but it must be sent by the deadline.

Combine all comments from your organisation (if applicable) into 1 response. We cannot accept more than 1 set of comments from each organisation.

Please underline all confidential information, and separately highlight information that is submitted as 'confidential [CON]' in turquoise, and all information submitted as 'depersonalised data [DPD]' in pink. If confidential information is submitted, please also send a second version of your comments with that information redacted. See Health technology evaluations: interim methods and process guide for the proportionate approach to technology appraisals (section 3.2) for more information.

The deadline for your response is **5pm** on **Thursday 8 February 2024.** Please log in to your NICE Docs account to upload your completed form, as a Word document (not a PDF).

Thank you for your time.

We reserve the right to summarise and edit comments received, or not to publish them at all, if we consider the comments are too long, or publication would be unlawful or otherwise inappropriate.

Comments received are published in the interests of openness and transparency, and to promote understanding of how recommendations are developed. The comments are published as a record of the comments we received, and are not endorsed by NICE, its officers or advisory committees.

Part 1: Treating relapsed or recurrent classical Hodgkin lymphoma and current treatment options

Table 1 About you, aim of treatment, place and use of technology, sources of evidence and equality

1. Your name	Dr Elizabeth Phillips
2. Name of organisation	University of Manchester and The Christie NHS Trust
3. Job title or position	Clinical Senior Lecturer and Honorary Consultant Haematologist
4. Are you (please tick all that apply)	☒ An employee or representative of a healthcare professional organisation that represents clinicians? ☒ A specialist in the treatment of people with relapsed or recurrent classical Hodgkin lymphoma? ☒ A specialist in the clinical evidence base for relapsed or recurrent classical Hodgkin lymphoma or pembrolizumab? ☐ Other (please specify):
5. Do you wish to agree with your nominating organisation's submission? (We would encourage you to complete this form even if you agree with your nominating organisation's submission)	☐ Yes, I agree with it ☐ No, I disagree with it ☐ I agree with some of it, but disagree with some of it ☒ Other (they did not submit one, I do not know if they submitted one etc.)
6. If you wrote the organisation submission and/or do not have anything to add, tick here. (If you tick this box, the rest of this form will be deleted after submission)	☐ Yes
7. Please disclose any past or current, direct or indirect links to, or funding from, the tobacco industry.	N/A- but none

8. What is the main aim of treatment for relapsed or recurrent classical Hodgkin lymphoma? (For example, to stop progression, to improve mobility, to cure the condition, or prevent progression or disability)	For younger/transplant-fit patients: to induce sufficient clinical response to facilitate autologous or allogeneic stem cell transplant, which is potentially curative- the ultimate aim is cure. For transplant-unfit patients: to treat symptoms (i.e. improve quality of life) and induce prolonged disease remission, i.e. prolong overall survival A small proportion of patients have durable remissions (>5 years) without further treatment and may potentially be cured with PD-1 inhibitors alone
9. What do you consider a clinically significant treatment response? (For example, a reduction in tumour size by x cm, or a reduction in disease activity by a certain amount)	The aim is to achieve at least partial response by standard criteria (most commonly Lugano 2014). Resolution of clinical symptoms alone may be sufficient in the early stages- it is reasonably common to have indeterminate response on PET, particularly within the first 12 weeks, where immune activity is indistinguishable from residual/progressive disease and may obscure response assessment- specific response criteria have been developed to account for this and determine true checkpoint inhibitor failure (Lymphoma response to immunomodulatory therapy criteria- LYRIC)
10. In your view, is there an unmet need for patients and healthcare professionals in relapsed or recurrent classical Hodgkin lymphoma?	Yes. PD-1 inhibitors are very effective drugs for the treatment of relapsed/refractory Hodgkin lymphoma but are currently only available for patients who have: 1) not had brentuximab vedotin, or 2) had both transplant and brentuximab vedotin. There are a proportion of patients who fail brentuximab-based treatment but are unable to proceed to transplant due to insufficient response. The only route to access a PD-1 inhibitor for these patients is via the CDF. They are a small but very high-risk group of patients with no other standard treatment options, who should not be excluded from an effective (and potentially curative) therapy.
11. How is relapsed or recurrent classical Hodgkin lymphoma currently treated in the NHS? • Are any clinical guidelines used in the treatment of the condition, and if so, which? • Is the pathway of care well defined? Does it vary or are there differences of opinion between professionals	BSH guidelines (Collins et al, 2014) are most commonly used, but pre-date the use of PD-1 inhibitors and are currently under revision Second-line treatment of Hodgkin lymphoma consists of multi-agent chemotherapy. In younger, fit patients (~80% patients with Hodgkin lymphoma), this is consolidated with a potentially curative stem cell transplant (SCT; usually

across the NHS? (Please state if your experience is from outside England.) • What impact would the technology have on the current pathway of care?	autologous). However, <50% of eligible patients achieve sufficient response to proceed with SCT. The remainder require third-line treatment There is significant variation in third-line treatment around the UK. Single-agent brentuximab vedotin and pembrolizumab are both available and NICE-approved. The latter is generally given first due to longer PFS compared with brentuximab (in the Keynote-204 trial) and retrospective data showing high PFS rates when combined with SCT. However, brentuximab may be used first in selected patients, e.g. those with concurrent immune conditions or where allogeneic transplant SCT is planned, due to toxicity concerns. Furthermore, many centres around the UK are now using brentuximab vedotin in combination with chemotherapy (usually brentuximab vedotin) in fit patients as third-line therapy in preference to pembrolizumab, usually as a bridge to SCT. Brentuximab-bendamustine has much higher response rates than single-agent pembrolizumab or brentuximab vedotin alone, therefore more patients are able to proceed to a potentially curative SCT (Shotton et al, ASH 2023). There is wide variation in access to brentuximab-bendamustine across the UK, which is currently being addressed as part of an NHS England Public Policy Proposal (led by myself), but it is widely used in Wales, Scotland and selected other centres. Brentuximab-bendamustine is also the recommended third-line treatment in children and teenagers/young adults with Hodgkin lymphoma according to international guidance (Daw et al, Hemasphere, 2020) A proportion of patients will fail third-line brentuximab or brentuximab-bendamustine without receiving SCT. There is no other established standard of care for these patients besides PD-1 inhibitors, which are currently unavailable outside the CDF.
12. Will the technology be used (or is it already used) in the same way as current care in NHS clinical practice? • How does healthcare resource use differ between the technology and current care?	There is no other established standard of care for these patients. In young, fit patients who fail 3 lines of therapy including brentuximab vedotin, intensive combination chemotherapy is sometimes given, which is resource-intensive, often requires inpatient admission and has significant toxicity. Other options include clinical trials or single-agent palliative chemotherapy. The evidence base for such treatments is very poor

• In what clinical setting should the technology be used? (for example, primary or secondary care, specialist clinic) • What investment is needed to introduce the technology? (for example, for facilities, equipment, or training)	By contrast, pembrolizumab is delivered at a flat dose over 30 minutes every 3 weeks as an outpatient in a secondary care setting. It is often reconstituted by nursing staff so does not require much pharmacy input. It is widely used in other cancers and as third-line therapy in Hodgkin lymphoma, so there is already considerable experience in delivering this treatment in most cancer centres.
13. Do you expect the technology to provide clinically meaningful benefits compared with current care? • Do you expect the technology to increase length of life more than current care? • Do you expect the technology to increase health-related quality of life more than current care?	Yes. Given the magnitude of the PFS benefit with pembrolizumab in Keynote-204 and evidence of sustained remissions beyond 5 years in single-arm studies of PD-1 inhibitors (pembrolizumab in Keynote-087 and nivolumab in Checkmate-205), it is very likely that these agents increase survival in relapsed/refractory Hodgkin lymphoma. Given the widespread availability of PD-1 inhibitors in developed countries, randomised trials (such as Keynote-204) are unlikely to definitively show an OS benefit, but it is likely to exist. An improvement in HRQOL is very likely, given the favourable toxicity profile and potential for symptomatic benefit.
14. Are there any groups of people for whom the technology would be more or less effective (or appropriate) than the general population?	I am not aware of any such groups
15. Will the technology be easier or more difficult to use for patients or healthcare professionals than current care? Are there any practical implications for its use? (For example, any concomitant treatments needed, additional clinical requirements, factors affecting patient acceptability or ease of use or additional tests or monitoring needed)	It is usually easier to deliver that further cytotoxic chemotherapy, albeit noting that treatment duration is longer, whereas chemotherapy is usually a fixed duration of < 6 months. Some UK clinicians will give pembrolizumab on a 6-weekly basis (400mg) rather than 3-weekly once an initial response has been established to minimise resource use. The flat dosing and short infusion times mean that it is straightforward to deliver, requiring less pharmacy input than weight/BSA-based cytotoxic chemotherapy.

	The supportive care needs for patients receiving pembrolizumab are minimal- growth factors and prophylactic antimicrobials are not required, unlike with most cytotoxic chemotherapy regimens. A small proportion of patients require subsequent endocrine replacement therapy due to immune-related adverse events (largely levothyroxine, less commonly hydrocortisone)- these agents are inexpensive
16. Will any rules (informal or formal) be used to start or stop treatment with the technology? Do these include any additional testing?	Ideally the rules should follow those for pembrolizumab in Hodgkin lymphoma in its currently approved indications (i.e. for brentuximab-naïve patients), for consistency, although there are some caveats: As a clinician, I would like to see inclusion of patients with ECOG PS 2, even though these patients were not included in the pivotal trials- there is a real need for less toxic treatments in this frailer/less fit patient group. Pembrolizumab is usually much better tolerated than cytotoxic chemotherapy. The latter is largely palliative and unlikely to be offered in the fourth-line setting- these patients do not have any other options. There is controversy about when to stop pembrolizumab. 'Indeterminate responses' are commonly seen, particularly at early points during treatment, where immune activity may be indistinguishable from residual/progressive disease. There are also data to suggest that some patients may benefit from treatment beyond progression if they have low-volume asymptomatic progression. The current CDF criteria for this indication (i.e. treatment after brentuximab for transplant-naïve patients) allow for treatment until loss of clinical benefit, which is pragmatic. Otherwise, treatment should be allowed until confirmed progression- either clinically or radiologically. Early radiological progression often requires a confirmatory repeat scan 8-12 weeks later.
17. Do you consider that the use of the technology will result in any substantial health-related benefits that are unlikely to be included in the quality-adjusted life year (QALY) calculation? • Do the instruments that measure quality of life fully capture all the benefits of the technology or have some been missed? For example, the treatment regimen	As someone who is unfamiliar with the specifics of the health economic modelling used, I do not feel confident to answer this question

may be more easily administered (such as an oral tablet or home treatment) than current standard of care	
18. Do you consider the technology to be innovative in its potential to make a significant and substantial impact on health-related benefits and how might it improve the way that current need is met? • Is the technology a 'step-change' in the management of the condition? • Does the use of the technology address any particular unmet need of the patient population?	Yes- it is definitely a 'step-change' in the treatment of Hodgkin lymphoma, with a completely different mechanism of action to other established therapies in this condition. PD-1 inhibitors have revolutionised the treatment of many malignancies and are approved across a wide range of indications. Single-agent response rates to PD-1 inhibitors are higher in Hodgkin lymphoma than in any other cancer. These responses are seen even in heavily pre-treated patients who have failed multiple standard therapies. There is increasing recognition of the curative potential of SCT after PD-1 inhibition (see Q22).
19. How do any side effects or adverse effects of the technology affect the management of the condition and the patient's quality of life?	The majority of patients who receive pembrolizumab tolerate it very well and do not feel they are on cancer treatment at all. A minority experience immune-related adverse effects, some of which can be permanent (hypothyroidism, hypoadrenalism) requiring long-term hormone replacement therapy. This does not seem to have a major impact on the quality of life for patients that I treat, nor does it interfere with subsequent Hodgkin lymphoma therapy, if needed.
20. Do the clinical trials on the technology reflect current UK clinical practice? • If not, how could the results be extrapolated to the UK setting? • What, in your view, are the most important outcomes, and were they measured in the trials? • If surrogate outcome measures were used, do they adequately predict long-term clinical outcomes? • Are there any adverse effects that were not apparent in clinical trials but have come to light subsequently?	The analyses for this appraisal include measures of quality of life and overall survival, which are the most relevant endpoints. Overall survival analyses have also been performed using UK a real-world SACT dataset that is directly relevant to current UK practice, with access to pembrolizumab via the CDF. I am not aware of any new data on adverse effects for pembrolizumab that are not already well known and reflected in the trial data. Pembrolizumab is already licensed and widely used across a broad range of malignancies
21. Are you aware of any relevant evidence that might not be found by a systematic review of the trial evidence?	It is important to consider that most of the data in relapsed/refractory Hodgkin lymphoma are retrospective and often low quality therefore a systematic review of trial evidence will return few results. This emphasises the lack of a clear

	standard of care/alternative treatment for patients who fail both chemotherapy and brentuximab vedotin, and have not received SCT. Real-world data provide supportive evidence where trial data are lacking- see comments in Q22
22. Are you aware of any new evidence for the comparator treatment(s) since the publication of NICE technology appraisal guidance [TA540]?	The most interesting data pertain to outcomes with autologous transplant after PD-1 inhibition in Hodgkin lymphoma presented at the American Society for Haematology in December 2023- this is particularly relevant to this transplant-naïve patient group. Post-autoSCT PFS rates in a large US, real-world, retrospective dataset were >90% at 2 years for patients who received PD-1 inhibition prior to autoSCT, compared with 70-75% for patients treated with other therapies, including brentuximab and cytotoxic chemotherapy (Desai et al, ASH 2023, abstract 182). There are caveats; the data are not published in full and most patients received only one line of salvage treatment prior to SCT `in this dataset. However, smaller datasets in more heavily pre-treated Hodgkin lymphoma patients (median 4 lines of therapy prior to autoSCT) demonstrate similarly impressive outcomes with autoSCT after PD-1 inhibition, with an 18-month PFS rate of 81% (Merryman et al, Blood Advances, 2021). These outcomes surpass those with transplant after standard chemotherapy and highlight the curative potential of PD-1 inhibitors when combined with subsequent autoSCT
23. How do data on real-world experience compare with the trial data?	The technology appraisal provides a comprehensive comparison of the trial data versus real-world SACT dataset
24. NICE considers whether there are any equalities issues at each stage of an evaluation. Are there any potential equality issues that should be taken into account when considering this condition and this treatment? Please explain if you think any groups of people with this condition are particularly disadvantaged. Equality legislation includes people of a particular age, disability, gender reassignment, marriage and civil partnership, pregnancy and maternity, race, religion or	Yes. Brentuximab-bendamustine is mostly widely given to a paediatric/young adult population, where it is the recommended third-line option in international guidelines in preference to pembrolizumab (Daw et al, Hemasphere, 2020). Therefore, lack of access to PD-1 inhibitors after brentuximab failure will predominantly affect this younger population. The current criteria for access to pembrolizumab will also disadvantage anyone who received frontline brentuximab and subsequently relapsed- predominantly patients treated outside of the UK or on clinical trials

belief, sex, and sexual orientation or people with any other shared characteristics. Please state if you think this evaluation could • exclude any people for which this treatment is or will be licensed but who are protected by the equality legislation • lead to recommendations that have a different impact on people protected by the equality legislation than on the wider population • lead to recommendations that have an adverse impact on disabled people. Please consider whether these issues are different from issues with current care and why. More information on how NICE deals with equalities issues can be found in the NICE equality scheme. Find more general information about the Equality Act and equalities issues here.	I am not aware of any ethnicity or sex-related issues

Part 2: Key messages

In up to 5 sentences, please summarise the key messages of your statement:

Click or tap here to enter text.

Click or tap here to enter text.

Click or tap here to enter text.

Click or tap here to enter text.

Click or tap here to enter text.

Thank you for your time.

Your privacy

The information that you provide on this form will be used to contact you about the topic above.

☐ **Please tick this box** if you would like to receive information about other NICE topics.

For more information about how we process your personal data please see our privacy notice.

in collaboration with:

Erasmus School of Health Policy & Management

Maastricht University

Pembrolizumab for treating relapsed or refractory classical Hodgkin's lymphoma after brentuximab vedotin (review of TA540) [ID5084]]

Produced by Kleijnen Systematic Reviews (KSR) Ltd, in collaboration with Erasmus University Rotterdam (EUR) and Maastricht University Medical Center+ (UMC+)

Authors
Evangelos Danopoulos, Systematic Reviewer, KSR Ltd, United Kingdom (UK)
Sabine Grimm, Health Economist, Maastricht UMC+, Netherlands (NL)
Bram Ramaekers, Health Economist, Maastricht UMC+, NL
Teebah Abu-Zahra, Health Economist, Maastricht UMC+, NL
Jiongyu Chen, Heath Economist/Systematic Reviewer, KSR Ltd, UK
Caro Noake, Information Specialist, KSR Ltd, UK
Nigel Armstrong, Health Economics Manager, KSR Ltd, UK
Manuela Joore, Health Economist, Maastricht UMC+, Netherlands (NL)
Robert Wolff, Managing Director, KSR Ltd, UK

Correspondence to
Robert Wolff, Managing Director
Kleijnen Systematic Reviews Ltd
Unit 6, Escrick Business Park
Riccall Road, Escrick
York, United Kingdom
YO19 6FD

Date completed 23/11/2023

Source of funding: This report was commissioned by the National Institute for Health Research (NIHR) Evidence Synthesis Programme as project number STA 13/61/28.

Declared competing interests of the authors 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 Evidence Synthesis Programme. Any errors are the responsibility of the authors.

This report should be referenced as follows:

Danopoulos E, Grimm S, Ramaekers B, Abu-Zahra T, Chen J, Noake C, Armstrong N, Joore M, Wolff R. Pembrolizumab for treating relapsed or refractory classical Hodgkin's lymphoma after brentuximab vedotin (review of TA540) [ID5084]: a Single Technology Assessment. York: Kleijnen Systematic Reviews Ltd, 2023.

Contributions of authors

Evangelos Danopoulos 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, Teebah Abu-Zahra, and Nigel Armstrong acted as health economists on this assessment, critiqued the Company's economic evaluation and contributed to the writing of the report. Jiongyu Chen acted as systematic reviewer, critiqued the clinical effectiveness methods and evidence 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 acted as health economist on this assessment, critiqued the company's economic evaluation, contributed to the writing of the report and provided general guidance. Robert Wolff 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 event
AEOSI	Adverse events of special interest
AlloSCT	Allogeneic stem cell transplant
ASaT	All subjects as treated
ASCO	American Society of Clinical Oncology
AutoSCT	Autologous stem cell transplant
AVD	Doxorubicin, vinblastine and dacarbazine
BEACOPP regimen	Bleomycin, etoposide, doxorubicin, cyclophosphamide, vincristine, procarbazine and prednisone
BICR	Blinded independent central radiologists'
BNF	British National Formulary
BSA	Body surface area
BSC	Best supportive care
BV	Brentuximab vedotin
CADTH	Canadian Agency for Drugs and Technologies in Health
CDF	Cancer Drugs Fund
CEA	Cost effectiveness analysis
CENTRAL	Cochrane Central Register of Controlled Clinical Trials
cHL	Classical Hodgkin lymphoma
CHOP	Cyclophosphamide, doxorubicin, prednisolone, vincristine
CI	Confidence interval
CiC	Commercial in confidence
C(M)R	Complete (metabolic) response
COVID-19	Coronavirus disease 2019
CR	Complete response
CRR	Complete remission rate
CRu	Unconfirmed complete response
CS	Company submission
CSR	Clinical Study Report
CT	Computed tomography
CTCAE	Common Terminology Criteria for Adverse Events
DECC	Dexamethasone, etoposide, chlorambucil, lomustine
DHAP	Dexamethasone, cytarabine, cisplatin
DOR	Duration of response
DP	Disease progression
DSA	Deterministic sensitivity analyses
EAG	Evidence Assessment Group
EBM	Evidence Based Medicine
ECOG	Eastern Cooperative Oncology Group
EHA	European Haematology Association
eMIT	Electronic market information
Emtree	Embase subject heading terms
EORTC-QLQC30	European Organisation for Research and Treatment Cancer Quality of Life Questionnaire
EQ-5D	European quality of life-5 dimensions
ESMO	European Society for Medical Oncology
EUCTR	EU Clinical Trials Register
EUR	Erasmus University Rotterdam
FAD	Final Appraisal Document
FAS	Full analysis set
FE	Fixing errors
FV	Fixing violations

GDP	Gemcitabine, dexamethasone, cisplatin
GEM-P	Gemcitabine, cisplatin, methylprednisolone
GVD	Gemcitabine, vinorelbine, doxorubicin
GVHD	Graft Versus Host Disease
HES	Hospital Episodes Statistics
HL	Hodgkin Lymphoma
HR	Hazard ratio
HRG	Healthcare Resource Group
HRQoL	Health-related quality of life
HRU	High resource user
HTA	Health Technology Assessment
HUI	Health Utility Index
ICE	Ifosfamide, carboplatin, etoposide
ICER	Incremental cost-effectiveness ratio
IGEV	Ifosfamide, gemcitabine, vinorelbine
IPD	Individual patient data
ISPOR	International Society for Pharmacoeconomics and Outcomes Research
ITC	Indirect treatment comparison
IVAC	Cytrabine, etoposide, ifosfamide, mesna
IWG	International Working Group
KM	Kaplan-Meier
KN	KEYNOTE (trial)
KSR	Kleijnen Systematic Reviews Ltd
LY	Life year
LYG	Life years gained
MAIC	Matching-adjusted indirect treatment comparison
mAPaT	Analysis populations used to report comparator study results
MedDRA	Medical Dictionary for Regulatory Activities
MeSH	Medical subject headings
Mini-BEAM	Carmustine, etoposide, cytarabine, melphalan
MJ	Matters of judgement
MK-3475	Pembrolizumab - Keytruda®
MSD	Merck Sharp and Dohme
NA	No assessment
N/A	Not applicable
NCI	National Cancer Institute
NHS	National Health Service
NHSE	National Health Services England
NICE	National Institute for Health and Care Excellence
NIHR	National Institute for Health Research
NL	Netherlands
NR	Not reached
N/R	Not reported
NSCLC	Non-small cell lung cancer
ONS	Office for National Statistics
ORR	Objective/overall response rate
OS	Overall survival
PAS	Patient Access Scheme
PbR	Payment-by-results
PD	Progressive disease
PD-1	Programmed death 1 protein
PD-L1	Programmed death ligand 1
PET	Positron Emission Tomography
PFS	Progression-free survival

PMitCEBO	Prednisolone, mitoxantrone, cyclophosphamide, etoposide, bleomycin and oncovin
P(M)R	Partial (metabolic) response
PR	Partial response
PRESS	Peer Review of Electronic Search Strategies
PRO	Patient-reported outcome
PSA	Probabilistic sensitivity analyses
PSS	Prescribed/Personal Social Services
PSSRU	Personal and Social Services Research Unit
Q3W	Every 3 weeks
Q6W	Every 6 weeks
QALY	Quality-adjusted life year
QoL	Quality of life
RCR	Royal College of Radiologists
RCT	Randomised controlled trial
RDI	Elative dose intensity
RR	Response rate
R/R cHL	Relapsed or refractory classical Hodgkin lymphoma
RVIG	Gemcitabine, ifosfamide, mesna, prednisolone, rituximab, vinorelbine
SACT	Systemic anti-cancer therapy
SAE	Serious adverse event
ScHARR	Sheffield Centre for health and Related Research
SCT	Stem cell transplant
SD	Stable disease; standard deviation
SE	Standard error
SEE	Structured expert elicitation
SIGN	Scottish Intercollegiate Guidelines Network
SLR	Systematic literature review
SMDM	Society for Medical Decision Making
SmPC	Summary of product characteristics
SoC	Standard of care
STA	Single Technology Appraisal
TA	Technology Appraisal
ToT	Time on treatment
TRAEs	Treatment-related adverse events
UK	United Kingdom
UMC+	University Medical Center+
US	United States
USA	United States of America
VAS	Visual analogue scale

Table of Contents

Table of Tables

Table of Figures

1. EXECUTIVE SUMMARY

This summary provides a brief overview of the key issues identified by the Evidence Assessment Group (EAG) as being potentially important for decision making. If possible, it also includes the EAG's preferred assumptions and the resulting incremental cost-effectiveness ratios (ICERs).

Section 1.1 provides an overview of the key issues. Section 1.2 presents the key model outcomes. Section 1.3 discusses the decision problem, Section 1.4 issues related to the clinical effectiveness, and Section 1.5 issues related to the cost effectiveness. A summary in presented in Section 1.7.

Background information on the condition, technology and evidence and information on key as well as non-key issues are in the main EAG report, see Sections 2 (decision problem), 3 (clinical effectiveness) and 4 (cost effectiveness) for more details.

All issues identified represent the EAG's view, not the opinion of the National Institute for Health and Care Excellence (NICE).

1.1 Overview of the EAG's key issues

Table 1.1: Summary of key issues

Issue #	Summary of issue	Report Sections
1	Choice of comparators: There is a misalignment of comparators to the NICE final scope and a lack of evidence to support the choice of basket of comparators.	2.3
2	Quality of SLR: The searches used in the SLR for clinical evidence present serious limitations that could have led to missing potentially relevant records. There are further inconsistencies and a lack of clarity of the methods used to identify and select evidence in this CS as well as adopting eligibility criteria that could exclude relevant records limit the output and quality of the SLR.	3.1, 3.1.2, 3.2
3	The recommended data collection from the TA540 has not been fulfilled as the majority of the presented outcomes of the SACT dataset do not differentiate between alloSCT and autoSCT.	2.4, 3.2.1.2, 3.2.5.3
4	There are extensive uncertainties in the entirety of the ITC analyses. The naïve ITCs, MAIC and Bucher ITCs have major limitations, and none can be considered fully reliable for decision making.	3.3, 3.4
5	The Company model structure is inconsistent with good modelling practices.	4.2.2
6	The composition and proportions of the SoC in the comparator arm.	4.2.4
7	Uncertain comparative effectiveness.	4.2.6
8	Uncertain duration of treatment effect.	4.2.6
9	KEYNOTE-087 utilities estimated through a mixed effects model.	4.2.8
10	Uncertainty about subsequent therapies.	4.2.9

alloSCT = allogeneic stem cell transplant; autoSCT = autologous stem cell transplant; CS = company submission; ITC = indirect treatment comparison; MAIC = matching-adjusted indirect treatment comparison; NICE = National Institute for Health and Care Excellence; SACT = systemic anti-cancer therapy; SLR = systematic literature review; SoC = standard of care; TA = Technology Appraisal

The key differences between the Company's preferred assumptions and the EAG's preferred assumptions are related to the comparators to be included in standard of care (SoC), the comparative effectiveness estimates and assumptions, and the health-related quality of life (HRQoL) estimates used.

1.2 Overview of key model outcomes

National Institute for Health and Care Excellence (NICE) Technology Appraisals (TAs) compare how much a new technology improves length (overall survival) and quality of life (QoL) in a quality-adjusted life year (QALY). An ICER is the ratio of the extra cost per QALY gained.

Overall, the technology is modelled to affect QALYs by:

- Increased QALYs for pembrolizumab, by impacting HRQoL pre-landmark, when cured after stem cell transplant (SCT) and when not cured after SCT; and adverse event (AE) and SCT disutilities (QALY increased by [redacted] compared with SoC without the severity modifier).
- Increased life years gained (LYG) for pembrolizumab up to the landmark but especially after the landmark (LYG increased by 8.72 years compared with SoC overall).

Overall, the technology is modelled to affect costs by:

- Increased drug costs (additional costs of [redacted] compared with SoC).
- The higher health state costs (additional costs of [redacted] compared with SoC).
- Cost-savings in terminal care, AE and subsequent treatment costs (cost-saving of [redacted] compared with SoC).

The following parameters were identified as most influential on the cost effectiveness of pembrolizumab versus SoC:

- Overall survival (OS) hazard ratio (HR) up to landmark: SoC versus pembrolizumab.
- Applying treatment waning (years 5 to 7).
- Probability that SCT will be curative: pembrolizumab.

Based on the Company's scenario analyses, modelling assumptions that have the greatest effect on the ICER were:

- All SoC and subsequent treatment costs were removed.
- Exponential survival curve after landmark.
- Treatment waning effect on no/failed SCT OS.

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

The decision problem addressed in the company submission (CS) is broadly in line with the final scope issued by NICE in terms of population, intervention and outcomes. However, the comparators were not completely aligned with the NICE remit, see Table 1.2.

Table 1.2: Key issue 1: Choice of comparators

Report Section	**2.3**
Description of issue and why the EAG has identified it as important	The Company did not adhere to the list of comparators in the NICE final scope. Instead, they reported a different list of comparators for the decision problem and an additional list of comparators in a basket of treatments used in the economic modelling. There is a lack of clarity on how these lists were compiled and a lack of evidence to support their use. Two observational studies as well as expert clinical opinion, elucidated by the Company, were cited as the source of data. Nevertheless, the data presented in these sources do not align to the lists compiled by the Company. In addition, BSC was not included in the comparators at all.

Report Section	2.3
What alternative approach has the EAG suggested?	The lists of comparators defined in the NICE final scope should be used if possible. Should the Company wish to explore further comparators they should be backed up by hard evidence.
What is the expected effect on the cost effectiveness estimates?	Unclear.
What additional evidence or analyses might help to resolve this key issue?	Adhere to the comparators listed in the NICE final scope. Alternatively use the interventions reported in Cheah (2016) or the interventions reported in Eyre (2017) that were administered subsequently to brentuximab vedotin.
BSC = best supportive care; EAG = Evidence Assessment Group; NICE = National Institute for Health and Care Excellence	

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

The EAG identified three further concerns with the evidence presented on the clinical effectiveness. Two concerns relate to the execution of the systematic literature review (SLR), namely limitations of the searches used in the SLR as well as inconsistencies and lack of clarity of the methods used to identify and select evidence (see Table 1.3). Regarding the presented evidence, the recommended data collection from Technology Appraisal (TA) 540 whose review this Single Technology Appraisal (STA) is, has not been fulfilled (see Table 1.4). In addition, extensive uncertainties were identified in the entirety of the indirect treatment comparison (ITC) analyses (see Table 1.5).

Table 1.3: Key issue 2: Quality of SLR

Report Section	3.1, 3.1.2, 3.2
Description of issue and why the EAG has identified it as important	A series of limitations were identified in the search strategy of the SLR which might have led to not retrieving all relevant records. In addition, excluding records due to language might have excluded further relevant records. The source of identification of certain evidence in the CS is not clear. This STA is a review of TA540 that was concluded in 2018, the CS does not contain any newly published evidence (apart from a SACT dataset, which only informs the effectiveness of pembrolizumab). Missing potentially relevant records/studies is even more important as the underlying basis of the analysis is that no new evidence have been published since TA540.
What alternative approach has the EAG suggested?	The EAG has proposed that a more sensitive search strategy should be executed so that all potentially relevant records should be retrieved. In addition, the eligibility criteria should be reviewed: the exclusion criterion should be removed; the list of comparators should align to the NICE final scope. The Company did not proceed with the EAG's suggestions stating that they believe their strategy to be overall competent.
What is the expected effect on the cost effectiveness estimates?	Unclear.
What additional evidence or analyses might help to resolve this key issue?	Update and re-run the search strategy to maximize sensitivity. Update the eligibility criteria, aligned to NICE final scope and best practice guidelines for SLRs.

Report Section	3.1, 3.1.2, 3.2
	In addition, historical SACT data should be sought to inform the effectiveness of the standard of care most relevant to UK clinical practice.

CS = company submission; EAG = Evidence Assessment Group; NICE = National Institute for Health and Care Excellence; SACT = systemic anti-cancer therapy; SLR = systematic literature review; STA = Single Technology Appraisal

Table 1.4: Key issue 3: Misaligned outcomes from the SACT dataset

Report Section	2.4, 3.2.1.2, 3.2.5.3
Description of issue and why the EAG has identified it as important	In the FAD for TA540 recommendations for data collection the proposals for further data collection included: time to alloSCT; duration of treatment with pembrolizumab before alloSCT; long-term follow-up of people having pembrolizumab with or without subsequent alloSCT (in particular, collection of data on overall survival). These data were not presented in the CS and were, according to the company, not available in the SACT dataset. Instead, data for autoSCT and alloSCT combined were presented for the outcome time to SCT. In the SACT dataset out of the 30.2% of patients who received an SCT, 35.4% received autoSCT and 64.6% received alloSCT. Combining these patients in an aggregate outcome distorts the interpretation of the data.
What alternative approach has the EAG suggested?	The EAG requested that these data should be provided.
What is the expected effect on the cost effectiveness estimates?	Unclear.
What additional evidence or analyses might help to resolve this key issue?	The Company has stated that the SACT dataset outcomes shared by the NHSE did not differentiate between the two types of SCT. It is unclear to the EAG if such evidence has been captured.

alloSCT = allogeneic stem cell transplant; autoSCT = autologous stem cell transplant; CS = company submission; EAG = Evidence Assessment Group; FAD = Final Appraisal Document; NHSE = National Health Services England; SACT = systemic anti-cancer therapy; SCT = stem cell transplant; TA = Technology Appraisal

Table 1.5: Key issue 4: Major uncertainties in the ITC analyses

Report Section	3.3, 3.4
Description of issue and why the EAG has identified it as important	There are extensive uncertainties in the entirety of the ITC analyses. Seven different OS HR estimates were presented including naïve ITCs, MAICs and Bucher ITCs, all of which have major limitations, and none can be considered fully reliable for decision making. The Company has chosen a Bucher ITC using data from KEYNOTE-087 and TA524. The EAG does not consider this estimate to be the most appropriate as both data sources have key limitations. KEYNOTE-204 examines pembrolizumab compared with BV which is not a comparator relevant to this STA, in addition TA524 which examined BV also did not use comparators relevant to this STA.

Report Section	3.3, 3.4
What alternative approach has the EAG suggested?	The EAG suggests that the naïve-ITC of SACT versus Cheah (2016) was the most appropriate choice, albeit with its own limitations.
What is the expected effect on the cost effectiveness estimates?	Unclear.
What additional evidence or analyses might help to resolve this key issue?	Adopt the ITC proposed by the EAG for the base-case scenario.

BV = brentuximab vedotin; EAG = Evidence Assessment Group; HR = hazard ratio; ITC = indirect treatment comparison; MAIC = matching-adjusted indirect treatment comparison; NICE = National Institute for Health and Care Excellence; OS = overall survival; SACT = systemic anti-cancer therapy; STA = Single Technology Appraisal; TA = Technology Appraisal

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

A full summary of the cost effectiveness evidence review conclusions can be found in Section 6.4 of this report. The Company's cost effectiveness results are presented in Section 6, the EAG's summary and detailed critique in Section 4, and the EAG's amendments to the Company's model and results are presented in Section 6. The key issues in the cost effectiveness evidence are discussed in the issue Tables below.

Table 1.6: Key issue 5: The Company model structure is inconsistent with good modelling practices

Report Section	4.2.2
Description of issue and why the EAG has identified it as important	The current model structure is inconsistent with good modelling practices as it is suboptimal from a transparency perspective as outcomes are not estimated explicitly and secondly, it violates the homogeneity within health states assumption. Although the EAG acknowledges that some degree of within-state heterogeneity is almost always present, the current model structure, pre-landmark consists of an alive state including patients that had no, failed, and successful SCT. Given SCT is an important (if not the main) mechanism through which pembrolizumab affects outcomes as well as in the disease pathway, the EAG is concerned that the Company's deviation from best modelling practices (i.e., the within-state heterogeneity in the Company's model structure) might produce substantially biased results.
What alternative approach has the EAG suggested?	Alternative model structure (suggested by the EAG in clarification question B6d) with three health states: i. no/failed SCT (without distinction before/after an arbitrary landmark) ii. successful SCT iii. death This model structure is simple, transparent, adheres to best modelling practices and does not require tunnel states. Moreover, compared with the Company's model structure, it would not necessitate using an arbitrary landmark point and allow including the probability of transitioning to successful SCT every cycle. The Company indicated they have "have no objections to the suggested structure and we were

	initially interested in implementing something like this, which is why we requested the "time to SCT or death" data from SACT". Moreover, this is more consistent with a) the original TA540 FAD data collection recommendations that mainly focused on (time to) SCT as well as b) the importance of SCT in the disease pathway and mechanism through which pembrolizumab affects outcomes. Hence, it would be very informative to estimate "time to SCT or death" and incorporate this explicitly in the economic model.
What is the expected effect on the cost effectiveness estimates?	An alternative model structure could substantially affect the estimated ICER, however, the magnitude and direction of the impact is unclear (the Company did not explore the alternative model structure in response to clarification question B6).
What additional evidence or analyses might help to resolve this key issue?	The EAG believes an alternative model structure (see suggestion above) should be adopted/explored to address the current decision problem, as requested in clarification question B6.
EAG = Evidence Assessment Group; FAD = Final Appraisal Document; ICER = incremental cost-effectiveness ratios; SCT = stem cell transplant; SACT = systemic anti-cancer therapy; TA = Technology Appraisal	

Table 1.7: Key issue 6: The composition and proportions of the SoC in the comparator arm

Report Section	**4.2.4**
Description of issue and why the EAG has identified it as important	There is uncertainty in how the comparator is reflected in the economic model. The Company deviated from the NICE final scope by creating a blended comparator reflecting SoC. The Company assumed equal proportions for all the regimens that were included in SoC. This will affect costs, and it is unknown to what extent.
What alternative approach has the EAG suggested?	In base-case proportions of treatment in the comparator were amended as the following: bendamustine 23%, mini-BEAM 23%, gemcitabine 12%, radiotherapy 12%, chemotherapy, oral chemotherapy, and ICE 10% each.
What is the expected effect on the cost effectiveness estimates?	With amending the proportion of treatments in the SoC arm, the ICER decreased, but uncertainty remains.
What additional evidence or analyses might help to resolve this key issue?	Further real-world evidence on this patient population and their treatment use.
EAG = Evidence Assessment Group; ICE= ifosfamide, carboplatin, etoposide; ICER = incremental cost-effectiveness ratios; mini-BEAM = carmustine, etoposide, cytarabine, melphalan; NICE = National Institute for Health and Care Excellence; SoC= standard of care	

Table 1.8: Key issue 7: Uncertain comparative effectiveness

Report Section	**4.2.6**
Description of issue and why the EAG has identified it as important	New SACT evidence only provides effectiveness estimates for the pembrolizumab arm of model, and this differs substantially from the data collected in KEYNOTE-087. There is no new comparative evidence, and all indirect comparisons are extremely uncertain and likely biased. There is no relevant evidence to inform the post-landmark relative treatment effect for people in the no/failed SCT state. The OS in the cured SCT state is also uncertain.

What alternative approach has the EAG suggested?	Use in base-case ITC SACT versus Cheah (2016) and explore as an alternative scenario the use of KEYNOTE-087 baseline OS with MAIC KEYNOTE-087 versus Cheah (2016). In base-case use exponential for post-landmark OS. Explore in scenario exponential for pre-landmark OS. Explore in scenario a 1.5 multiplier to general population mortality in cured SCT state (might be reasonable for base-case). Explore in a scenario the effect of equal proportions of having SCT and equal probability of cure after SCT.
What is the expected effect on the cost effectiveness estimates?	With almost all performed alternative analyses, the ICER increases (see results). The expected effect of having better comparative data is unknown.
What additional evidence or analyses might help to resolve this key issue?	Comparative effectiveness evidence: the use of historical SACT data to inform comparator effectiveness (OS) and composition as well as probability of SCT and probability of cure after SCT would be an improvement over the indirect comparisons and data sources currently used. Enabling pre-landmark treatment waning would help with less extreme scenario analysis than the one performed by the EAG.
EAG = Evidence Assessment Group; ICER = incremental cost-effectiveness ratio; ITC = indirect treatment comparison; MAIC = matching-adjusted indirect treatment comparison; OS = overall survival; SACT = systemic anti-cancer therapy; SCT = stem cell transplant	

Table 1.9: Key issue 8: Uncertain duration of relative treatment effect

Report Section	**4.2.6**
Description of issue and why the EAG has identified it as important	Uncertain duration of relative treatment effect. The duration of treatment effect is unknown. This can have a significant on the ICER.
What alternative approach has the EAG suggested?	In base-case assume no treatment effect post-landmark. Explore in a scenario no relative treatment effect on OS in the pre-landmark health state.
What is the expected effect on the cost effectiveness estimates?	With the performed alternative analyses, the ICER increases (see results).
What additional evidence or analyses might help to resolve this key issue?	Enable in the model treatment waning in the pre-landmark health state and provide expert opinion.
EAG = Evidence Assessment Group; ICER = incremental cost-effectiveness ratios; OS = overall survival	

Table 1.10: Key issue 9: KEYNOTE-087 utilities estimated through a mixed effects model

Report Section	**4.2.8**
Description of issue and why the EAG has identified it as important	Based on clarification responses B18, it became clear that the utility values reported in CS Table 42 were “simple naïve means”, i.e., averaging the longitudinal utility data without considering the missing data (i.e., complete case analysis). Unfortunately, in response to clarification question B18b, the Company did not provide utility values estimated based on the pivotal trial (KEYNOTE-087) using a mixed effects model as requested. Health state utility values based on KEYNOTE-087 are informative for the model and validating the

Report Section	4.2.8
	health state utility values used. Notably, in the original TA540, the Company provided a mixed effects model to estimate utility values based on KEYNOTE-087 by responses status.
What alternative approach has the EAG suggested?	Providing similar analyses as done based on KEYNOTE-204 in response to clarification question B19 (but without the treatment covariate) would be informative and could potentially be used to inform the pembrolizumab utility value.
What is the expected effect on the cost effectiveness estimates?	The impact of including KEYNOTE-087 utility values, on the estimated ICER is unclear.
What additional evidence or analyses might help to resolve this key issue?	See above.
EAG = Evidence Assessment Group; CS = company submission; ICER = incremental cost-effectiveness ratios; TA = Technology Appraisal	

Table 1.11: Key issue 10: Uncertainty about subsequent therapies

Report Section	4.2.4
Description of issue and why the EAG has identified it as important	The Company assumed a lower subsequent therapy proportion in the pembrolizumab arm compared to the SoC arm. The proportion of patients who receive subsequent therapy in both arms was informed using the proportions in KEYNOTE-204, with the BV arm representing the SoC arm in this submission.
What alternative approach has the EAG suggested?	In a scenario analysis, the subsequent treatment costs in both arms were set to 0.
What is the expected effect on the cost effectiveness estimates?	With assuming a cost of 0 for the subsequent treatments in both treatment arms, the ICER increased.
What additional evidence or analyses might help to resolve this key issue?	Further real-world evidence on the actual proportion of subsequent therapies to be received after pembrolizumab and the treatments included in the SoC arm.
BV = brentuximab vedotin; EAG = Evidence Assessment Group; ICER = incremental cost-effectiveness ratios; SoC = standard of care	

1.6 Summary of the EAG's view

With the majority of changes made by the EAG, the ICERs increased. While most of the individual changes only had a relatively modest impact on the ICER, when most EAG scenarios were included in one analysis, the ICER went up substantially. Hence, there is substantial uncertainty about the most appropriate ICER and there is remaining uncertainty about the model structure and its impact that could not be fully resolved.

In conclusion, large uncertainty remains about the cost effectiveness of pembrolizumab versus SoC in relapsed or refractory classical Hodgkin lymphoma (R/R cHL), which can be at least partly resolved by the Company by conducting further analyses. The EAG considers the current model structure (both in the CS and EAG base-case) flawed and this could conceivably change the ICER. Therefore, the EAG believes that neither the CS nor the EAG report contain an unbiased ICER of pembrolizumab compared

with SoC. Further data on the comparator, for example by obtaining access to historical systemic anti-cancer therapy (SACT) data (if available), could be helpful in informing relative treatment effectiveness and treatment use.

Table 1.12: Deterministic EAG base-case, no severity modifier

Technologies	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER (£/QALY)
CS base-case					
Pembrolizumab	[redacted]	[redacted]			
Standard of care	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
EAG 1 Comparator proportions amended					
Pembrolizumab	[redacted]	[redacted]			
Standard of care	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
EAG 2 Exponential in no/failed SCT OS (instead of log-logistic)					
Pembrolizumab	[redacted]	[redacted]			
Standard of care	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
EAG 3 Set no/failed SCT OS HR to 1					
Pembrolizumab	[redacted]	[redacted]			
Standard of care	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
EAG 4 ITC 5 for pre-landmark OS HR					
Pembrolizumab	[redacted]	[redacted]			
Standard of care	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
EAG 5 Pre-landmark utilities from KEYNOTE-204 Utility analysis: SoC 0.73041 and pembrolizumab 0.81573					
Pembrolizumab	[redacted]	[redacted]			
Standard of care	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
EAG 6 Set post-landmark HRQoL (pembrolizumab and SoC) equal to SoC pre-landmark					
Pembrolizumab	[redacted]	[redacted]			
Standard of care	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
EAG 7 Utility of 0.77 for successful SCT					
Pembrolizumab	[redacted]	[redacted]			
Standard of care	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
EAG base-case					
Pembrolizumab	[redacted]	[redacted]			
Standard of care	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
EAG base-case (probabilistic)					
Pembrolizumab	[redacted]	[redacted]			
Standard of care	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

CS = company submission; EAG = Evidence Assessment Group; HR = hazard ratio; HRQoL = Health-related quality of life; ICER = incremental cost-effectiveness ratio; LYG = life years gained; OS = overall survival; QALYs = quality-adjusted life years; SCT = stem cell transplant; SoC= standard of care
Results deterministic unless indicated.

2. CRITIQUE OF COMPANY'S DEFINITION OF DECISION PROBLEM

Table 2.1: Statement of the decision problem (as presented by the Company)

	Final scope issued by NICE	Decision problem addressed in the CS	Rationale if different from the final NICE scope	EAG comment
Population	People with relapsed or refractory classical HL who have had BV and cannot have autoSCT.	As per final scope. Adults with relapsed or refractory classical HL who have relapsed after or not responded to treatment with BV and are ineligible for autoSCT. Note of clarification: pembrolizumab was not recommended for treating R/R cHL in adults who have had autoSCT and BV (TA540[1]) and so the population of interest to this STA are those who transplant naïve.	N/A	The population is narrower than the one defined in the final scope issued by NICE[2], which is discussed in Section 2.1.
Intervention	Pembrolizumab	As per final scope	N/A	The EAG has no comments.
Comparator(s)	• Single or combination chemotherapy including drugs such as gemcitabine, vinblastine and cisplatin • BSC	MSD recognise that patients with R/R cHL who are not suitable for SCT after BV have few treatment options, and most people are likely to receive single-agent chemotherapy in this setting. However, no new evidence on treatment post BV was identified. For clinical effectiveness, MSD have considered the SoC to be as set out in Cheah (2016)[3], including:	The new SLR for studies on clinical effectiveness, and updates to the economic SLRs that formed the basis of MSD's original submission to TA540[1], identified no new study evaluating the comparators specified by NICE. Thus, MSD consider that Cheah (2016)[3] remains the most relevant study to generate estimates of comparative clinical effectiveness versus pembrolizumab for those not responding to BV. As noted in the original submission, Cheah (2016)[3]	The comparators are not aligned with the final scope issued by NICE[2], as detailed in Section 2.3.

	Final scope issued by NICE	Decision problem addressed in the CS	Rationale if different from the final NICE scope	EAG comment
		• Investigational agent • Gemcitabine • Bendamustine • Other alkylatory • BV retreatment • Platinum based • AutoSCT • Other To inform the economic model, MSD have created a blended comparator (Table 50 of the CS[4]) based on Cheah (2016)[3], Eyre (2017)[5] and expert opinion[6].	reported combined outcome data for all included chemotherapy regimens, and the lack of IPD for any of the treatments precludes MSD from providing estimates of comparative effectiveness for pembrolizumab versus specific regimens. MSD understand the term BSC to mean "no active treatment". Based on feedback from clinicians this is not a comparator of interest and is therefore not included in the submission.	
Outcomes	The outcome measures to be considered include: • OS • PFS • RRs • Adverse effects of treatment • HRQoL • Time to alloSCT	As per final scope, with the exception that time to SCT is not broken down by type of transplant (autologous versus allogeneic) as separate data were not available.	N/A	The outcome time to alloHSCT is not included in the CS, as detailed in Section 2.4.

Based on Table 1 and pages 10 - 11 of the CS[4]

autoSCT = autologous stem cell transplant; BSC = best supportive care; BV = brentuximab vedotin; CS = company submission; EAG = Evidence Assessment Group; HL = Hodgkin lymphoma; IPD = individual patient data; MSD = Merck Sharp and Dohme; N/A = not applicable; NICE = National Institute for Health and Care Excellence; R/R cHL = relapsed or refractory classical Hodgkin lymphoma; SLR = systematic literature review; SoC = standard of care; STA = Single Technology Appraisal; TA= Technology Appraisal

2.1 Population

The population defined in the scope is: people with relapsed or refractory classical Hodgkin lymphoma (R/R cHL) who have had brentuximab vedotin (BV) and cannot have autologous stem cell transplant auto(SCT).[2] The population in the company submission (CS) is limited to adults with R/R cHL who have relapsed after or not responded to treatment with BV and are ineligible for autoSCT as well as transplant naïve. According to the Company they have narrowed the scope addressed by the decision problem to only transplant naïve patients because pembrolizumab was not recommended for treating R/R cHL in adults who have had autoSCT and BV in Technology Appraisal (TA) 540[1] whose review this submission is.

According to the European marketing authorisation, last updated in 2022[7], *"pembrolizumab is indicated for the treatment of adult and paediatric patients aged 3 years and older with R/R cHL who have failed autoSCT or following at least two prior therapies when autoSCT is not a treatment option"*.

The recommendation from TA540[1] was for use within the Cancer Drugs Fund (CDF) as an option for treating R/R cHL in adults who have had BV and cannot have autoSCT, only if:

- Pembrolizumab is stopped after 2 years of treatment or earlier if the person has a stem cell transplant (SCT) or the disease progresses; and
- The conditions in the managed access agreement for pembrolizumab are followed.

EAG comment: There is a misalignment between the population defined by a) the National Institute for Health and Care Excellence (NICE) final scope[2], b) the decision problem addressed in the CS[4], c) the marketing authorisation[7], d) the recommendation from TA540[1] and e) the inclusion criteria for the systemic anti-cancer therapy (SACT) data:

- The decision problem is narrowed to only adult R/R cHL patients.
- The decision problem was narrowed to R/R cHL patients which are autoSCT naïve.
- The inclusion criteria for SACT were narrowed to R/R cHL patients which are autoSCT naïve and are ineligible for SCT but might be eligible for SCT after treatment.
- The marketing authorisation includes adult and paediatric R/R cHL patients (aged 3 years and older) who have failed autoSCT or following at least two prior therapies when autoSCT is not a treatment option.

The Company was asked in the clarification letter[8] to confirm and justify the above discrepancies. The Company confirmed that "*the population relevant to the decision problem is adults who have not had autologous stem cell transplant (autoSCT) and remain ineligible for autoSCT*." (page 6[9]), which is also narrower than the marketing authorisation. The Company reiterated that the decision to only focus on transplant naïve patients is based on the fact that pembrolizumab was not recommended for treating relapsed or R/R cHL in adults who had received autoSCT and BV in TA540[1].

Regarding current and future eligibility of patients for autoSCT, the Company stated that "*All patients were deemed ineligible for autoSCT at the point of treatment initiation. However, some patients could become suitable candidates for SCT, should treatment induce a sufficient level of response. Because of older age and presence of comorbidities, some patients would be deemed not suitable for SCT at any time, be that either autologous or allogeneic SCT, and would continue treatment with pembrolizumab up to 35 cycles as long as clinical benefit remains.*" (page 7[9]). In addition, the Company has confirmed that "*the goal of treatment subsequent to BV is, in some cases, to act as a bridge therapy, triggering*

sufficient response to treatment to facilitate SCT." (page 16[9]). Focusing the decision problem to only adult patients was not discussed.

2.2 Intervention

The intervention (pembrolizumab) is in line with the scope.

The recommended dose of KEYTRUDA in adults is either 200 mg every 3 weeks (Q3W) or 400 mg every 6 weeks (Q6W) administered as an intravenous infusion over 30 minutes, as outlined in the summary of product characteristics (SmPC)[10]. Therapy is to be initiated and supervised by specialist physicians experienced in the treatment of cancer.

2.3 Comparators

The description of the comparators in the NICE final scope is as follows: *"Single or combination chemotherapy including drugs such as gemcitabine, vinblastine and cisplatin; Best supportive care"*.[2] The Company has chosen to include in their comparators a list of drugs they consider to be 'standard care' based on the retrospective observational study by Cheah (2016)[3] which was also part of the evidence base for TA540[1]. This list of comparators was to include the following treatments:

- Investigational agent
- Gemcitabine
- Bendamustine
- Other alkylatory
- BV retreatment
- Platinum based
- AutoSCT
- Other

Nevertheless, the economic model was informed by a different list of comparators (see Table 50 of the CS[4]). According to the Company this list further informed by another retrospective observational Eyre (2017)[5] as well as expert opinion[6].

EAG comment:

- The Evidence Assessment Group (EAG) requested that the Company provide justification for the inconsistency in comparators with the NICE final scope and to provide the list of comparators that would be standard of care (SoC) in the United Kingdom (UK) with percentage use based on objective evidence. The Company responded that the basket of comparators was informed by four relevant sources: TA462[11], TA540[1], Eyre (2017)[5] and Cheah (2016)[3]. They went on stating that using these sources "*The eight advisors at the UK advisory board then excluded or added options so that the final list reflected those that are available and in use in UK clinical practice. There are no guidelines, practice is heterogeneous and also depends on patient level factors. The advisors were not able to provide percentages given that pembrolizumab has been the SoC for a number of years and there are only ~50 patients per year nationally in this treatment line. Prior to the introduction of pembrolizumab many patients went into clinical trials. There is no objective evidence available so we felt that an even split was the starting point that made the fewest explicit assumptions. The blend of comparators is only relevant for determining the cost in the model and so is best covered using sensitivity analyses rather than explicit breakdowns that are based on indirectly applicable populations in the literature.*" (page 8-9[9]).
  TA462[11] is NICE guidance on nivolumab for treating R/R cHL. The Company has stated in their response to the request for clarification that "*Patients previously treated with nivolumab*

are outside of the population in the decision problem." (page 11[9]) The comparators that were used by the Company in TA540[1] addressed a broader population as the NICE final scope was for R/R cHL patients who have received autoSCT and BV or BV when autoSCT is not a treatment option. In fact, the TA540[1] Final Appraisal Document (FAD) stated that Cheah (2016)[3] had limited application partly because it was conducted in the United States of America (USA) and partly because 70% of patients received SCT. The EAG inquired whether the basket of comparators was informed from the 30% of patients in Cheah (2016)[3] who had not received a SCT, and to adjust the comparators accordingly. The Company confirmed that no subgroups were considered but did not adjust the analysis as requested.

- Eyre (2017)[5], which was conducted in the UK, was mentioned in the TA540[1] FAD as an alternative, albeit also with some limitations. The EAG suggested that the Company would redo all analysis using the list of comparators based on Eyre (2017)[5] alone, to which the Company did not respond.
- The EAG further noted in the request for clarification[8] that the list of comparators provided by the Company in Table 1 of the CS[4] which is taken from Cheah (2016)[3] is not the same list of drugs used in the blended comparator for the economic model reported in Table 50 of the CS[4]. Table 50 includes further drugs: ifosfamide, carboplatin, etoposide (ICE); prednisolone, mitoxantrone, cyclophosphamide, etoposide, bleomycin and oncovin (PMitCEBO); dexamethasone, etoposide, chlorambucil, lomustine (DECC); radiotherapy and carmustine, etoposide, cytarabine, melphalan (mini-BEAM). The Company supports that these additional drugs have been informed by Eyre (2017)[5] and expert opinion. Nevertheless, some of these drugs (ICE, ifosfamide, gemcitabine, vinorelbine (IGEV)), gemcitabine, cisplatin, methylprednisolone (GEM-P)) were reported in Eyre (2017)[5] to have been administered pre-BV treatment while others are not reported in either Eyre (2017)[5] or the expert opinion report[6] at all (gemcitabine, dexamethasone, cisplatin (GDP)), (gemcitabine, vinorelbine, doxorubicin (GVD)). The Company did not respond specifically to these inquiries but rather reiterated that the final list of comparators in the treatment basket was defined by the advisors at the UK Advisory Board. The Company has provided a report for UK KEYNOTE-087 CDF exit virtual Advisory Board meeting[6]. In this report only fifth line treatments subsequent to KEYTRUDA or SoC therapy and treatments between KEYTRUDA and SCT were discussed.
- Best supportive care (BSC) was included in the NICE final scope as a comparator. The Company did not include in the CS as according to feedback from clinicians it is not a comparator of interest. The EAG requested evidence to support this decision[8] to which the Company responded "*Clinical experts engaged by MSD stated during an advisory board that, in the absence of pembrolizumab as a treatment option, active chemotherapy would typically be given as a 4th line therapy.(2) Clinical experts went on to comment that 5th line treatment (after either pembrolizumab or standard of care) would commonly consist of oral palliative chemotherapy (e.g., DECC), with radiotherapy, single agent gemcitabine or bendamustine and participation in clinical trials also potential options. Based on this advice, MSD considered it appropriate to exclude BSC (defined as "no active treatment") as a comparator in the economic model. However, we contacted a clinician for information during responding to CQs who said that up to 10% of patients might actually get BSC. A sensitivity analysis could be done on the model that reduced SoC costs by 10%, which would very slightly increase the ICER.*" (page 10[9]). The Company did not provide any evidence on why BSC should be defined as 'no active treatment'. In line with their experts panel suggestions, it is the EAG's opinion that the Company should have included the available BSC treatments in their comparator treatment basket.

The Company keeps reiterating that the choices of comparators were made according to advice received from clinical experts. It is the understanding of the EAG that the NICE final scope[2] is also informed by clinical experts as well as other stakeholders. It is unclear why the Company did not adhere to the comparators defined in the NICE final scope[2]. The misalignment of comparators to the NICE final scope, not including BSC treatments as well as the lack of evidence to support the choice of basket of comparators is a **key issue**. Further discussion is provided in the cost effectiveness Section (4.2.4).

2.4 Outcomes

The NICE final scope[2] lists the following outcome measures:

- Overall survival (OS)
- Progression-free survival (PFS)
- Response rates ((RRs)
- Adverse effects of treatment
- Health-related quality of life (HRQoL)
- Time to allogeneic stem cell transplant (alloSCT)

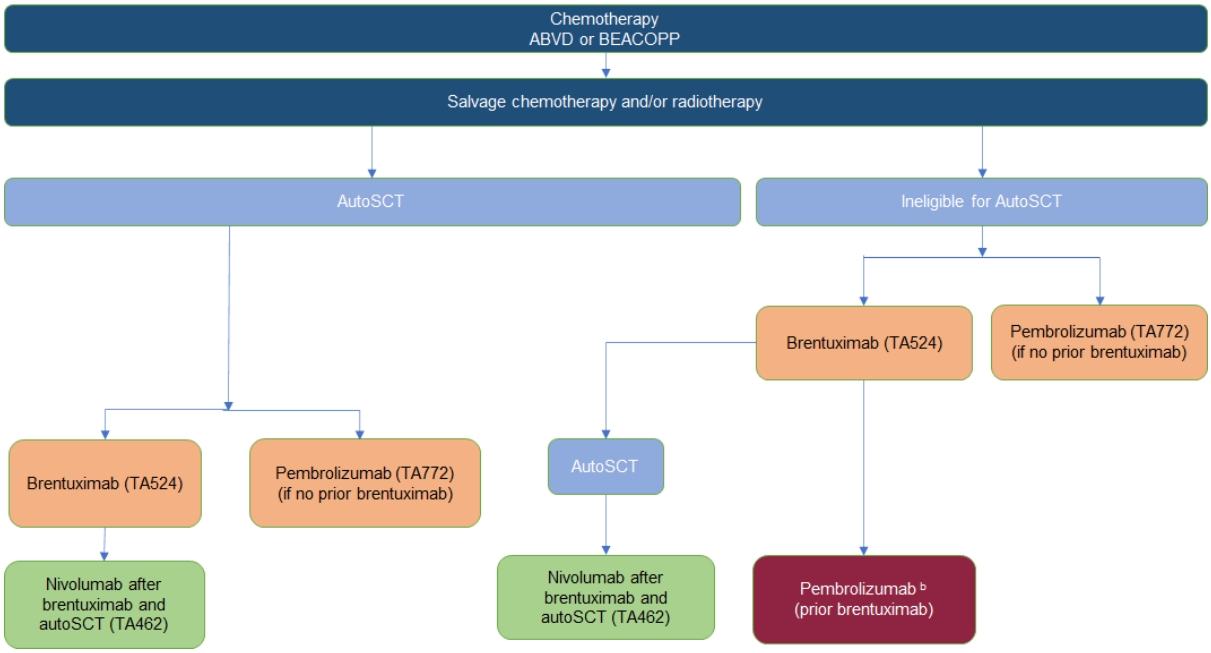
The Company stated that due to the lack of evidence specifically on alloSCT, time to stem cell transplant (SCT; both auto and allo) was used instead.

EAG comment: In the FAD for TA540[1] recommendations for future data collection included time to alloSCT. Nevertheless, the Company stated that these data were not available since "*the data provided by NHSE included patients who received both alloSCT and autoSCT (rather than alloSCT alone) as both types of SCT were used in clinical practice*." (page 17[9]). This issue is further discussed in Section 3.2.1.2 of the report.

2.5 Other relevant factors

According to the Company, the overall goal of classical Hodgkin lymphoma (cHL) treatment is the "*achievement of sufficient clinical response to enable stem cell transplantation*."[4] The Company states that there is an unmet need for R/R cHL patients. The Company's interpretation of the treatment pathway as well as the proposed position of pembrolizumab is illustrated in Figure 2.1. The Company anticipates that the use of pembrolizumab after the failed response to BV will be at the fourth-line setting. No equity or equality considerations were reported.

Figure 2.1: R/R cHL treatment pathway in the UK[a]



Based on Figure 2 of the CS[4]

[a] The Company's interpretation of the treatment pathway for R/R cHL based on NICE recommendations and available guidelines[12-14].

[b] Pembrolizumab's proposed position in the treatment pathway that is under consideration in this STA.

ABVD = doxorubicin, bleomycin, vinblastine and dacarbazine; autoSCT = autologous stem cell transplant; BEACOPP = bleomycin, etoposide, doxorubicin, cyclophosphamide, vincristine, procarbazine and prednisone; CS = company submission; NICE = National Institute for Health and Care Excellence; R/R cHL = relapsed or refractory classical Hodgkin lymphoma; STA = Single Technology Appraisal; TA = Technology Appraisal; UK = United Kingdom

EAG comment: The implications of whether pembrolizumab is considered a bridge therapy to SCT is discussed in the cost effectiveness Section of this report (4.2.2).

3. CLINICAL EFFECTIVENESS

3.1 Critique of the methods of review(s)

3.1.1 Searches

The following paragraphs contain summaries and critiques of the searches related to clinical effectiveness presented in the CS. 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.[15, 16] The EAG has presented only the major limitations of each search strategy in the report. A summary of the sources searched is provided in Table 3.1.

Table 3.1: Data sources for Appendix D: Identification, selection and synthesis of clinical evidence (as reported in CS)

Resource	Host/Source	Date Ranges	Date searched/ updated
Electronic databases			
Embase	Ovid	2010-Current	1/5/23
MEDLINE (Inc. Epub Ahead of Print, In-Process, In-Data-Review & Other Non-Indexed Citations and Daily)	Ovid	2010-Current	1/5/23
CENTRAL	EBM Reviews (Ovid)	2010-Current	1/5/23
Conferences			
Northern Light Life Sciences Conference Abstracts (searched for four named conference proceedings)	Ovid	• ASCO (2021) • American Society of Hematology (2021–2022) • EHA (2021) • ESMO (2021)	11/5/23
ASCO annual meeting	Internet	2022	Not reported
EHA Annual Congress	Internet	2022	Not reported
ESMO Congress	Internet	2022	Not reported
Trials registries			
ClinicalTrials.gov	https://clinicaltrials.gov		Not reported

Resource	Host/Source	Date Ranges	Date searched/ updated
EUCTR	https://www.clinicaltrialsregister.eu		Not reported
Additional searches			
Hand searches			
ASCO = American Society of Clinical Oncology; CENTRAL = Cochrane Central Register of Controlled Clinical Trials; CS = company submission; EBM = Evidence-Based Medicine; EHA= European Haematology Association; ESMO = European Society for Medical Oncology; EUCTR = EU Clinical Trials Register			

EAG comment:

- Searches were undertaken in May 2023 to identify observational studies and randomised controlled trials (RCTs) that evaluated treatments used in the management of patients with R/R cHL who have previously received BV and are ineligible autoSCT. The CS, Appendix D[4, 17] and the Company's response to the request for clarification[9] provided sufficient details for the EAG to appraise the literature searches. A good range of databases and grey literature including trials registers and named conference proceedings were searched. Where appropriate, strategies utilised study design filters recommended by the Scottish Intercollegiate Guidelines Network (SIGN).
- The EAG noted that the previous search strategies used in the 2017 SLR were not rerun for this update. The new searches were run in May 2023, and carried a date limit of 2010-Current. The EAG noted that the searches contained a number of limitations that affected both their reproducibility and recall. The following key areas for concern were detailed in the EAG's request for clarification[8]:
    - The EAG queried the removal of conference proceedings from the Embase strategy, the Company replied that this was intended to limit the number of irrelevant results from conference abstracts from Embase and instead chose to conduct a search of five named conferences (American Society of Clinical Oncology (ASCO), American Society of Hematology, European Hematology Association (EHA), European Society of Medical Oncology (ESMO)) from 2021-2022 via the Northern Light database and handsearching. The EAG does not accept this rationale and remains concerned that relevant papers from other conferences may have been missed.
    - The EAG was concerned about the inclusion of an English language limit in the search strategies, the Company responded that whilst it was "*not possible to estimate the exact number of non-English records that were excluded due to how the restriction was applied in the searches, it is anticipated to be fewer than 80 records*" (page 4[9]) and felt that this combined with issues regarding the uniformity of the language to describe eligibility for autoSCT, would make it unlikely that any key papers would be missed. Despite the Company's suppositions, the EAG would still refer to current best practice which 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"* (Morrison, 2012[18]) and that *"research related to language bias supports the inclusion of non-English studies in systematic reviews"*. (Egger (1997); Lefebvre (2022)[19, 20]).
    - Further to the issues above, the EAG also queried the lack of synonyms for the term 'relapsed' (suggested additional terms include: resist$ or persist$ or return$ or reocur$

or reoccur$ or recurren$ or recidiv$ or regenerat$). Of more concern was an increase in the number of facets utilised in the update searches. Whilst the Company provided justification through the change in scope, the EAG felt that these changes raised issues of specificity versus sensitivity, the more specific the search strategy, the greater the chance of missing relevant papers. In particular, the EAG was concerned that the inclusion of terms related to 'post chemo', etc. (line #20) in title and abstract (later combined with terms for 'brentuximab' OR 'stem cell transplant') was particularly restrictive and contained some redundant terms. Given the low number of relevant papers retrieved, the EAG recommended that this facet be removed, and searches rerun taking this, and the other areas highlighted, into consideration. To make this request more manageable in the limited time available, the EAG suggested using a 2017 date limit to coincide with the original review rather than the broader 2010 limit used in the update searches.

The Company's response to these requests was to state that "*For the remaining critiques on the search strategies, based on confirmatory hand searches of relevant materials (e.g., published systematic reviews, narrative reviews, and treatment guidelines), applying the suggested changes would not be expected to yield additional studies and MSD consider it unnecessary to rerun the searches as proposed by the EAG*" (page 4 [9]). In light of this response, the EAG itself reran the main Embase search (which originally retrieved 814 records) implementing the suggested changes and excluding those results that would have been retrieved by the Company's original Embase search. This resulted in an additional 3,309 (approx. 2,100 if the suggested 2017 date limit was applied) records to screen. Unfortunately, the EAG was unable to screen the results within the Single Technology Appraisal (STA) timeline, as this would be outside of the EAG remit, but the numbers suggest that the Company's approach may well have missed potentially relevant records which is **a key issue.**

3.1.2 Inclusion criteria

A SLR was conducted to identify relevant clinical evidence. Full details of the SLR search strategy, study selection process and results were reported in Appendix D of the CS[17].

The eligibility criteria used in the search strategy is presented in Table 3.2. Two reviewers independently reviewed the retrieved records both at title and abstract, and full text screening stages.

Table 3.2: Eligibility criteria used in the SLR

	Inclusion criteria	Exclusion criteria
Population	Adult (≥18 years old) patients with R/R cHL who have previously received BV and cannot undergo autoSCT	• Newly diagnosed cHL patients • R/R cHL patients deemed suitable to receive autoSCT or who have failed prior autoSCT
Interventions	Pembrolizumab Single or combination chemotherapy including drugs such as gemcitabine, vinblastine, and cisplatin	Interventions not listed
Comparators	No restrictions	N/A
Outcomes	Efficacy outcomes:	Outcomes not listed

	Inclusion criteria	Exclusion criteria
	• ORR (including CR, PR, PD, and SD, if available) • Duration of response • PFS • OS • Time-to subsequent anti-cancer therapy Safety outcomes: • AEs (any and grade ≥3) • TRAEs (any and grade ≥3) • SAEs (any and treatment-related) • Dose-limiting toxicity • Discontinuation due to AEs (overall and treatment-related) • Death due to AEs (overall and treatment-related) Patient reported outcomes i.e., HRQoL	
Time	No restrictions	N/A
Study design	RCTs Non-randomised trials Single-arm trials Prospective and retrospective observational cohort studies	Case-control studies Cross-sectional studies Case reports, and case series
Language restrictions	English language	Non-English language

Based on Table 5 of Appendix D[17]
AEs = adverse events; autoSCT = autologous stem cell transplant; BV = brentuximab vedotin; cHL = classical Hodgkin lymphoma; CR = complete response; HRQoL = health-related quality of life; N/A = not applicable; ORR = objective response rate; OS = overall survival; PD = progressive disease; PFS = progression-free survival; PR = partial response; RCT = randomised controlled trial; relapsed or refractory classical Hodgkin lymphoma; SAEs = serious adverse events; SLR = systematic literature review

EAG comment:

- The inclusion criteria of the clinical SLR are not completely aligned to the NICE final scope[2]. According to the eligibility criteria the only interventions to be included were "*Pembrolizumab; Single or combination chemotherapy including drugs such as gemcitabine, vinblastine, and cisplatin*" while "*Interventions not listed*" were to be excluded. The EAG requested that this criterion should be removed (from the updated searches) since an exhaustive list of interventions was not provided in the NICE final scope[2]. The Company refused to rerun the searches as discussed in Section 3.1.1. Regarding the proposed amendment of the criterion, they stated "*Specific intervention terms were not added to the search strategies to maximise the number of records retrieved on the available evidence base. Rescreening of studies excluded for "interventions" did not yield any additional relevant studies*." (page 12[9]). Rescreening of only the studies excluded for interventions is not a foolproof approach, and most importantly it is not following best practice SLR methodology.[21]

- The EAG pointed out in the request for clarification that excluding records due to language is against best practice in conducting SLRs and requested that this exclusion criterion would be removed in conjunction with rerunning the searches as described in the previous Section (3.1.1). The Company refused to update the criterion as suggested. The English language only restriction is inappropriate for obtaining robust evidence on the treatment and outcomes of adults with R/R cHL who have had BV and cannot have autoSCT. Potentially, relevant studies might have been missed which is a key issue.
- According to the eligibility criteria studies with patients "*Relapsed/refractory cHL patients deemed suitable to receive auto-SCT or who have failed prior auto-SCT*" should be excluded. Nevertheless, Cheah (2016)[3] included only a subgroup that fulfilled this criterion, but it was included. The EAG requested that the Company to clarify whether studies were excluded that might have had a subgroup that fulfilled this population criterion. The Company stated that "*no study with relevant subgroup data was excluded*" (page 11[9]).

3.1.3 Critique of data extraction

According to the CS, two reviewers independently extracted data. Following consensus, any discrepancies were to be resolved with the aid of a third reviewer. The EAG is satisfied with the methodological approach taken by the Company, which reflects best practice in systematic reviews.[21]

3.1.4 Quality assessment

The Company conducted the risk of bias assessment of the observational and the single arm trials using the Newcastle-Ottawa scale[22]. The assessment was executed independently by two reviewers, with a third reviewer resolving discrepancies. Quality assessment is further examined in Section 3.2.4.

3.1.5 Evidence synthesis

Given that only one single arm trial was identified and used by the Company as a source of evidence for this CS, no meta-analysis was performed. Additional data were provided from a SACT cohort collected after CDF endorsement. Two further observational studies have been used in the indirect comparison analysis. These studies and the methods used are discussed in Section 3.4.

3.2 Critique of trials of the technology of interest, their analysis and interpretation

The SLR retrieved 1,959 records, after both phases of the inclusion screening 10 records making up two unique studies were identified, KEYNOTE-013 and KEYNOTE-087. These are both single arm studies evaluating the clinical efficacy and safety of pembrolizumab in R/R cHL. KEYNOTE-013 was not included in the CS due to the small number of patients (N=9). The included publications are reported in Table 3.3.

Table 3.3: Included publications

Author	Year	Title
KEYNOTE-087		
Chen et al[23]	2017	Phase ii study of the efficacy and safety of pembrolizumab for relapsed/refractory classic hodgkin lymphoma
Chen et al[24]	2019	Pembrolizumab in relapsed or refractory hodgkin lymphoma: 2-year follow-up of keynote-087
Merck Sharp and Dohme[25]	2021	A Phase II Clinical Trial of MK-3475 (Pembrolizumab) in Subjects with Relapsed or Refractory (R/R) Classical Hodgkin Lymphoma (cHL)

Author	Year	Title
Moskowitz et al[26]	2016	Multicohort phase 2 study of pembrolizumab for relapsed/ refractory classical hodgkin lymphoma (r/r chl): Keynote-087
Moskowitz et al[27]	2016	Pembrolizumab for relapsed/refractory classical hodgkin lymphoma: Multicohort, phase 2 keynote-087 study
Moskowitz et al[28]	2016	Pembrolizumab in relapsed/refractory classical hodgkin lymphoma: Primary end point analysis of the phase 2 keynote-087 study
von Tresckow et al[29]	2019	Patient-reported outcomes in keynote-087, a phase 2 study of pembrolizumab in patients with classical hodgkin lymphoma
Zinzani et al[30]	2018	Two-year follow-up of keynote-087 study: Pembrolizumab monotherapy in relapsed/refractory classic hodgkin lymphoma
Zinzani et al[31]	2019	Three-year follow-up of keynote-087: Pembrolizumab monotherapy in relapsed/refractory classic hodgkin lymphoma
KEYNOTE-013		
Armand et al[32]	2016	Programmed death-1 blockade with pembrolizumab in patients with classical hodgkin lymphoma after brentuximab vedotin failure
Based on Table 6 of Appendix D[17]		

EAG comment:

- The observational studies Cheah (2016)[3] and Eyre (2017)[5] were not reported to have been identified in the SLR. It is not clear how these studies were retrieved or why the main clinical SLR failed to identify them.
- The study by Hanel (2023)[33] was identified by the SLR but was excluded with population given as the reason for exclusion. This was a Phase 2 trial of ibrutinib and nivolumab in patients with R/R cHL. 76.5% of the participants had prior brentuximab therapy, while 47.1% had prior autoSCT. The EAG questioned why this study was not included in the SLR. The Company responded that "*58.8% of patients had prior nivolumab. Patients previously treated with nivolumab are outside of the population in the decision problem. There are only 7 patients in the no prior nivolumab group and baseline characteristics are not reported for this group (so it is not clear how many of these patients received brentuximab). Therefore, the study was excluded as a subgroup that aligns with the decision problem cannot be identified*." (page 11[9]). This creates an inconsistency in CS since both Cheah (2016)[3] and Eyre (2017)[5] which were used in the CS were not entirely comprised of patients meeting the criteria of the decision problem and no subgroup outcomes for the relevant sub-populations were available.
- The EAG requested that the Company would consider the use of Hanel (2023)[33] in the indirect treatment comparison (ITC) analysis, given the lack of alignment of the population in Cheah (2016)[3] as well as the limitations of an observational study. The Company stated that besides the population limitations the study also did not report OS outcomes so an ITC was not possible.
- In Table 63 of Appendix M[17], the study NCT02824029 of ibrutinib is reported as an ongoing study of ibrutinib. Since this study was not reported in the SLR results the EAG inquired how it was identified. This study includes patients with relapsed or refractory HL who have failed at least two lines of prior therapy and are not eligible for autoSCT, therefore the EAG also inquired why the Company is claiming that the population does not align with the population relevant to the decision problem and asked the Company to consider using the study in the CS (efficacy and safety). The Company responded that "*A search for ongoing studies was*

carried out to fulfil the request detailed in Section B.2.11 of the NICE STA template. A simple search of ClinicalTrials.gov was carried out, using the population term of classical Hodgkin lymphoma, and applying the restriction of studies scheduled to report in the subsequent 12 months. Retrieved records were evaluated further to identify studies relevant to the decision problem. The identified study is a single arm trial of ibrutinib, and MSD are not aware of any results being published. At the time of writing, ibrutinib is not part of established clinical practice in the UK for the management of R/R cHL, therefore, had data been available, they would not have been used to address the decision problem. However, MSD considered that the study could be relevant in the future, should ibrutinib be evaluated through the NICE TA process." (page 12[9]). The United States National Institutes of Health Clinical Trial Registry (http://www.clinicaltrials.gov) was reported in Appendix D, Section D1.1[17] as one of the registries searched within the SLR. The Company's statement regarding the intervention's suitability is not entirely pertinent due to the lack of relevant data for this CS. The inconsistencies and the lack of clarity of the methods used to identify evidence in this CS as well as the weaknesses of the SLR methodology is a key issue.

3.2.1 Details of the included studies/data

3.2.1.1 KEYNOTE-087

The CS[4] identified KEYNOTE-087 as the only trial relevant to the decision problem. It a single arm study evaluating the clinical efficacy and safety of safety of pembrolizumab in R/R cHL. The trial is still running; hence data collection is ongoing. The evidence presented in the CS[4] are based on a 5-year data cutoff (data cutoff date 15 March 2021).

KEYNOTE-087 (NCT02453594) is a phase II, multicentre, multi-cohort, single arm, trial of pembrolizumab in patients with R/R cHL. The trial has three cohorts, only one of which was identified to be relevant for the CS, namely cohort 2. This cohort included patients "*Unable to achieve a CR or PR to salvage chemotherapy and did not receive autoSCT, but have relapsed after treatment with, or failed to respond to, BV*" (page 33 of CS[4]). KEYNOTE-087 is a global study, between 26-06-2015 and 21-03-2016, 210 patients were enrolled from 51 study sites. Only three of those were in the UK. Regarding Cohort 2, 81 patients were enrolled, 10 of which came from the three UK study sites.

Patients were administered 200 mg of pembrolizumab via intravenous infusion (infused over 30 minutes) on day 1 of each 3-Week cycle for up to 35 cycles. When patients achieved complete response (CR), partial response (PR) or stable disease (SD), they could continue to be treated with pembrolizumab 200 mg every 3 weeks (Q3W) for up to 2 years or until they experienced unacceptable toxicity or documented disease progression. After disease progression patients were followed for OS until death, withdrawal of consent, or the end of the study. Study methodology is presented in Table 3.4.

EAG comment:

- Although the trial was well conducted the underlying issue is that a single arm, non-comparative study has inherently serious limitations. We cannot know if the outcomes reported entirely reflect the intervention as confounding parameters might be in play. In addition, open-label trials can lead to bias in interventions as well as reporting outcomes.
- KEYNOTE-087 is a multinational trial that had three sites in the UK. Nevertheless, in cohort 2, out of the 81 patients only 10 came from the UK sites. The generalisability of the outcomes to the UK setting in terms of both patients and clinical practice is questionable.

3.2.1.2 Systemic anti-cancer therapy cohort data set

After publication of TA540[1], data were collected prospectively for pembrolizumab on specific outcomes during its time in the CDF. The cohort for the SACT dataset comprised 215/220 (98%) unique patients with CDF applications. The methodology of the SACT dataset is reported in Table 3.5. A total of 215 patients entered into treatment with pembrolizumab through the CDF.

EAG comment:

- In the FAD for TA540[1] recommendations for data collection the proposals for further data collection included:
  - proportion of people having pembrolizumab who have an alloSCT
  - time to alloSCT
  - duration of treatment with pembrolizumab before alloSCT
  - long-term follow-up of people having pembrolizumab with or without subsequent alloSCT (in particular, collection of data on OS)

  The EAG requested that the Company provide clarifications on how the above recommendations were met. The Company responded that "*The above recommendations were met through collection of data in the SACT database. MSD note that the data provided by NHSE included patients who received both alloSCT and autoSCT (rather than alloSCT alone) as both types of SCT were used in clinical practice. Data for time on treatment (ToT) was presented for all patients in the SACT, regardless of whether they received an SCT.*" (page 17[9]). The distinction between patients receiving alloSCT and autoSCT was highlighted in the FAD and is underpinned by the NICE final scope[2]. All of the recommended data collection has not been fulfilled as all of the presented data do not differentiate between them. It is not clear to the EAG how this is affecting the outcomes. This uncertainty is therefore a key issue.

- The EAG inquired whether the Company had access to SACT or historical data on any of the comparators. Collection of SACT data only on the intervention is not entirely informative. The Company confirmed that "*MSD do not have access to either contemporary or historical data from SACT for the comparators of interest, and is unaware of any such data being published*." (page 13[9]).

Table 3.4: Study methodology for KEYNOTE-087

Study	**KEYNOTE-087, NCT02453594**
Study design	Multi-centre, multi-cohort, single-arm, non-randomised study
Population	KEYNOTE-087 enrolled three cohorts, of which cohort 2 is relevant to the decision problem that is the focus of this STA. Cohort 2: participants with R/R cHL who were unable to achieve a CR or PR to salvage chemotherapy and who did not receive autoSCT, but relapsed after treatment with, or failed to respond to treatment with BV.
Inclusion criteria	• Be ≥18 years of age on day of signing informed consent • Have relapsed or refractory *de novo* cHL and meet one of the following inclusions for **Cohort 2**: o 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. Relapsed was defined as DP after most recent therapy, and refractory to treatment as failure to achieve CR or PR to most recent therapy. o Have measurable disease, which was defined as at least one lesion that can be accurately measured in at least two dimensions with spiral CT scan. Minimum measurement must be >15 mm in the longest diameter or >10 mm in the short axis o 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 o Must have a performance status of 0 or 1 on the ECOG performance scale
Exclusion criteria	Patients were excluded from participating in the trial if they met any of the following criteria: • Was participating and receiving study therapy or has participated in a study of an investigational agent and received study therapy or used an investigation device within 4 weeks of the first dose of treatment • Had a diagnosis of immunosuppression or was 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 • Had received a prior monoclonal antibody within 4 weeks prior to study day 1 or had not recovered (i.e., ≤Grade 1 or at baseline) from AEs due to agents administered more than 4 weeks earlier • Had received prior chemotherapy, targeted small molecule therapy, or radiation therapy within 2 weeks prior to study day 1 or had not recovered (i.e., ≤Grade 1 or at baseline) from AEs due to a previously administered agent

Study	KEYNOTE-087, NCT02453594
	• Had undergone 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 versus host disease • Had 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 • Had evidence of active, non-infectious pneumonitis • Had an active infection requiring intravenous systemic therapy • Was 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 • Had received prior therapy with an anti-PD-1, anti-PD-L1, anti-PD-L2, anti-CD137, or anti-CTLA-4 antibody (including ipilimumab or any other antibody or drug specifically targeting T-cell co-stimulation or checkpoint pathways).
Intervention(s)	**Pembrolizumab** 200 mg administered via intravenous infusion (infused over 30 minutes) on day 1 of each 3-week cycle for up to 35 cycles
Comparator(s)	N/A
Concomitant medication	Allowed at the discretion of the investigator and in keeping with the community standards of medical care
Prohibited concomitant medication	• 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 resulted in discontinuation of study therapy • Live vaccines within 30 days prior to the first dose of trial treatment and while participating in the trial

Study	KEYNOTE-087, NCT02453594
	Live vaccines include, but are not limited to, measles, mumps, rubella, and chicken pox. Seasonal influenza vaccines for injection are generally killed virus vaccines and are permitted, but, intranasal influenza vaccines are live attenuated vaccines, and are prohibited • Glucocorticoids for any purpose other than to modulate symptoms from an event of clinical interest of suspected immunologic aetiology
Primary outcomes	ORR by independent central review (BICR)
Secondary outcomes	• ORR by investigator assessment according to the IWG response criteria • ORR by BICR using the 5-point scale according to the Lugano Classification • CRR by BICR and investigator assessment according to the IWG response criteria • CRR by BICR using the 5-point scale according to the Lugano Classification • PFS by BICR and investigator assessment according to the IWG response criteria • DOR by BICR and investigator assessment according to the IWG response criteria • OS

Based on Table 7 and Sections B.2.3.1.2 - B.2.3.1.4-5 of the CS[4]
AEs = adverse events; autoSCT = autologous stem cell transplant; alloSCT = allogeneic stem cell transplant; BICR = blinded independent central radiologists; BV = brentuximab vedotin; CRR = complete remission rate; CS = company submission; CT = computed tomography; DOR = duration of response; DP = disease progression; ECOG = Eastern Cooperative Oncology Group; IWG = International Working Group; N/A = not applicable; ORR = overall response rate; OS = overall survival; PFS = progression-free survival; PR = partial response; R/R cHL = relapsed or refractory classical Hodgkin lymphoma; STA = Single Technology Appraisal

Table 3.5: Methodology for SACT data set

Study	SACT dataset
Study design	Real-world evidence Data captured prospectively between 25-07-2018 and 30-09-2022
Population – inclusion criteria	• Adult with histologically documented cHL • Failed two lines of chemotherapy and also BV • Has not received SCT of any kind and is ineligible for SCT after BV • Patient is **either** a candidate for future SCT if there is sufficient benefit from pembrolizumab, **or** not a candidate for SCT however good the response to pembrolizumab may be • ECOG performance status of 0 or 1

Study	SACT dataset
Population - Exclusion criteria	Previous treatment with PDL-1, PDL-2, CD137 or CTLA-4 inhibitors
Intervention(s)	Pembrolizumab 200 mg administered via intravenous infusion (infused over 30 minutes) on day 1 of each 3-week cycle for up to 35 cycles (maximum treatment duration with pembrolizumab of 2 years)
Comparator(s)	N/A
Reported outcomes specified in the decision problem	• Time to SCT • Proportion of patients who receive a SCT • OS
All other reported outcomes	• Treatment duration

Based on Table 8 of the CS[4] and Section D2.1.2 of the Appendices[17]
BV = brentuximab vedotin; cHL = classical Hodgkin lymphoma; CS = company submission; CTLA-4 = cytotoxic T-lymphocyte antigen 4; ECOG = Eastern Cooperative Oncology Group; N/A = not applicable; OS = overall survival; PDL = programmed death-ligand; SACT = systemic anti-cancer therapy; SCT = stem cell transplant

3.2.2 Statistical analysis

3.2.2.1 Statistical analysis of KEYNOTE-087 trial

The statistical methods of the trial are presented in Table 3.6. Primary and secondary study endpoints were evaluated within cohorts. The primary study endpoints involved hypothesis testing but the secondary did not. The analysis population was based on all subjects as treated (ASaT) population (i.e., all enrolled patients who received at least one dose of pembrolizumab) for the primary efficacy endpoint as well as the safety outcomes (safety population). On the other hand, the full analysis set (FAS) population were used for the rest of the efficacy endpoints. This population *"comprised all patients who: received at least one dose of pembrolizumab; had a baseline disease assessment; and had a post baseline disease assessment OR discontinued the trial due to progressive disease/drug related AE."* (page 46 of CS[4]).

Table 3.6: Analysis strategy for primary and key secondary efficacy endpoints for KEYNOTE-087

Endpoint/variable	Statistical method	Analysis population	Missing data approach
Primary outcome			
Overall response rate IWG criteria (2007)[34] Independent central review	Exact test of binomial parameter; 2-sided 95% exact CI	ASaT/FAS	People with missing data were considered non-responders
Secondary outcomes			
Overall response rate IWG criteria (2007)[34] Study site Lugano criteria (2014)[35] Independent central review	Point estimate; 2-sided 95% exact CI	ASaT/FAS	People with missing data were considered non-responders
Complete remission rate IWG criteria (2007)[34] Independent central review Study site Lugano criteria (2014)[35] Independent central review	Point estimate; 2-sided 95% exact CI	ASaT/FAS	People with missing data were considered non-responders
Progression-free survival IWG criteria (2007)[34] Independent central review Study site	Summary statistics using Kaplan–Meier method	ASaT/FAS	Censored at last assessment
Duration of response IWG criteria (2007)[34] Independent central review Study site	Summary statistics using Kaplan–Meier method	All responders	Non-responders were excluded from the analysis
Overall survival	Summary statistics using Kaplan–Meier method	ASaT/FAS	Censored at last assessment

Based on Table 11 of the CS[4]
ASaT = all subjects as treated; CI = confidence interval; CS = company submission; FAS = full analysis set; IWG = International Working Group

3.2.3 Baseline characteristics

3.2.3.1 Baseline characteristics of KEYNOTE-087 trial

Baseline characteristics of cohort 2 of the trial are presented in Table 3.7 alongside the baseline characteristics of the SACT cohort. The majority of patients were <65 years old at baseline. Patients were heavily pre-treated, with 96.3% (n=78) having ≥3 prior lines of therapy, with an overall range of 1-11 and a mean of 4 prior lines. The majority of the patients were of white race (90.1%) and ethnicity (90.1%), diagnosed with the subtype nodular sclerosis (80.2%). Eastern Cooperative Oncology Group (ECOG) performance status was balanced between 0 (54.3%) and 1 (45.7%).

Table 3.7: Baseline characteristics of KEYNOTE-087, cohort 2 and the SACT dataset

Characteristic	Cohort 2 (N=81), n, %	SACT cohort (N=215), n (%)
Gender		
Male	43 (53.1)	130 (60)
Female	38 (46.9)	85 (40)
Age, years		
Mean	42.3	N/R
SD	17.4	N/R
Median	40	N/R
Range	20 to 76	N/R
<65	66 (81.5)	N/R
≥65	15 (18.5)	N/R
20–24	16 (19.8)	75 (35)
25–29	12 (14.8)	
30–34	7 (8.6)	
35–39	4 (4.9)	
40–44	7 (8.6)	22 (10)
45–49	5 (6.2)	
50–54	8 (9.9)	37 (17)
55–59	5 (6.2)	
60–64	2 (2.5)	32 (15)
65–69	8 (9.9)	
70–74	6 (7.4)	43 (20)
75–79	1 (1.2)	
80+	NR	6 (3)
Race		
American Indian or Alaska native	1 (1.2)	N/R
Asian	4 (4.9)	N/R
Black or African American	2 (2.5)	N/R
Missing	1 (1.2)	N/R
Multi-racial	0	N/R

Characteristic	Cohort 2 (N=81), n, %	SACT cohort (N=215), n (%)
White	73 (90.1)	N/R
Ethnicity		
Hispanic or Latino	5 (6.2)	N/R
Not Hispanic or Latino	65 (80.2)	N/R
Not reported	7 (8.6)	N/R
Unknown	4 (4.9)	N/R
Race group		
White	73 (90.1)	N/R
Non-White	7 (8.6)	N/R
Missing	1 (1.2)	N/R
US region		
US	20 (24.7)	N/R
Ex-US	61 (75.3)	N/R
Disease subtype		
cHL– nodular sclerosis	65 (80.2)	N/R
cHL – mixed cellularity	10 (12.3)	N/R
cHL – lymphocyte rich	1 (1.2)	N/R
cHL – lymphocyte depleted	4 (4.9)	N/R
Missing	1 (1.2)	N/R
ECOG performance status		
0	44 (54.3)	59 (27)
1	37 (45.7)	86 (40)
2	0 (0)	16 (7)
3	0 (0)	0 (0)
Missing	0 (0)	52 (24)
Prior lines of therapy group		
≥3	78 (96.3)	N/R
<3	3 (3.7)	N/R
Prior lines of therapy		
Subjects with data	81	N/R
Mean	4.0	N/R
SD	1.7	N/R
Median	4.0	N/R
Range	1 to 11.0	N/R
Prior radiation		
Yes	21 (25.9)	N/R
No	60 (74.1)	N/R

Characteristic	Cohort 2 (N=81), n, %	SACT cohort (N=215), n (%)
Bulky lymphadenopathy		
Yes	12 (14.8)	N/R
No	69 (85.2)	N/R
Baseline B symptoms		
Yes	26 (32.1)	N/R
No	55 (67.9)	N/R
Baseline bone marrow involvement		
Yes	5 (6.2)	N/R
No	75 (92.6)	N/R
Missing	1 (1.2)	N/R

Based on Tables 9-10 of the CS[4]
CS = company submission; cHL = classical Hodgkin lymphoma; ECOG = Eastern Cooperative Oncology Group; N/R = not reported; SACT = systemic anti-cancer therapy; SD = standard deviation; US = United States

3.2.3.2 Baseline characteristics of the SACT cohort dataset

Limited baseline characteristics data were presented for the SACT cohort (Table 3.7). The majority of the patients were <40 years old (35%), but a big proportion is also >70 years old (23%). Although the inclusion criteria for SACT was ECOG performance status of either 0 or 1, according to the reported baseline characteristics a 7% had an ECOG of 2, while for a large proportion of the population (24%) ECOG was not reported.

Compared to KEYNOTE-087, the patients in the SACT cohort were older and had a worse ECOG performance status. The Company attributes these differences partially to the stringent inclusion criteria of clinical trials. In addition, the Company based partially on expert opinion[6] highlighted that the collection of SACT date took place during the coronavirus disease 2019 (COVID-19) pandemic, which might have affected the typical treatment pathway due to greater concerns around immune suppression, while there were further concerns around the quality of collected data as well as the generalisability to the UK setting.

EAG comment: The EAG requested that the Company would provide a breakdown of previous lines of therapy for the participants of KEYNOTE-087. The Company only provided three OS Kaplan-Meier (KM) plot figures for participants ≤3 versus >3 prior lines of therapy (n=37 versus n=44), with 4 prior lines of therapy (n=25), and with ≥5 prior lines of therapy (n=18). The OS outcomes are discussed in Section 3.2.5.5. According to the CS, pembrolizumab is placed at the fourth line of therapy. It is still not clear to the EAG how many participants had 3 previous lines of therapy at the time they were treated with pembrolizumab. The EAG also requested that a breakdown of previous lines of therapy would be provided for the SACT cohort. The Company responded that these data were not available from National Health Services England (NHSE). Implications of previous lines of therapy are discussed in Section 3.2.5.6.

3.2.4 Risk of bias assessment of the KEYNOTE-087 trial

The quality assessment of KEYNOTE-087 was based on the Newcastle–Ottawa scale[22]. A score of 6 was awarded to the study out of a maximum of 9. A summary of the assessment is presented in Table 3.8.

Table 3.8: Quality assessment of KEYNOTE-087

Newcastle–Ottawa scale domain	KEYNOTE-087 score
Selection	
Representativeness of cohort	*
Selection of Controls	Not applicable
Ascertainment of exposure	*
Outcome of interest	*
Comparability of cohorts	Not applicable
Outcome bias	
Outcome assessment	*
Duration of follow-up	*
Loss to follow-up	*
Based on Table 12 of the CS[4] CS = company submission	

EAG comment: The EAG agrees with the quality assessment presented in the above Table. The overall quality issue is related to the robustness of the evidence coming from any single-arm, non-comparative study. The observed outcomes might not be a true reflection of the intervention.

3.2.5 Efficacy results of the KEYNOTE-087 trial and the SACT cohort

Efficacy results are presented in this Section for both cohort 2 of KEYNOTE-087 as well as the SACT cohort, when outcomes were presented. Eighty-one patients comprised cohort 2 of KEYNOTE-087. The first patient enrolled on 24 June 2015 and the last on 16 December 2016. The database cutoff for the CS was 15 March 2021 with a median duration of follow-up at 62.2 months (range: 2.1 months to 67.5 months). The disposition of patients at the time of the cutoff is presented in Table 3.9. The data for the SACT cohort were captured between 25 July 2018 and 30 September 2022. A total of 215 patients comprises the cohort.

Table 3.9: Disposition of patients forming cohort 2 of KEYNOTE-087

Status for study medication in trial segment treatment	Cohort 2 (N=81), n (%)
Started	81
Completed	13 (16.0)
Discontinued:	68 (84.0)
Adverse event	5 (6.2)
Bone marrow transplant	2 (2.5)
Clinical progression	1 (1.2)
Complete response	9 (11.1)
Lost to follow-up	2 (2.5)
Physicians decision	6 (7.4)
Pregnancy	1 (1.2)
Progressive disease	37 (45.7)
Withdrawal by subject	5 (6.2)
Based on Table 13 of the CS[4] CS = company submission	

3.2.5.1 Overall response rate

Overall response rate (ORR) was the primary outcome for KEYNOTE-087, in the ASaT population. It was based on classification of response for CR, PR, SD and PD, by blinded, independent central review (BICR), based on International Working Group (IWG) response criteria[34], with a secondary endpoint assessed following the Lugano classification system[35].

Fifty-two (64.2%) and 55 (67.9%) of 81 patients achieved either CR or PR as per IWG and Lugano criteria, respectively. A summary of results is presented in Table 3.10. Median time to response by IWG criteria for participants achieving CR or PR (n=52) was 2.8 months (2.2 to 11.0 months), while median duration of response (DOR) was 11.1 months (0.0+ to 59.0+ months) (Table 3.11). Response durations of ≥12 and ≥24 months are illustrated in Figure 3.1, were achieved by 15 (45.6% by KM estimation) and 10 (32.6% by KM estimation) participants, respectively.

Table 3.10: Summary of best overall response based on BICR per IWG and Lugano classifications for KEYNOTE-087, cohort 2

Level of response	Cohort 2 (N=81)			
	n (%)	95% CI[a]	n (%)	95% CI[a]
	IWG criteria		Lugano criteria	
ORR (CRR + PR)	52 (64.2)	52.8 to 74.6	55 (67.9)	56.6 to 77.8
CRR	21 (25.9)	16.8 to 36.9	23 (28.4)	18.9 to 39.5
PR	31 (38.3)	27.7 to 49.7	32 (39.5)	28.8 to 51.0
SD	8 (9.9)	4.4 to 18.5	6 (7.4)	2.8 to 15.4
PD	19 (23.5)	14.8 to 34.2	18 (22.2)	13.7 to 32.8
NA	2 (2.5)	0.3 to 8.5	2 (2.5)	0.3 to 8.5

Based on Table 14 of the CS[4]
[a] Based on binomial exact CI method
BICR = blinded = independent central review; CI = confidence interval; CR = complete remission rate; CS = company submission; IWG = International Working Group; NA = no assessment; OR = objective response rate; PD = progressive disease; PR = partial remission; SD = stable disease

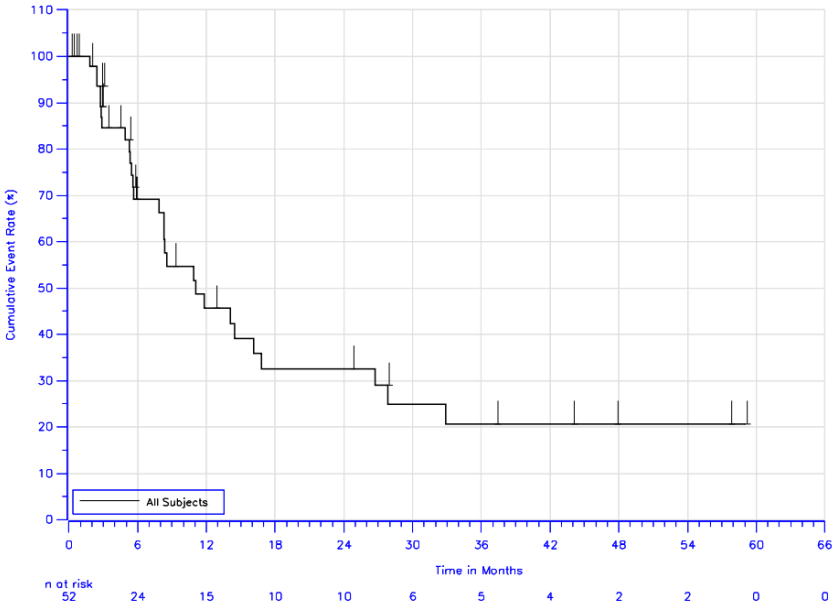
Table 3.11: Summary of time to response and response duration for those achieving CR or PR based on BICR and IWG criteria for KEYNOTE-087, cohort 2

Outcome	Cohort 2 (n=52)
Time to response (months)	
Mean (SD)	3.2 (1.4)
Median (range)	2.8 (2.2 to 11.0)
Response duration (months)	
Median (range)	11.1 (0.0+ to 59.0+)
95% CI	7.9 to 16.8
Number with response lasting ≥3 months (%)[a]	36 (84.5)
Number with response lasting ≥6 months (%)[a]	24 (69.1)
Number with response lasting ≥9 months (%)[a]	19 (54.7)
Number with response lasting ≥12 months (%)[a]	15 (45.6)
Number with response lasting ≥24 months (%)[a]	10 (32.6)
Number with response lasting ≥36 months (%)[a]	5 (20.7)

Outcome	Cohort 2 (n=52)
Number with response lasting ≥48 months (%)[a]	2 (20.7)

Based on Table 15 of the CS[4]
[a] From product-limit (Kaplan-Meier) method for censored data
BICR = blinded independent central review; CI = confidence interval; CR = complete remission; CS = company submission; IWG = International Working Group; PR = partial remission; SD = standard deviation.

Figure 3.1: Kaplan-Meier estimates of duration of objective response for cohort 2 based on BICR and IWG criteria



Based on Figure 3 of the CS[4]
BICR = blinded, independent central review; CS = company submission; IWG = International Working Group

3.2.5.2 Health-related quality of life

Health-related quality of life outcomes were presented for the entire KEYNOTE-087 trial and not separately for cohort 2. In the entire ASaT population (N=210), pembrolizumab was found to be associated with a clinically meaningful improvement in European quality of life-5 dimensions (EQ-5D) from baseline to Week 12, with a mean change in score of 8.4 points. An increase >8 points represents the threshold for a minimal clinically important difference. As shown in Table 3.12, the greatest improvements (change from baseline) in HRQoL were recorded for those patients classed as responders (achieving CR/PR).

Table 3.12: Summary of HRQoL endpoints in KEYNOTE-087

Population	Change from baseline at Week 12 EORTC QLQ-C30		Change from baseline at Week 12 EQ-5D-VAS	
	N[a]	Mean (SD)	N[a]	Mean (SD)
ASaT	184	8.6 (1.6)	191	8.4 (1.4)
CRR/PR	110	10.4 (2.1)	116	10.9 (1.8)
SD	48	7.3 (3.2)	49	5.4 (3.0)
PD	26	3.5 (3.6)	26	2.6 (2.7)

Based on Table 16 of the CS[4]

Population	Change from baseline at Week 12 EORTC QLQ-C30		Change from baseline at Week 12 EQ-5D-VAS	
	N[a]	Mean (SD)	N[a]	Mean (SD)
[a] Number of patients in the ASaT population with observations at baseline and Week 12. Data cutoff of 25 September 2016. ASaT = all subjects as treated; CR = complete remission rate; CS = company submission; EORTC QLQ-C30 = European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire-Core 30; EQ-5D VAS = European quality of life-5 dimensions visual analogue scale; HRQoL = health-related quality of life; PD = progressive disease; PR = partial remission; SD = stable disease; SD = standard deviation				

EAG comment: It is not clear why the Company chose to present the HRQoL outcomes for the entire trial and not for cohort 2, which is relevant to this CS. Any interpretation of these outcomes should be done with caution.

3.2.5.3 Stem cell therapy

Two outcomes related to SCT are presented in this Section. The proportion of patients receiving SCT as well as time to receiving SCT. Data were available for both KEYNOTE-087, cohort 2 and the SACT cohort.

Proportion of people receiving SCT

A 29.6% and a 30.2% of people received SCT in KEYNOTE-087, cohort 2 and in the SACT cohort, respectively. In cohort 2 of KEYNOTE-087, out of the 24 (29.6%) patients who received SCT, 14 (58.3%) and 9 (37.5%) underwent autoSCT and alloSCT, respectively. In the SACT cohort, out of the 65 (30.2%) patients who received SCT, 23 (35.4%) and 42 (64.6%) underwent autoSCT and alloSCT, respectively. Although the proportion of patients receiving SCT between the two cohorts is comparable, the majority of patients received autoSCT in KEYNOTE-087 (cohort 2), while the majority of patients in SACT received alloSCT (Table 3.12). Within the SACT cohort, 132 (61%) patients were identified in Blueteq as being suitable candidates for SCT, further details are presented in Table 3.13.

According to the Company, clinical expert opinion suggests that the ratio for alloSCT and autoSCT in the SACT is not representative to the English clinical practice, stating that lower alloSCT would be expected. The Company attributes this difference between autoSCT and alloSCT ratios to the fact that the patients in the SACT cohort are older and less fit, compared to the participants of cohort 2 from KEYNOTE-087, as well as being SCT-naïve.

Table 3.13: Proportion of people receiving SCT

Study/population	SCT	autoSCT	alloSCT
KEYNOTE-087, cohort 2	24	14	9
SACT	65	42	23
alloSCT = allogeneic stem cell therapy; autoSCT = autologous stem cell therapy; SACT = systemic anti-cancer therapy; SCT – stem cell therapy			

Table 3.14: Applications to CDF for pembrolizumab for treating cHL – SCT suitability in Blueteq and SCT procedures in HES

SCT suitability	Blueteq SCT suitability[a] (N)	HES alloSCT (N)	HES autoSCT (N)	HES SCT (N)	HES SCT (%)
Candidate for future SCT	132	42	23	65	49
Not a candidate for SCT	83	3	1	–	–

SCT suitability	Blueteq SCT suitability[a] (N)	HES alloSCT (N)	HES autoSCT (N)	HES SCT (N)	HES SCT (%)
Total	215	45	24	65[b]	–

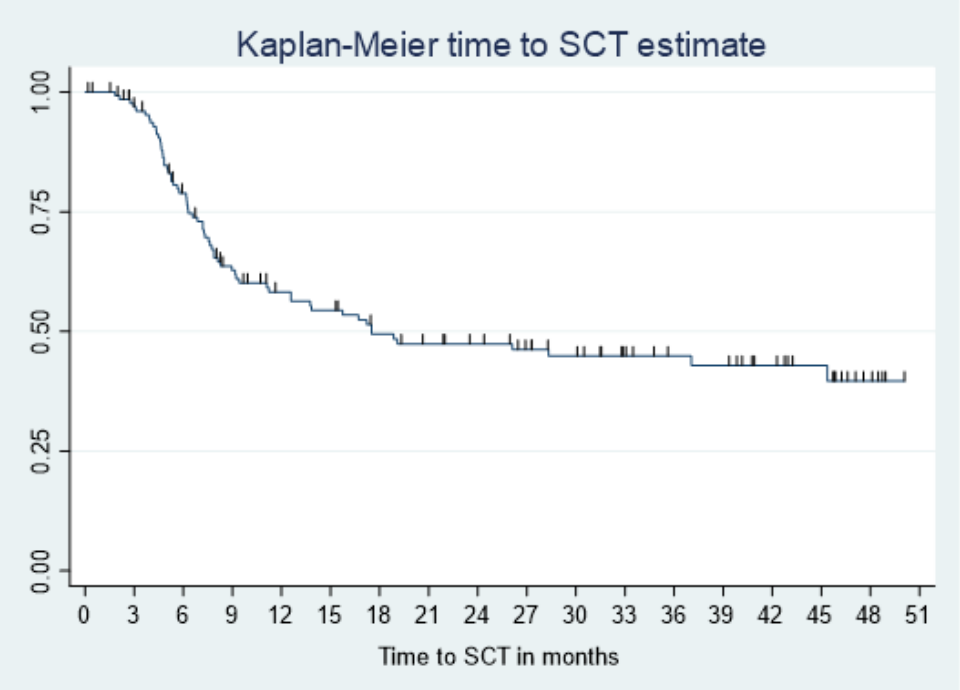
Based on Table 17 of the CS[4]
[a] Applications made between 25 July 2018 and 30 September 2022.
[b] Total number of SCTs including those who were not initially deemed a candidate for SCT seems to be 69 rather than 65. MSD have taken the 65/215 at face value and discuss an alternate assumption of 69/215 in the economic analysis.
alloSCT = allogeneic stem cell therapy; autoSCT = autologous stem cell therapy; CDF = Cancer Drugs Fund; cHL = classical Hodgkin lymphoma; CS = company submission; HES = Hospital Episodes Statistics; MSD = Merck Sharp and Dohme; SCT = stem cell therapy

Time to SCT

Time to alloSCT is one of the outcomes specified in the NICE final scope[2]. The CS only reports on time to SCT (autoSCT and alloSCT combined) stating that only such data were available for both KEYNOTE-087, cohort 2 and the SACT dataset. The estimated mean time to SCT in cohort 2 of KEYNOTE-087 was 30.3 months (131.4 weeks). In the SACT dataset, median time to SCT of patients eligible for SCT was 17.5 months (532 days) (Figure 3.2). Time to SCT limited to patients receiving an SCT was not reported separately.

According to the Company the notable difference between the two cohorts can be potentially attributed to variation in clinical practice. KEYNOTE-087, cohort 2, only included 10 patients from the UK, five of which went on to receive SCT. Information on whether these patients would be classified as potential future candidates for SCT is not available. The Company stated that they consider the SACT data to be more generalizable to clinical practice in England.[4]

Figure 3.2: Kaplan-Meier estimates for time to SCT in HES (N=132)[a]



Based on figure 4 of the CS[4]
[a] Patients identified in Blueteq as being suitable candidates for SCT
CS = company submission; HES = Hospital Episodes Statistics; SCT = stem cell transplant

EAG comment:

- There is a notable difference between the proportion of people receiving SCT between the two cohorts (cohort 2 of KEYNOTE-087 versus SACT dataset). According to the Company the clinical experts they consulted stated that the proportion of patients receiving alloSCT is higher than what is observed in UK clinical practice. Nevertheless, three additional notes were made by these clinical experts:
  - *"It was noted by an advisor that this rate is reflective of the last few years of clinical practice; however, there is also an increasing use of autoSCT across the general population*
  - *One advisor commented that they perform 5–10 times more autoSCTs than alloSCTs – alloSCTs are generally now only performed following relapse post autoSCT*
  - *It was highlighted that this variation may be due to the specific sub-cohort of patients who have failed on two lines of chemotherapy and BV"* (page 9[6])

  Therefore, the clinical experts highlighted that the rates are reflective of recent clinical practice and that this variation could be related to the specific sub-population addressed in this CS.
- There is a notable difference between time to SCT in KEYNOTE-087, cohort 2 and the SACT dataset: 30.3 months (mean) versus 17.5 months (median). The Company attributed this difference to clinical practice in the different countries the data came from. If this is the case, this also speaks to the generalisability of the rest of the outcomes coming from KEYNOTE-087 to the UK setting.
- The outcome time to SCT was presented for combined alloSCT and autoSCT. As already stated in Section 3.2.1.2, according to NICE final scope[2] as well as the recommendations on the FAD for TA540[1], time to alloSCT was of interest. The effect of including both types of SCTs in the outcome is not clear to the EAG, which leads to uncertainty of the presented evidence. As stated in Section 3.2.1.2 this is a key issue.

3.2.5.4 Progression-free survival

Progression-free survival data were available only for KEYNOTE-087, cohort 2, and was defined as the time from first dose to the first documented disease progression (by BICR) and by site assessment according to IWG criteria or death due to any cause, whichever occurred first. Fifty-seven patients (70.4%) experienced an event and median PFS in the ASaT population was 11.1 months (95% confidence interval (CI): 7.5 to 13.7) (BICR and IWG criteria). A summary of PFS outcomes is presented in Table 3.15. Kaplan-Meier estimates of PFS are illustrated in Figure 3.3.

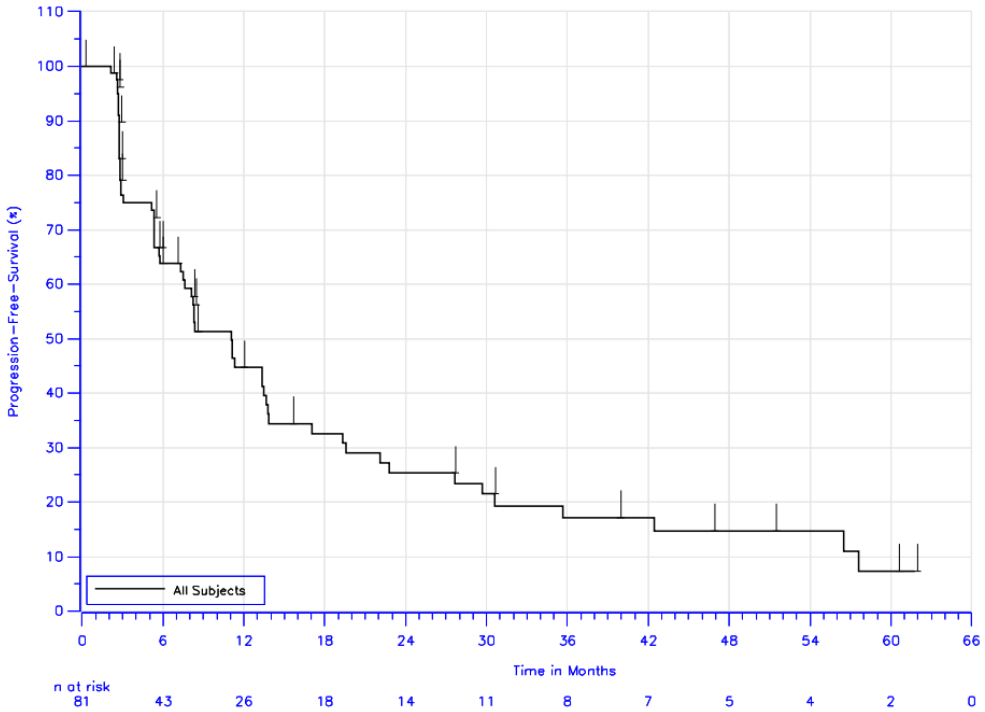
Table 3.15: Summary of PFS for cohort 2 based on BICR as per IWG criteria

Outcome	Cohort 2 (N=81)
Number of events, n (%)	57 (70.4)
Person-months	1080
Event rate/100 person-months (%)	5.3
Median PFS (95% CI), months	11.1 (7.5 to 13.7)
PFS rate at various time points	
6 months (%)[a]	63.8
12 months (%)[a]	44.7
24 months (%)[a]	25.4
36 months (%)[a]	17.2

Outcome	Cohort 2 (N=81)
48 months (%)[a]	14.7
60 months (%)[a]	7.4

Based on Table 18 of the CS[4]
[a] From product-limit (Kaplan-Meier) method for censored data.
BICR = blinded independent central review; CI = confidence interval; CS = company submission; IWG = International Working Group; PFS = progression-free survival; SD = standard deviation

Figure 3.3: Kaplan-Meier estimates of PFS for KEYNOTE-087, cohort 2 based on BICR and IWG criteria



Based on Figure 5 of the CS[4]
BICR = blinded independent central review; CS = company submission; IWG = International Working Group; PFS = progression-free survival

3.2.5.5 OS

Mean OS for KEYNOTE-087, cohort 2 was 53.7 months (standard error (SE) 1.8 months), while median OS was not reached. At a median follow-up of 62.2 months, 24 out of the 81 patients (29.6%) had died (Table 3.16). In the SACT dataset, at a median follow-up of 19.2 months, 73 out of the 215 patients (34%) had died. Median OS was also not reached, while mean OS was not available. A summary of OS data is presented in Table 3.16 for both cohorts. Kaplan–Meier estimates of OS for KEYNOTE-087 and the SACT dataset are illustrated in Figure 3.4 and Figure 3.5 respectively.

There is a notable difference between the proportion of patients that were still alive at the set time points presented in Table 3.16. The Company attributes this to the difference in patients' characteristics. In addition, the Company considers that the prognosis of patients in the SACT cohort is better than those not receiving pembrolizumab in this setting.

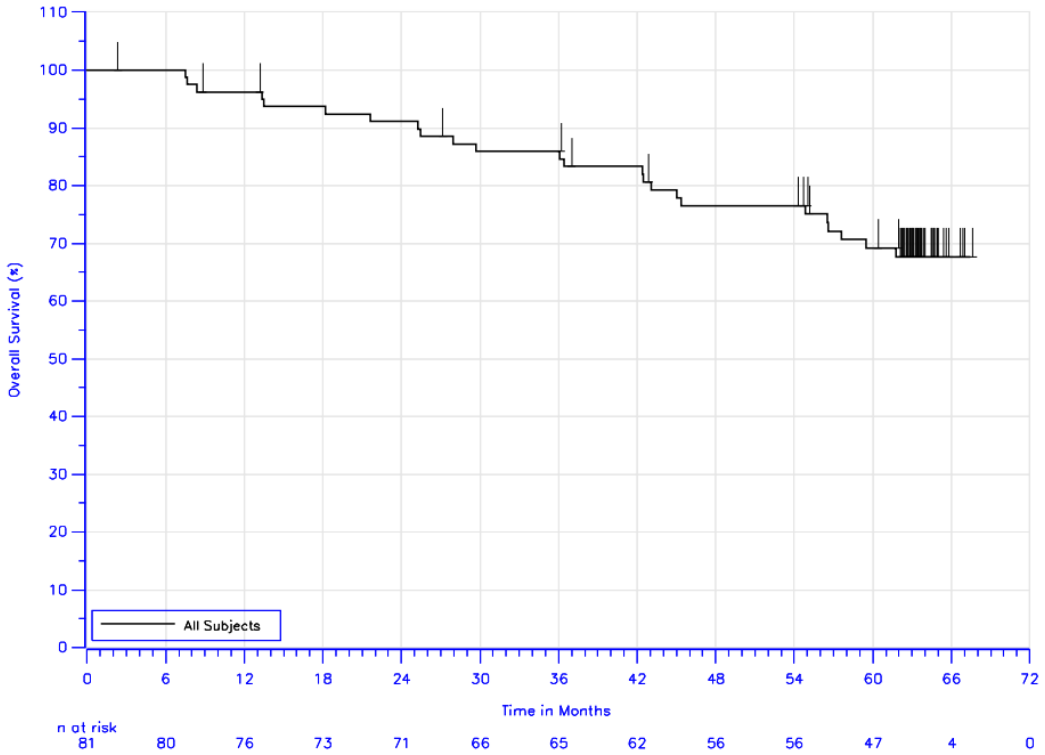
Table 3.16: Summary of OS for KEYNOTE-087, cohort 2 and for the SACT dataset

Outcome	Cohort 2 (N=81)	SACT dataset (N=215)
Number of events, n (%)	24 (29.6)	73 (34.0%)

Outcome	Cohort 2 (N=81)	SACT dataset (N=215)
Median OS (95% CI), months	Not reached	Not reached
OS rate at various time points		
6 months (%)	100.0[a]	88% (95% CI: 83% to 92%)
12 months (%)	96.3[a]	82% (95% CI: 76% to 87%)
18 months (%)	93.7[a]	75% (95% CI: 68% to 80%)
24 months (%)	91.1[a]	68% (95% CI: 61% to 75%)
36 months (%)	85.9[a]	56% (95% CI: 47% to 64%)
48 months (%)	76.5[a]	55% (95% CI: 46% to 63%)
60 months (%)	69.2[a]	N/A

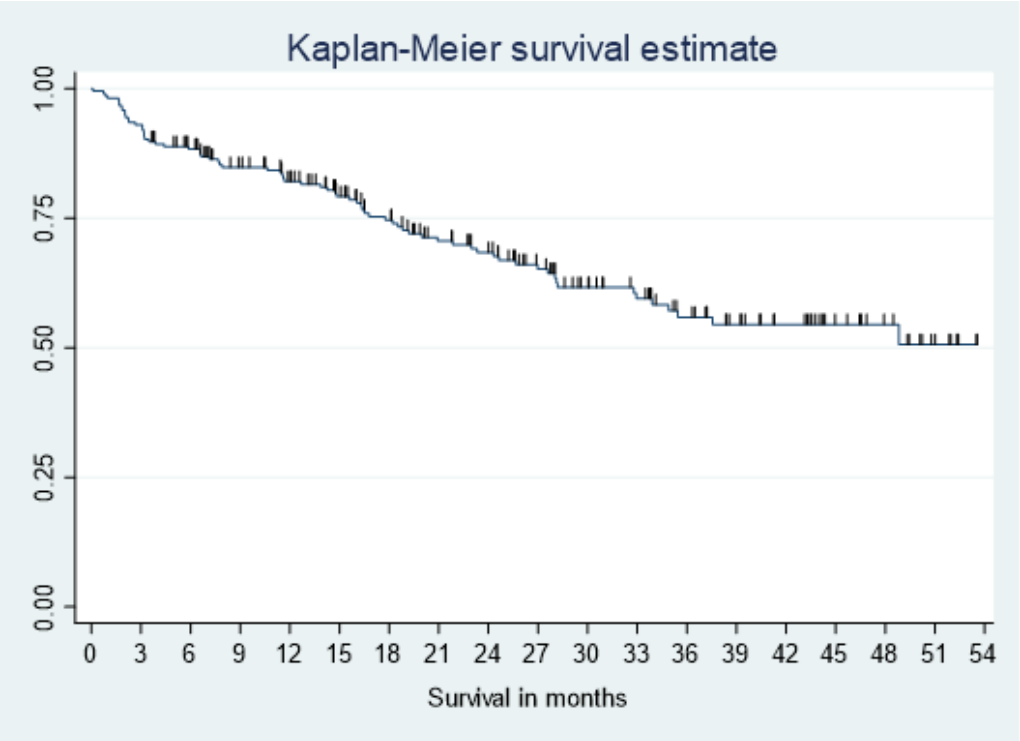
Based on Table 19 of the CS[4]
[a] From product-limit (Kaplan-Meier) method for censored data.
CI = confidence interval; CS = company submission; N/A = not available; OS = overall survival; SACT = systemic anti-cancer therapy

Figure 3.4: Kaplan-Meier estimates of OS for KEYNOTE-087, cohort 2



Based on Figure 6 of the CS[4]
CS = company submission; OS = overall survival

Figure 3.5: Kaplan-Meier estimates of OS for the SACT dataset



Based on Figure 7 of the CS[4]
CS = company submission; OS = overall survival; SACT = systemic anti-cancer therapy

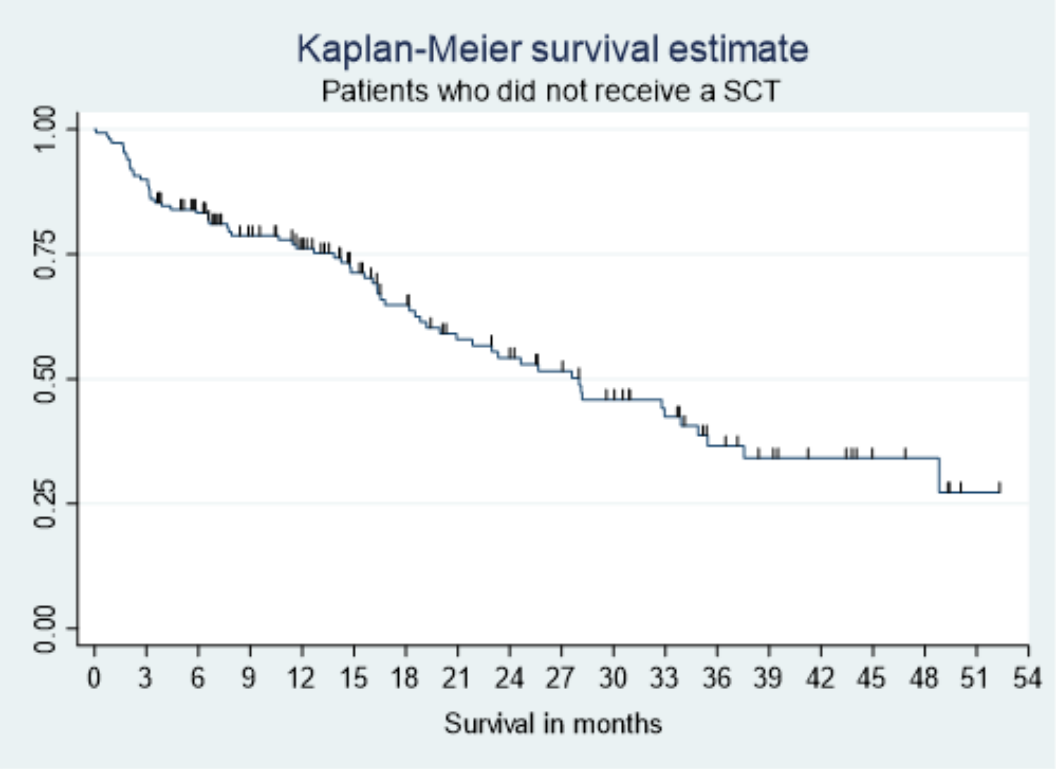
OS for patients who did and did not receive a SCT

In patients with R/R cHL SCT can be curative. In KEYNOTE-087 the impact of SCT on OS was examined in a *post hoc* subgroup analysis separately for patients who received and did not receive SCT. Such subgroup data was also available for the SACT dataset. The Company noted the limitation of *post hoc* analysis, which warrantee caution in the interpretation of the outcomes.

In KEYNOTE-087, cohort 2, out of the 24 patients who received SCT, 19 (79.2%) remained alive at the last follow-up. Mean OS was 53.7 months (SE 2.7 months), and mean OS after SCT was 42.5 months (SE 2.5 months). In the SACT cohort, of the 65 patients who received SCT, 59 (91%) remained alive at the time of data cutoff. Median and mean OS were not available. For those not receiving SCT, 38 (66.7%) from KEYNOTE-087, cohort 2 remained alive, and mean OS was 52.1 months (SE 2.3 months), while median OS was not reached. In the SACT dataset among the 150 (70%) patients who did not receive an SCT, median OS was 28 months (95% CI: 20.0 to 34.9 months) (Figure 3.6).

In KEYNOTE-087, cohort 2, the mean OS is similar for those receiving and not receiving SCT. Nevertheless, despite a median follow-up of 62.2 months, median OS has not been reached in either subgroup and the Company noted that data may be too immature to distinguish the impact of SCT on OS.

Figure 3.6: Kaplan-Meier survival plot for patients from the SACT dataset who did not receive SCT



Based on Figure 8 of the CS[4]
CS = company submission; SACT = systemic anti-cancer therapy; SCT = stem cell transplant

EAG comment: A comparison between the median OS of patients receiving SCT in cohort 2 of KEYNOTE-087 and the SACT dataset was not possible since this outcome was not presented for the latter. When comparing OS for patients who did not receive SCT there is a notable difference where in cohort 2 of KEYNOTE-087 the mean OS was 52.1 months (median not reached) while the median OS was 28 months in the SACT dataset.

3.2.5.6 Subgroup analysis

The prespecified subgroup analysis in KEYNOTE-087 focused on ORR was consistent in the following groups:

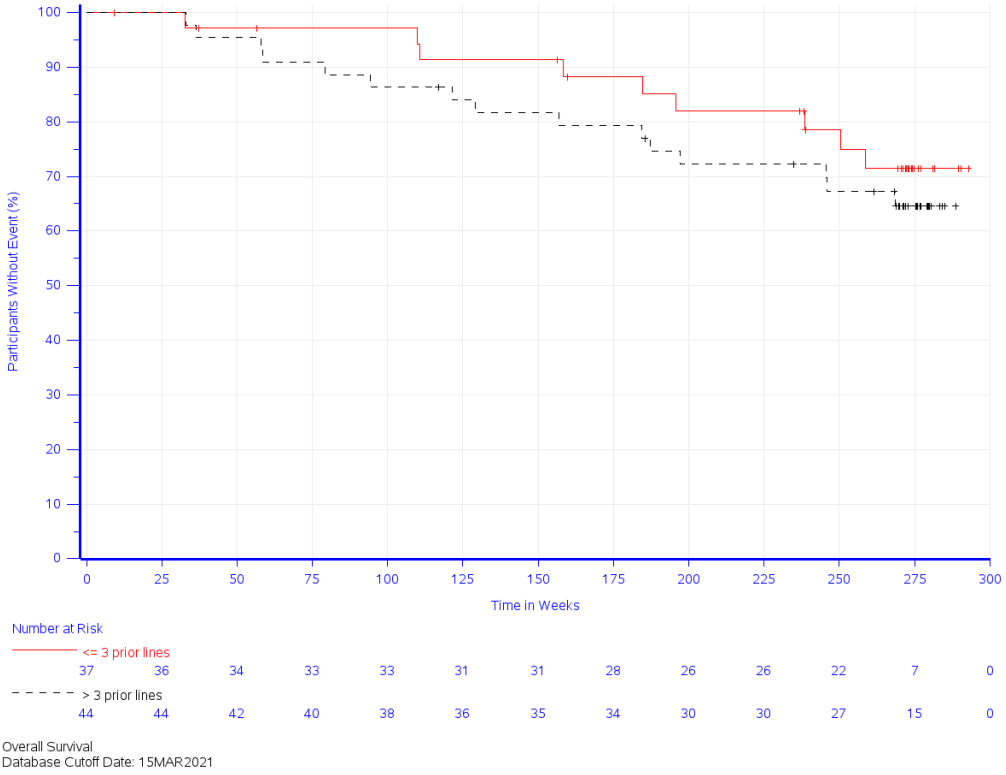
- Age category (≤65 versus >65 years)
- Sex (female versus male)
- Race (white versus non-white)
- Region (US versus ex-US)
- Number of prior therapies (<4 versus ≥4)

According to the CS[4] "*If the observed numbers for a particular subgroup were too small to make a clinically meaningful interpretation, the subgroup analysis was not conducted. Data on ORR by subgroup are not available in the CSR.*". In addition, since subgroup analyses for the outcomes of PFS and OS were not prespecified, these were also not available.

EAG comment:

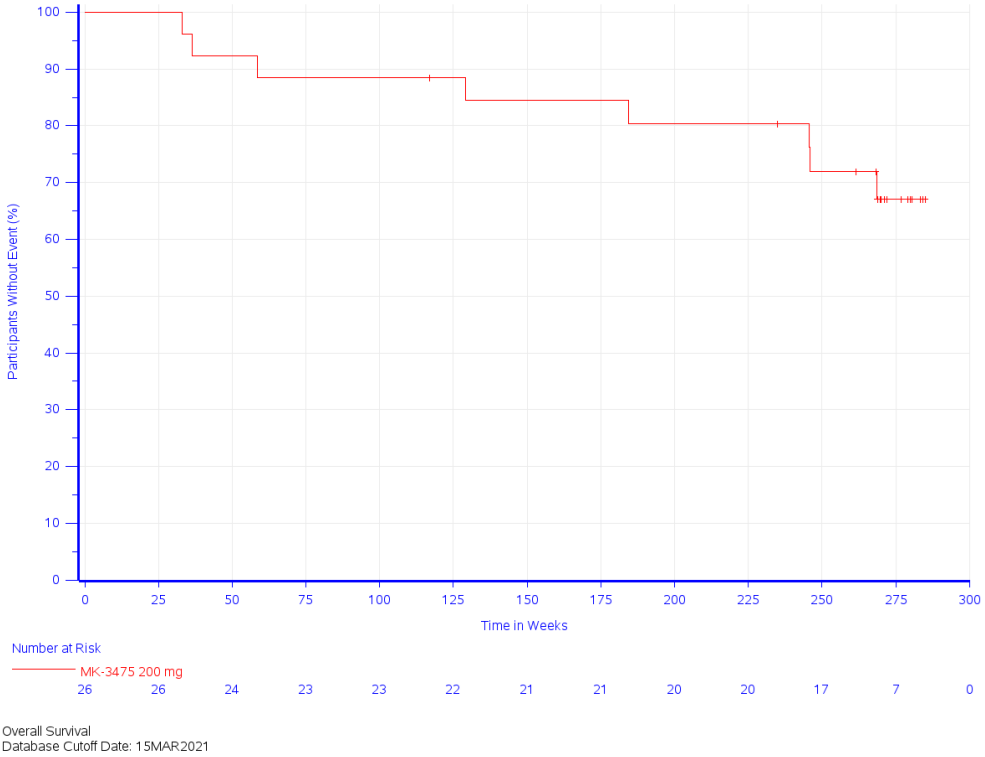
- The EAG requested[8] that the Company would provide the results of the subgroup analysis for age category (≤65 versus >65 years), sex (female versus male), race (white versus non-white), region (US versus ex-US) and number of prior therapies (<4 versus ≥4). The Company reiterated that such analyses was not available. "*The CSR stipulates that if the observed numbers for a particular subgroup are too small to make a meaningful clinical interpretation, then that subgroup analysis would not be conducted.*" (pages 18-19[9]). The Company did not provide any further information on the number of patients in the subgroups so that the EAG could make an informed critique, nor define what is the number of patients that was considered 'too small' to conduct a subgroup analysis. Since all the prespecified subgroup analysis for ORR were expressed as two categorical variables: age of ≤65 versus >65 years; female versus male; white race versus non-white race; US region versus ex-US region; number of prior therapies <4 versus ≥4 and the number of participants in cohort 2 of KEYNOTE-087 was 81, it is unlikely that the number of patients was too small to conduct a subgroup analysis.
- In the Company's response to a different clarification question, the Company provided KM curves for OS for Cohort 2 from KEYNOTE-087 based on number of lines of prior therapy for ≤3 versus >3 prior lines of therapy (Figure 3.7), 4 prior lines of therapy (Figure 3.8) and ≥5 prior lines of therapy (Figure 3.9). As expected, further line of therapy is associated with lower OS outcomes.

Figure 3.7: Kaplan-Meier for OS for cohort 2 of KEYNOTE-087 by ≤3 versus >3 prior lines of therapy



Based on Figure 2 of the response to the request for clarification[9]
OS = overall survival

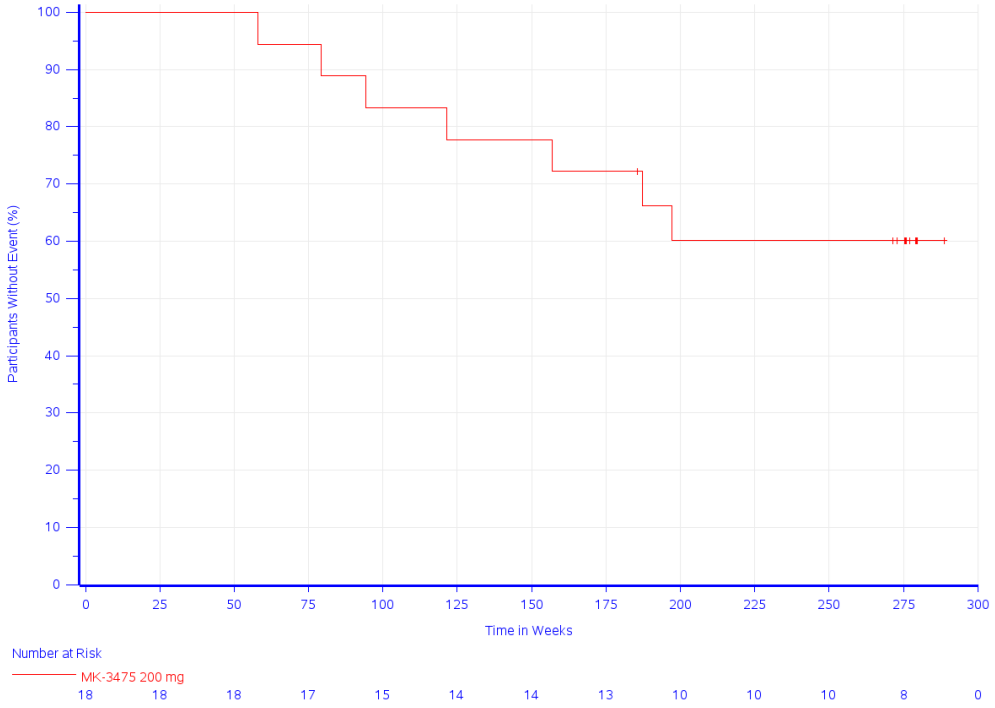
Figure 3.8: Kaplan-Meier for OS for cohort 2 of KEYNOTE-087 with 4 prior lines of therapy



Overall Survival
Database Cutoff Date: 15MAR2021

Based on Figure 3 of the response to the request for clarification[9]
OS = overall survival

Figure 3.9: Kaplan-Meier for OS for cohort 2 of KEYNOTE-087 with ≥5 prior lines of therapy



Overall Survival
Database Cutoff Date: 15MAR2021

Based on Figure 4 of the response to the request for clarification[9]
OS = overall survival

3.2.6 Adverse effects

Safety analyses were based on the ASaT population up to the data cutoff of 15 March 2021, which corresponds to approximately 5 years after the last patient-initiated study treatment. KEYNOTE-087 comprised three cohorts of patients, of which cohort 2 (N=81) is the population relevant to the STA, that is, those who are SCT-naïve and have relapsed after treatment with, or failed to respond to, BV.

The median duration of exposure in cohort 2 was 254.0 days, with a median of 13.0 administrations and 63.0% of patients remaining on treatment for ≥6 months (Table 3.17).

Table 3.17: Summary of drug and clinical trial exposure for cohort 2 and for the full trial population of KEYNOTE-087

Characteristic	Cohort 2 (N=81)		KEYNOTE-087 (N=210)	
Number of days on therapy				
Mean (SD)	393.2 (342.4)		461.4 (355.6)	
Median (range)	254.0 (1 to 1696)		387.0 (1 to 1880)	
Number of administrations				
Mean (SD)	17.5 (12.8)		20.3 (12.8)	
Median (range)	13.0 (1 to 52)		18.5 (1 to 52)	
Duration of exposure				
	n (%)	Person-years	n (%)	Person-years
>0 months	81 (100)	87.2	210 (100)	265.3
≥1 months	80 (98.8)	87.2	206 (98.1)	265.1
≥3 months	73 (90.1)	85.9	194 (92.4)	262.8
≥6 months	51 (63.0)	77.3	155 (73.8)	248.2
≥12 months	35 (43.2)	66.6	108 (51.4)	214.9

Based on Table 21 of the CS[4]
CS = company submission; SD = standard deviation

3.2.6.1 Overall adverse events

In cohort 2, 80 of 81 (98.8%) patients experienced at least one adverse event (AE); 26 of 81 (32.1%) patients experienced Grade 3-5 AE (Table 3.18). Adverse events resulting in death occurred in three, two in cohort 2. The Company stated that the death was attributed to one each of acute Graft Versus Host Disease (GVHD), post-procedural infection, and septic shock, and no death was deemed to be related to a drug-related AE.

Table 3.18: Summary of AEs for KEYNOTE-087 by cohort for the ASaT population

Characteristic	Cohort 1 (N=69) n (%)	Cohort 2 (N=81) n (%)	Cohort 3 (N=60) n (%)	Total (N=210) n (%)
One or more AE, n (%)	68 (98.6)	80 (98.8)	57 (95.0)	205 (97.6)
Drug-related[a] AE, n (%)	54 (78.3)	52 (64.2)	47 (78.3)	153 (72.9)
Toxicity grade 3–5 AEs, n (%)	22 (31.9)	26 (32.1)	21 (35.0)	69 (32.9)
Toxicity grade 3–5 drug-related AEs, n (%)	12 (17.4)	10 (12.3)	5 (8.3)	27 (12.9)
Non-serious AEs, n (%)	68 (98.6)	79 (97.5)	57 (95.0)	204 (97.1)

Characteristic	Cohort 1 (N=69) n (%)	Cohort 2 (N=81) n (%)	Cohort 3 (N=60) n (%)	Total (N=210) n (%)
Serious AEs, n (%)	15 (21.7)	18 (22.2)	15 (25.0)	48 (22.9)
Serious drug-related AEs, n (%)	8 (11.6)	4 (4.9)	5 (8.3)	17 (8.1)
Died, n (%)	0 (0.0)	2 (2.5)	1 (1.7)	3 (1.4)
Died due to a drug-related AE, n (%)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Discontinued[b] due to an AE, n (%)	8 (11.6)	5 (6.2)	5 (8.3)	18 (8.6)
Discontinued due to a drug-related AE, n (%)	6 (8.7)	4 (4.9)	4 (6.7)	14 (6.7)
Discontinued due to a serious AE, n (%)	5 (7.2)	3 (3.7)	2 (3.3)	10 (4.8)
Discontinued due to a serious drug-related AE, n (%)	3 (4.3)	2 (2.5)	2 (3.3)	7 (3.3)

Based on Table 22 of the CS[4]
The following footnotes were included in the CS[4]
[a] Determined by the investigator to be related to the drug.
[b] Study medication withdrawn.
Grades are based on NCI CTCAE version 4.0. Non-serious AEs up to 30 days of last dose and serious AEs 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.
ASaT = all subjects as treated; AE = adverse event; MedDRA = Medical Dictionary for Regulatory Activities; NCI CTCAE = National Cancer Institute Common Terminology Criteria for Adverse Events

3.2.6.2 Most frequently reported AEs

The most frequently reported AEs (>15%) in cohort 2 included cough (27.2%), pyrexia (23.5%), fatigue (21.0%), nasopharyngitis (19.8%), and hypothyroidism (16.0%).

Of the 81 patients treated in KEYNOTE-087 cohort 2, 52 (64.2%) experienced ≥1 treatment-related adverse events (TRAEs). The most frequently reported drug-related AEs in cohort 2 included hypothyroidism (14.8%), pyrexia (8.6%), fatigue (7.4%), and rash (6.2%).

An overview of the 10 most commonly experienced AEs and drug-related AEs is available in Table 3.19.

Table 3.19: Overview of the most frequently reported AEs and drug-related AEs across cohorts from KEYNOTE-087

Adverse effect	Cohort 1 (N=69) n (%)	Cohort 2 (N=81) n (%)	Cohort 3 (N=60) n (%)	Total (N=210) n (%)
Any cause (>15%)				
Pyrexia	27 (39.1)	19 (23.5)	17 (28.3)	63 (30.0)
Cough	19 (27.5)	22 (27.2)	14 (23.3)	55 (26.2)
Fatigue	15 (21.7)	17 (21.0)	16 (26.7)	48 (22.9)
Diarrhoea	20 (29.0)	12 (14.8)	11 (18.3)	43 (20.5)
Upper respiratory tract infection	23 (33.3)	7 (8.6)	13 (21.7)	43 (20.5)
Nausea	16 (23.2)	11 (13.6)	11 (18.3)	38 (18.1)

Adverse effect	Cohort 1 (N=69) n (%)	Cohort 2 (N=81) n (%)	Cohort 3 (N=60) n (%)	Total (N=210) n (%)
Vomiting	16 (23.2)	9 (11.1)	13 (21.7)	38 (18.1)
Nasopharyngitis	12 (17.4)	16 (19.8)	7 (11.7)	35 (16.7)
Arthralgia	14 (20.3)	12 (14.8)	7 (11.7)	33 (15.7)
Hypothyroidism	8 (11.6)	13 (16.0)	12 (20.0)	33 (15.7)
Drug-related (incidence ≥5% in one or more treatment groups)				
Experienced ≥1 AE	54 (78.3)	52 (64.2)	47 (78.3)	153 (72.9)
Hypothyroidism	6 (8.7)	12 (14.8)	12 (20.0)	30 (14.3)
Pyrexia	12 (17.4)	7 (8.6)	5 (8.3)	24 (11.4)
Fatigue	10 (14.5)	6 (7.4)	7 (11.7)	23 (11.0)
Rash	10 (14.5)	5 (6.2)	8 (13.3)	23 (11.0)
Diarrhoea	9 (13.0)	4 (4.9)	4 (6.7)	17 (8.1)
Headache	10 (14.5)	3 (3.7)	3 (5.0)	16 (7.6)
Nausea	7 (10.1)	2 (2.5)	6 (10.0)	15 (7.1)
Arthralgia	4 (5.8)	5 (6.2)	4 (6.7)	13 (6.2)
Cough	3 (4.3)	4 (4.9)	6 (10.0)	13 (6.2)
Pruritus	4 (5.8)	5 (6.2)	4 (6.7)	13 (6.2)

Based on Table 23 of the CS[4]
The following footnotes were included in the CS[4]
Non-serious AEs up to 30 days of last dose and serious AEs 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.
AE = adverse event; CS = company submission; MedDRA = Medical Dictionary for Regulatory Activities

3.2.6.3 Grade 3 to 5 AEs

In cohort 2, AEs of Grade 3-5 were reported 32.1% of participants. Ten patients (12.3%) experienced ≥1 Grade 3 or 4 AE that was considered related to pembrolizumab, but no drug-related Grade 5 AE was reported. The most commonly reported Grade 3–5 AE was anaemia (three patients, Table 3.20), whereas the most frequent drug-related Grade 3 or 4 AE was neutropenia (two patients, Appendix F[17]), which of blood and lymphatic system disorders. An overview of the 10 most commonly experienced Grade 3–5 AEs (any cause and drug-related by organ class) is available in Table 3.20.

Table 3.20: Overview of Grade 3-5 AEs (incidence >0% in one or more treatment groups) experienced across cohorts in KEYNOTE-087

Organ class	Cohort 1 (N=69) n (%)	Cohort 2 (N=81) n (%)	Cohort 3 (N=60) n (%)	Total (N=210) n (%)
Any cause				
Experienced ≥1 AE Grade 3-5	22 (31.9)	26 (32.1)	21 (35.0)	69 (32.9)
Anaemia	4 (5.8)	3 (3.7)	0 (0.0)	7 (3.3)
Neutropenia	3 (4.3)	3 (3.7)	0 (0.0)	6 (2.9)
Pneumonia	4 (5.8)	1 (1.2)	0 (0.0)	5 (2.4)
Diarrhoea	3 (4.3)	1 (1.2)	0 (0.0)	4 (1.9)

Organ class	Cohort 1 (N=69) n (%)	Cohort 2 (N=81) n (%)	Cohort 3 (N=60) n (%)	Total (N=210) n (%)
Herpes zoster	0 (0.0)	2 (2.5)	1 (1.7)	3 (1.4)
Leukopenia	1 (1.4)	2 (2.5)	0 (0.0)	3 (1.4)
Pyrexia	2 (2.9)	0 (0.0)	1 (1.7)	3 (1.4)
Thrombocytopenia	0 (0.0)	2 (2.5)	1 (1.7)	3 (1.4)
Acute graft versus host disease	0 (0.0)	1 (1.2)	1 (1.7)	2 (1.0)
Alanine aminotransferase increased	1 (1.4)	1 (1.2)	0 (0.0)	2 (1.0)
Drug-related				
Experienced ≥1 AE Grade 3-5	12 (17.4)	10 (12.3)	5 (8.3)	27 (12.9)
Blood and lymphatic system disorders	3 (4.3)	2 (2.5)	1 (1.7)	6 (2.9)
Cardiac disorders	3 (4.3)	0 (0.0)	0 (0.0)	3 (1.4)
Gastrointestinal disorders	1 (1.4)	2 (2.5)	0 (0.0)	3 (1.4)
General disorders and administration site conditions	1 (1.4)	1 (1.2)	1 (1.7)	3 (1.4)
Hepatobiliary disorders	1 (1.4)	0 (0.0)	0 (0.0)	1 (0.5)
Immune system disorders	0 (0.0)	2 (2.5)	0 (0.0)	2 (1.0)
Infections and infestations	2 (2.9)	1 (1.2)	2 (3.3)	5 (2.4)
Investigations	2 (2.9)	1 (1.2)	0 (0.0)	3 (1.4)
Metabolism and nutrition disorders	1 (1.4)	0 (0.0)	0 (0.0)	1 (0.5)
Musculoskeletal and connective tissue disorders	2 (2.9)	1 (1.2)	0 (0.0)	3 (1.4)

Based on Table 25 of the CS[4]
The following footnotes were included in the CS[4]
Non-serious AEs up to 30 days of last dose and serious AEs 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. Grades are based on NCI CTCAE version 4.0.
AE = adverse event; CS = company submission; MedDRA = Medical Dictionary for Regulatory Activities; NCI CTCAE = National Cancer Institute Common Terminology Criteria for Adverse Events

3.2.6.4 Serious adverse events

In cohort 2, 18 patients (22.2%; Table 3.21) experienced a serious adverse events (SAE) during study treatment and through 90 days after the last dose of pembrolizumab. The most commonly reported SAE in cohort 2 was herpes zoster (2.5%). Across all cohorts of KEYNOTE-087, three cases of acute GVHD were recorded, one of which was fatal. An overview of the 10 most common SAEs is available in Table 3.21.

Table 3.21: Overview of SAEs incidence ≥1% in one or more treatment groups) experienced across cohorts in KEYNOTE-087

SAE	Cohort 1 (N=69) n (%)	Cohort 2 (N=81) n (%)	Cohort 3 (N=60) n (%)	Total (N=210) n (%)
Experienced ≥1 SAE	15 (21.7)	18 (22.2)	15 (25.0)	48 (22.9)
Pneumonia	4 (5.8)	1 (1.2)	1 (1.7)	6 (2.9)
Pneumonitis	1 (1.4)	1 (1.2)	2 (3.3)	4 (1.9)

SAE	Cohort 1 (N=69) n (%)	Cohort 2 (N=81) n (%)	Cohort 3 (N=60) n (%)	Total (N=210) n (%)
Pyrexia	0 (0.0)	1 (1.2)	3 (5.0)	4 (1.9)
Acute graft versus host disease	1 (1.4)	1 (1.2)	1 (1.7)	3 (1.4)
Bronchitis	0 (0.0)	1 (1.2)	1 (1.7)	2 (1.0)
Herpes zoster	0 (0.0)	2 (2.5)	0 (0.0)	2 (1.0)
Pericarditis	2 (2.9)	0 (0.0)	0 (0.0)	2 (1.0)
Acute kidney injury	1 (1.4)	0 (0.0)	0 (0.0)	1 (0.5)
Acute sinusitis	0 (0.0)	1 (1.2)	0 (0.0)	1 (0.5)
Anaemia	1 (1.4)	0 (0.0)	0 (0.0)	1 (0.5)

Based on Table 26 of the CS[4]
The following footnotes were included in the CS[4]
Non-serious AEs up to 30 days of last dose and serious AEs 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.
CS = company submission; MedDRA = Medical Dictionary for Regulatory Activities; SAE = serious adverse event

3.2.6.5 Discontinuation of study treatment due to AEs

Five (6.5%; Table 3.22) patients enrolled in KEYNOTE-087 cohort 2 discontinued pembrolizumab due to one or more AEs. The most commonly reported AEs that resulted in stopping pembrolizumab were pneumonitis (n=2; 2.5%) which of respiratory, thoracic and mediastinal disorders.

Four patients (4.9%; Table 3.22) discontinued pembrolizumab due to a drug-related AE in cohort 2. An overview of AEs leading to discontinuation by organ class is available in Table 3.22.

Table 3.22: Overview of AEs (incidence >0% in one or more treatment groups) leading to discontinuation of pembrolizumab across cohorts of KEYNOTE-087

Organ class	Cohort 1 (N=69) n (%)	Cohort 2 (N=81) n (%)	Cohort 3 (N=60) n (%)	Total (N=210) n (%)
AEs resulting in discontinuation				
Experienced ≥1 AE	8 (11.6)	5 (6.2)	5 (8.3)	18 (8.6)
Cardiac disorders	1 (1.4)	0 (0.0)	0 (0.0)	1 (0.5)
Immune system disorders	0 (0.0)	1 (1.2)	0 (0.0)	1 (0.5)
Infections and infestations	1 (1.4)	0 (0.0)	1 (1.7)	2 (1.0)
Injury, poisoning and procedural complications	1 (1.4)	1 (1.2)	0 (0.0)	2 (1.0)
Musculoskeletal and connective tissue disorders	2 (2.9)	0 (0.0)	0 (0.0)	2 (1.0)
Neoplasms benign, malignant and unspecified (including cysts and polyps)	2 (2.9)	1 (1.2)	0 (0.0)	3 (1.4)
Nervous system disorders	0 (0.0)	0 (0.0)	1 (1.7)	1 (0.5)
Respiratory, thoracic and mediastinal disorders	2 (2.9)	3 (3.7)	3 (5.0)	8 (3.8)

Drug-related AEs resulting in discontinuation				
Experienced ≥1 AE	6 (8.7)	4 (4.9)	4 (6.7)	14 (6.7)
Cardiac disorders	1 (1.4)	0 (0.0)	0 (0.0)	1 (0.5)
Immune system disorders	0 (0.0)	1 (1.2)	0 (0.0)	1 (0.5)
Infections and infestations	1 (1.4)	0 (0.0)	0 (0.0)	1 (0.5)
Injury, poisoning and procedural complications	1 (1.4)	1 (1.2)	0 (0.0)	2 (1.0)
Musculoskeletal and connective tissue disorders	2 (2.9)	0 (0.0)	0 (0.0)	2 (1.0)

Based on Table 27 of the CS[4] and Table 32 of Appendix F[17]
The following footnotes were included in the CS[4] and Appendix F[17]
Every subject is counted a single time for each applicable row and column.
Non-serious AEs up to 30 days of last dose and serious AEs 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.
AE = adverse event; CS = company submission; MedDRA = Medical Dictionary for Regulatory Activities

3.2.6.6 Adverse events of special interest of KEYNOTE-087

Of the 81 patients treated in KEYNOTE-087 cohort 2, 26 (32.1%; Table 3.23) experienced ≥1 adverse events of special interest (AEOSIs). Four patients (4.9%) experienced AEOSIs Grade 3-5. Twenty-five (30.9%) patients reported an AEOSI that was considered related to pembrolizumab. Four patients (4.9%) patients discontinued pembrolizumab due to a drug related AEOSI. No death was attributed to an AEOSI. The most common AEOSIs reported during the study were hypothyroidism (n=13; 16.0%), infusion-related reaction (n=7; 8.6%). An overview of the most commonly reported AEOSIs is available in Table 3.23.

Table 3.23: Overview of AEOSI (incidence >0% in one or more treatment groups) across cohorts of KEYNOTE-087

AEOSI by category	Cohort 1 (N=69) n (%)	Cohort 2 (N=81) n (%)	Cohort 3 (N=60) n (%)	Total (N=210) n (%)
Any cause				
Experienced ≥1 AEOSI	21 (30.4)	26 (32.1)	23 (38.3)	70 (33.3)
Experienced ≥1 AE Grade 3–5	4 (5.8)	4 (4.9)	0 (0.0)	8 (3.8)
Discontinued due to AEs	6 (8.7)	4 (4.9)	4 (6.7)	14 (6.7)
Hypothyroidism	8 (11.6)	13 (16.0)	12 (20.0)	33 (15.7)
Infusion reactions	6 (8.7)	7 (8.6)	6 (10.0)	19 (9.0)
Pneumonitis	3 (4.3)	4 (4.9)	4 (6.7)	11 (5.2)
Hyperthyroidism	1 (1.4)	4 (4.9)	3 (5.0)	8 (3.8)
Colitis	2 (2.9)	1 (1.2)	0 (0.0)	3 (1.4)
Uveitis	3 (4.3)	0 (0.0)	0 (0.0)	3 (1.4)
Myositis	2 (2.9)	0 (0.0)	0 (0.0)	2 (1.0)
Skin	1 (1.4)	1 (1.2)	0 (0.0)	2 (1.0)
Drug-related				
Experienced ≥1 AEOSI	20 (29.0)	25 (30.9)	22 (36.7)	67 (31.9)
Experienced ≥1 AE Grade 3–5	3 (4.3)	4 (4.9)	0 (0.0)	7 (3.3)

AEOSI by category	Cohort 1 (N=69) n (%)	Cohort 2 (N=81) n (%)	Cohort 3 (N=60) n (%)	Total (N=210) n (%)
Discontinued due to AEs	6 (8.7)	4 (4.9)	3 (5.0)	13 (6.2)

Based on Table 28 of the CS[4] and Table 12-10 of CSR[25]
The following footnotes were included in the CS[4]
Non-serious AEs up to 30 days of last dose and serious AEs 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.
AEs = adverse events; AEOSI = adverse event of special interest; CS = company submission; CSR = Clinical Study Report; MedDRA = Medical Dictionary for Regulatory Activities

3.2.6.7 Adverse events in the SACT cohort data set

There is no related AE data of SACT cohort data set reported.

EAG comment:

- The CS reports very high rates of AEs and drug-related AEs in cohort 2 of KEYNOTE-087. The Company were asked to explain how these results relate to the comparators (clarification question A30 a)[8]. The Company were also asked to describe how does the type of AEs experienced by the participants of cohort 2 of KEYNOTE-087 relate to the AEs experienced by patients treated with a comparator (clarification question A30 b). The Company commented that, *"As stated in the CS, in general, pembrolizumab was well tolerated by patients in Cohort 2 of KEYNOTE-087, with a manageable safety profile. The safety profile of pembrolizumab is considered acceptable in the context of alternative therapies, such as standard chemotherapy regimens. Clinicians in the NHS have considerable experience with pembrolizumab in a variety of indications, including cHL. No safety signal was identified in KEYNOTE-087 that differs from the large portfolio of pembrolizumab trials that have already reported. The data presented from KEYNOTE-087 show that most AEs experienced were low grade, and did not result in study discontinuation. We also spoke to a clinician to elicit further information for this CQ response who confirmed that pembrolizumab is typically a much better tolerated treatment than standard chemotherapy regimens. The AEs from comparator studies are available in the cost effectiveness section of the CS, the clinical effectiveness section focuses on the safety profile of pembrolizumab from KEYNOTE-087*." (page 24[9]) The EAG does not consider this to be a reasonable explanation for the relationship of these AEs to the comparator in clinical effectiveness. More data and analyses are still needed to illustrate the relationship with comparators. The EAG believes this also fails to account for AEs in cohort 2 in relation to comparators.
- The EAG noted that there are no data provided for AEs experienced by patients treated with any of the comparators. The Company were asked to provide the necessary data and an appropriate analysis and comparison (clarification question A31)[8]. The Company responded by stating that, *"These data are provided in the cost-effectiveness section of the CS. Grade 3+ AE rates are estimated separately for all of the comparators and weighed averages are used in the economic model along with utility decrements and durations to estimate QALY losses for each treatment arm. In the base case, QALY losses were -0.0009 for pembrolizumab and -0.0039 for the weighted SoC."* (page 24[9]). The EAG consider that there is remaining uncertainty as AEs for clinical effectiveness still lack the necessary comparator data and analyses. Data of AEs in cost effectiveness is not considered to be a reasonable justification for not providing data on AEs for clinical effectiveness. In Section B.3.3.4 the Company states that the sources used to identify AEs of Grade 3+ for all the SoC treatment options mostly used the same sources as

TA540 which are unlikely to match exactly the population of interest. The EAG does not consider their use appropriate for this submission, as discussed further in the cost effectiveness Section (4.2.7).

- The Company has chosen to present the AEs data for all the cohorts of KEYNOTE-087. The EAG requested justification for this decision (clarification question A32). The Company state that *"MSD consider AE data for Cohort 2 to be the most relevant, particularly in the economic evaluation, but decided to present AE results for all three cohorts within the clinical section to demonstrate the consistency of the safety profile of pembrolizumab across the groups of patients with R/R cHL."* The EAG need to emphasise that only the cohort 2 population is of interest, whereas companies described the entire population in the AEs Section. The EAG believes that the focus should be on the cohort 2 population and to avoid confusion with the total population. It would be reasonable to expect that the specific characteristics of the population in cohort 2 would affect their experience of AE.
- There is a lack of data on AEs experienced by patients from the SACT dataset. The Company has provided no explanation on this omission.

3.2.7 Ongoing studies

KEYNOTE-087 is an ongoing trial. The next database lock is anticipated in [redacted]. Trial NCT02824029 was also identified as an ongoing study that is anticipated to provide additional evidence in the next 12 months. The intervention of this trial is Ibrutinib for R/R cHL patients. The Company does not consider this trial to fully align with the population of interest specified in the decision problem.

EAG comment: The EAG inquired why the Company did not consider trial NCT02824029 to align to the population relevant to the decision problem. The Company responded that "*At the time of writing, ibrutinib is not part of established clinical practice in the UK for the management of R/R cHL, therefore, had data been available, they would not have been used to address the decision problem. However, MSD considered that the study could be relevant in the future, should ibrutinib be evaluated through the NICE TA process*." (page 12[9])

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

The Company did not identify any studies that evaluated a comparison of interest. Instead, they used two retrospective observational studies: Cheah (2016)[3] and Eyre (2017)[5]. These studies provided datasets for generating estimates of comparative effectiveness versus pembrolizumab in R/R cHL. In addition, a retrospective study based on the real world data derived from the SACT database were also used to inform estimates of pembrolizumab effectiveness. Data were provided as aggregated data and pseudo individual patient data (IPD), as appropriate, for OS. An overview of the pembrolizumab and comparator data sources utilised in the ITCs is presented in Table 3.24.

Table 3.24: Summary of data sources for ITC of pembrolizumab versus comparator

Target population	Pembrolizumab data source	Comparator data source	Outcome
Participants with R/R cHL who have failed to respond or relapsed BV and are ineligible for autoSCT (cohort 2 of KN087)	Pembrolizumab from KEYNOTE-087: IPD (Database cutoff: 15 March 2021) Pembrolizumab from SACT database: pseudo-IPD and summarised baseline characteristics	BV from Eyre (2017)[5] study pseudo-IPD and summarized baseline characteristics Post BV treatments from Cheah (2016)[3]. Pseudo-IPD and summarised baseline characteristics	OS

Target population	Pembrolizumab data source	Comparator data source	Outcome
Based on Table 9 of the CS Appendices[17] autoSCT = autologous stem cell transplant; BV = brentuximab vedotin; CS = company submission; IPD = individual patient data; ITC = indirect treatment comparison; R/R cHL = relapsed or refractory classical Hodgkin lymphoma; OS = overall survival; SACT = systemic anti-cancer therapy			

3.3.1 Cheah (2016)

Cheah (2016)[3] is a retrospective observational study that focused on outcomes after treatment with BV in patients with cHL who were either refractory to BV or experienced disease relapse. The study evaluated the records for patients treated in one centre in the USA, between June 2007 and January 2015. Inclusion criteria were histologically confirmed cHL as well as receiving treatment with BV for relapsed cHL and subsequently experiencing disease progression at any time after treatment with BV. Ninety-seven patients met the study inclusion criteria. Baseline characteristics for the full cohort and the subgroup of those for whom information was available at the time of documented disease progression are presented in Table 3.25. In the full cohort, the majority of participants were male (53%) and were at Stage 2 (43%) and had an ECOG performance status of 0 (84%). The median age of the participants was 28 years (range 16-83) and the median number of previous lines of therapy was three (range of 0 – 9). Before BV treatment, 70 (72%) patients had undergone SCT, the majority received autoSCT (n=66). An additional 10 (10.3%) patients received alloSCT after autoSCT but before BV. The Company highlighted that "*As noted by the EAG in TA540, the large proportion of people having undergone SCT does not align with the autoSCT naïve status of people enrolled into cohort 2 of KEYNOTE-087 or receiving pembrolizumab through the CDF*." (page 62[4]).

Treatments administered after BV comprised investigational agents, gemcitabine, bendamustine, another alkylator, BV retreatment, platinum-based treatment, and autoSCT. The Company highlighted that "*The single centre, retrospective design, US setting, youth and SCT history of the patients are major concerns. A priori, the patient level factors just mentioned suggest these patients would have a favourable prognosis compared to the population treated with SoC in this Technology Appraisal.*" (page 61-62[4]). OS was measured from the time of progression post-BV to death. A summary of results is presented in Table 3.26.

Table 3.25: Baseline characteristics reported in Cheah (2016)

Characteristic	Before treatment with BV (N=97)	At documented DP after treatment with BV (N=89)
Female	46 (47%)	N/R
Median age (range), years	28 (16–83)	32 (18–84)
Age >45	11 (11%)	14 (14%)
Stage		**n=84**
1	2 (2%)	2 (3%)
2	40 (43%)	25 (30%)
3	23 (25%)	18 (21%)
4	29 (31%)	39 (46%)
B symptoms	n= 94; 55 (60%)	n=86; 7 (8%)
Histologic subtype	**n=95**	**N/R**
Nodular sclerosing	84 (88%)	–
Mixed cellularity	7 (7%)	–

Characteristic	Before treatment with BV (N=97)	At documented DP after treatment with BV (N=89)
Lymphocyte rich	1 (1%)	–
Unknown	3 (3%)	–
Haemoglobin <105 g/l	n=35; 10 (29%)	n=51; 18 (35%)
Lymphocytes <0.6 x 10^9/I	n=33; 4 (12%)	n=46; 19 (41%)
White cell count >15 x 10^9/l	n=34; 11 (32%)	n=82; 4 (5%)
Albumin <40 g/l	n=31; 19 (61%)	n=82; 23 (28%)
ECOG	**n=61**	**n=81**
0	51 (84%)	33 (41%)
1	10 (16%)	44 (54%)
2	–	3 (4%)
3	–	1 (1%)
Any extranodal site	n=80; 22 (27%)	n=88; 31 (35%)
Max tumour bulk >10 cm	n=40; 15 (37%)	–
International prognostic score	**n=54**	
Good (0 to 1)	8 (15%)	N/R
Intermediate (2 to 3)	41 (76%)	N/R
High (4 to 7)	5 (9%)	N/R
Median prior lines of therapy (range)	n=94; 3 (0–9)	N/R
Initial chemotherapy		
ABVD	81 (84%)	N/R
ABD	4 (4%)	N/R
Other	11 (11%)	N/R
Initial radiotherapy	28 (29%)	N/R
Response to frontline therapy	**n=93**	
CR	33 (35%)	N/R
PR	28 (30%)	N/R
SD	10(11%)	N/R
PD	22 (24%)	N/R
Disease progression during BV therapy	N/R	n=97; 72 (74%)
Disease progression after BV diagnosed by imaging	N/R	n=92; 92 (100%)
Maximum tumour diameter ≥4 cm	N/R	n=69; 18 (26%)
Biopsy carried out	N/R	n=88; 18 (20%)
Biopsy CD30 positive by immunohistochemistry	N/R	n=18; 14 (78%)

Characteristic	Before treatment with BV (N=97)	At documented DP after treatment with BV (N=89)
Median number of treatments after BV	N/R	n=86; 2 (0–8)

Based on Table 10 of the CS Appendices[17]
ABVD = doxorubicin, bleomycin, vinblastine and dacarbazine; AVD = doxorubicin, vinblastine, dacarbazine; BV = brentuximab vedotin; CR = complete response; CS = company submission; DP = disease progression; ECOG = Eastern Cooperative Oncology Group; N/R = not reported; PD = progressive disease; PR = partial response; SCT = stem cell transplant; SD = stable disease

Table 3.26: Summary of results reported in Cheah (2016)

Outcome after progression after treatment with BV (n=67)	Result
CR, n (%)	12 (15)
PR, n (%)	15 (19)
Median PFS, months	3.5
Median OS, months	25.2

Based on Table 11 of the CS Appendices[17]
BV = brentuximab vedotin; CR = complete response; CS = company submission; OS = overall survival; PFS = progression-free survival; PR = partial response

3.3.2 Eyre (2017)

Eyre (2017)[5] is a retrospective, multicentre study evaluating outcomes in 99 SCT-naïve patients with R/R cHL treated with BV monotherapy. Data came from nine UK centres between May 2011 and July 2016. Further inclusion criteria were: at least 2 prior lines of treatment and being treated with BV with the intention of undergoing subsequent SCT. Overall survival was calculated from the initiation of BV monotherapy to date of death and censored at the date of last follow-up. Progression-free survival was calculated from the initiation of BV to the time of relapse, DP, death, or censored at the date of last follow-up. Baseline characteristics are reported in Table 3.27 and a summary of outcomes in Table 3.28. The majority of patients were female (55%), had an ECOG performance score of 0 (52%), were at Stage 3-4 (71%) and had 2 prior lines of therapy, prior-BV (71%).

Table 3.27: Baseline characteristics of SCT-naïve patients receiving BV in Eyre (2017)

Characteristic	BV SCT-naïve (N=99)
At diagnosis	
Median age (range), years at diagnosis	32 (13–70)
Gender	
Male	45 (45%)
Female	54 (55%)
ECOG performance status at diagnosis (n=86)	
0	45 (52%)
1	36 (42%)
>1	5 (5%)
Histological subtype at diagnosis (n=89)	
Nodular sclerosis	75 (84%)
Mixed cellularity	12 (13%)

Characteristic	BV SCT-naïve (N=99)
Lymphocyte deplete	1 (1%)
Lymphocyte rich	1 (1%)
Ann Arbor staging at diagnosis (n=98)	
1–2	28 (29%)
3–4	70 (71%)
Bulk at diagnosis >10 cm (n=95)	
No	75 (79%)
Yes	20 (21%)
Duration of first remission	
Earliest remission to relapse, months (n=66)	Median 6.0 (range 0.7–74.0)
Risk factors at first relapse	
Haemoglobin, g/l (n=80)	Median 122 (range 66–153)
≥120 g/l	41 (51%)
<120 g/l	39 (49%)
Extranodal disease (n=94)	
Yes	44 (47%)
No	50 (53%)
B symptoms (n=88)	
Yes	33 (38%)
No	55 (62%)
Ann Arbor stage (n=94)	
1–2	27 (29%)
3–4	67 (71%)
Median time from last treatment to BV, months (n=94)	2.5 (range 0.7–34.8)
Median time from initial diagnosis to BV, months (n=99)	14.5 (range 4.0–190.9)
Prior lines of therapy pre-BV (n=99), n (%)	
2	70 (71%)
3	24 (24%)
4	5 (5%)
Median number of prior chemotherapy lines	2 (range 2–4)
Response to BV (n=96)	
ORR	54 (56%)
C(M)R/CR/CRu	24 (25%)/3 (3%)/1 (1%)
PMR/PR	2 (2%)/24 (24%)
SD	8 (8%)
PD	34 (35%)
Cycles of BV given, median	4 (range 1–9)
Treatment summary after receiving BV (n=99)	
No further treatment	8 (8%)

Characteristic	BV SCT-naïve (N=99)
ASCT	15 (15%)
AlloSCT	19 (19%)
Chemotherapy followed by autoSCT	8 (8%)
Chemotherapy followed by alloSCT	19 (19%)
Chemotherapy with no SCT	30 (31%)

Based on Table 12 of the CS Appendices[17]
alloSCT = allogeneic stem cell transplantation; autoSCT = autologous stem cell transplantation; BV = brentuximab vedotin; CR =complete response; C(M)R = complete (metabolic) response; CRu = unconfirmed complete response; CS = company submission; ECOG = Eastern Cooperative Oncology Group; ORR = overall response rate; PR = partial response; P(M)R = partial (metabolic) response; PD = progressive disease; SCT = stem cell transplantation; SD = stable disease

Table 3.28: Summary of outcomes from Eyre (2017)

Outcome	Result
CR, n (%)	29%
PR, n (%)	27%
Median PFS (95% CI), months	5.6 (4.4 to 12.2)
No SCT (n=38)	3.0 (2.4 to 4.4)
AutoSCT (n=23)	NR (17.0 to NR)
AlloSCT (n=38)	NR (5.6 to NR)
Median OS (95% CI), months	37.2 (18.3 to NR)
No SCT (n=38)	12.2 (8.1 to 18.3)
AutoSCT (n=23)	NR (27.0 to NR)
AlloSCT (n=38)	NR (37.2 to NR)

Based on Table 13 of the CS Appendices[17]
alloSCT = allogeneic stem cell transplantation; autoSCT = autologous stem cell transplantation; CI = confidence interval; CR = complete response; CS = company submission; NR = not reached; OS = overall survival; PFS = progression-free survival; PR = partial response; SCT = stem cell transplantation; SD = stable disease

3.3.3 Further evidence for the ITC

Further sources of evidence for the ITC analyses came from KEYNOTE-204 and TA524[36]. KEYNOTE-204 is an ongoing, randomised, open-label, phase III study of pembrolizumab compared with BV in subjects with R/R cHL. According to the eligibility criteria, participants were to have R/R cHL and received at least 1 prior multi-agent chemotherapy regimen. Prior treatment with BV (or a BV-containing regimen) was allowed, as long as the participants had responded to the BV or BV-containing regimen. The Company acknowledged that the participants of this trial do not align with the population of the decision problem in that:

- "*KEYNOTE-204 includes patients who have had prior autoSCTs, whereas the decision problem focuses on patients who cannot have autoSCT*
- *"KEYNOTE-204 is at an earlier line of therapy and are eligible for treatment with BV, whereas the decision problem focuses on patients who have already been treated with BV.* (page 229[17])

Nevertheless, the Company states that the trial should be used as it the only source of randomised evidence of pembrolizumab versus an active comparator in R/R cHL.

TA524[36] examined BV for treating CD30-positive HL. According to the guidance: BV was recommended as an option for treating CD30-positive HL in adults with relapsed or refractory disease, only if:

- they have already had autoSCT or
- they have already had at least two previous therapies when autoSCT or multi-agent chemotherapy are not suitable.

EAG comment:

- The key limitation of Cheah (2016)[3] regarding the population of patients, is that out of the 97 patients, 70 (72%) had previous SCT, 66 had autoSCT and four had alloSCT. The remaining 27 (28%) patients did not undergo consolidative transplant. Therefore only 28% of patients match the population in the decision problem. Separate data for this subgroup are not available in the publication.[3] Further misalignments on the population baseline characteristics are evident in terms of age, disease stage, ECOG performance status, B symptoms, haemoglobin, lymphocytes, white cell count and albumin. In addition, the population comes from only one centre in the USA and the generalisability to the UK setting is questionable.
- The use of the entire population from Cheah (2016)[3] as the comparator in a naïve ITC will probably result in underestimating the effect of the intervention.
- Eyre (2017)[5] includes only SCT-naïve patients which is aligned with the population of the decision problem. The interventions in the study include both BV as well as treatments subsequent to BV, the latter being the appropriate comparator for this CS. Unfortunately, subgroup data for the patients receiving treatment further to BV is not readily available in the publication[5]. The Company has used the outcomes of the entire cohort in the ITC analysis. This misalignment is also of course, affecting the line of therapy besides the comparators. An earlier line of therapy would be associated with healthier more responsive patients. Therefore, the effect of the intervention in a naïve ITC would be underestimated.
- KEYNOTE-204 is and RCT comparing pembrolizumab to BV in subjects with R/R cHL. BV is not a comparator of interest for this STA. This issue is further discussed in Section 3.4.5.
- TA524 is used in two ITCs: Bucher ITC (KEYNOTE-204 and TA524) and Bucher ITC (SACT versus Eyre (2017)[5] and TA524) (see Section 3.4) the first one forming the Company's base-case scenario. The comparator in TA524 (i.e., SoC) was single-agent chemotherapy including:
    - vinblastine
    - etoposide
    - gemcitabine

  The composition of SoC in this CS as presented in Table 50 is notably different. Vinblastine is not included in the basket of comparators, while etoposide and gemcitabine are only included as part of a multi-agent regimen and not as a single-agent chemotherapy as in TA524.
- Given the profound limitations of the data sources used by the Company and the use of SACT data for pembrolizumab, the EAG requested if historical SACT data (pre-pembrolizumab) could be obtained to inform the effectiveness of SoC. The Company response to this was that they did not have access to any such data.[8]

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

The objective of the ITC analysis was to compare the clinical efficacy, in terms of OS, of pembrolizumab versus the comparators for patients with R/R cHL. The Company performed three types of analyses: naïve indirect comparison, matched adjusted indirect treatment comparison (MAIC) and Bucher ITC.

The naïve indirect comparison was used to compare pembrolizumab using data from SACT and KEYNOTE-087, cohort 2 with SoC using data from Cheah (2016)[3] and Eyre (2017)[5], generating four scenarios. Overall survival was compared using Cox proportional hazard models using pseudo-IPD of selected comparator arm as well as data from cohort 2 from KEYNOTE-087 and the SACT dataset. The model used treatment as a single covariate. Results were expressed in HRs, 95% CIs and p-value, together with median survival time, (95% CI) and the number and percentage of events by treatment arm.

The MAIC was used to compare pembrolizumab using data from KEYNOTE-087, cohort 2 with SoC using data from Cheah (2016)[3] and Eyre (2017)[5], generating two scenarios. Data from patients from cohort 2 of KEYNOTE-087 were re-weighted to match the average baseline characteristics of patients retrieved from Cheah (2016)[3] and Eyre (2017)[5]. The adjustment of IPD data from KEYNOTE-087, cohort 2 to match the aggregated baseline characteristics from the comparator studies was carried out following the methods described in Signorovitch (2010)[37] and Signorovitch (2012)[38]. As IPD were not available for the comparator studies, a method of moments was used to estimate the weight estimate instead the of the usual maximum likelihood approach to estimate the parameters of the propensity score model. In this method the IPD of participants receiving the treatment of interest was re-weighted to exactly match their mean baseline characteristics to the aggregated data available in the literature for the control. Three baseline characteristics were identified and selected as potential effect modifiers: age, sex and ECOG performance status score (0 versus 1). According to the Company the potential effect modifiers were identified by clinical experts[6].

Bucher ITCs were also used to provide supporting evidence on the clinical effectiveness of pembrolizumab versus SoC with BV as the common comparator. The HR for pembrolizumab versus SoC was calculated by multiplying the HR for pembrolizumab versus BV and the HR for BV versus SoC together. The standard error of the log HR was obtained by taking the square root of the sum of the two variances. The source of the presented evidence was the Markov trace from the BV versus SoC model used in TA524[36] on BV for treating CD30-positive HL. From KEYNOTE-204, data from a SCT-naïve subgroup of patients was used. In the Bucher ITCs the treatment effect of pembrolizumab versus SoC was assumed to be the sum of the treatment effects for pembrolizumab versus BV and BV versus SoC. Simple Bucher ITCs were used to calculate the implied 4-year OS HR for pembrolizumab versus SoC.

The Company acknowledges the presence of inherent bias in the ITC. Regarding the ITC of pembrolizumab (KEYNOTE-087, cohort 2) versus SoC (Cheah (2016)[3]), the baseline characteristics for SoC may not be totally representative for the set of patients used in the efficacy analyses. The characteristics used for matching of patients receiving pembrolizumab to those treated with SoC correspond to the number of patients for whom baseline characteristics at the time of DP are available (n=89), whereas the corresponding pseudo-IPD are derived from patients who actually received treatment after progression on BV (n=79). The Company stated that the direction of bias is unknown.

In addition, the eligibility criteria for KEYNOTE-087 and Eyre (2017)[5] do not completely match. Clinical opinion suggests that any bias would be against the experimental arm, and as such the ITC could be considered as a conservative comparison. The Company noted that "*the analyses should be interpreted with extreme caution, as the incomplete matching of characteristics is only one of many factors to be considered when judging the level of bias present in the resulting ITC and, there is the potential for residual bias. The patients in Eyre (2017) appeared to be healthier than those forming cohort 2 of KEYNOTE-087, being, for example, younger and having received fewer lines of previous therapy:*

- *Age: median age of 40 years for cohort 2 of KEYNOTE-087 versus 32 years in Eyre (2017)[5];*
- *Prior lines of therapy: median of 4 for cohort 2 of KEYNOTE-087 versus 2 in Eyre (2017)[5];*
- *Sex: proportion of females: 46.91% for cohort 2 of KEYNOTE-087 versus 54.55% in Eyre (2017)[5].*

Please also note that the assessment of baseline characteristics is limited to a few common characteristics as reported in both the CSR for KEYNOTE-087 and in Eyre (2017)." (page 47-18[17])

The Company replicated the MAIC presented in TA540[1] using Cheah (2016)[3] to represent SoC. The Company maintains that Cheah (2016)[3] "*...remains the most appropriate dataset to generate estimates of comparative clinical effectiveness versus pembrolizumab for those with R/R cHL who are SCT-naïve and did not respond to BV*" (page 60[4]).

3.4.1 Results of the ITC based on Cheah (2016)

A summary to the OS estimates derived from the ITC of pembrolizumab versus SoC based on Cheah (2016)[3] is presented in Table 3.29. Regarding cohort 2 from KEYNOTE-087, the naïve and MAIC analyses produced similar estimates of OS for pembrolizumab versus SoC based on Cheah (2016)[3], with HRs of 0.25 and 0.24, respectively, favouring pembrolizumab (Table 3.29); with statistically significant results (p <0.001). On the other hand, the naïve ITC for SACT versus SoC based on Cheah (2016)[3] resulted in a higher HR of 0.59, while the difference remained statistically significant (p=0.006) (Table 3.29). A comparison of baseline characteristics between KEYNOTE-087, cohort 2 and SACT cohort with Cheah (2016)[3] is presented in Tables 14 and 17 of Appendix D of the CS[17], respectively.

Kaplan-Meier curves generated for the indirect comparison of pembrolizumab versus SoC are presented in Figure 3.10 for KEYNOTE-087, cohort 2 and Figure 3.11 for the SACT dataset.

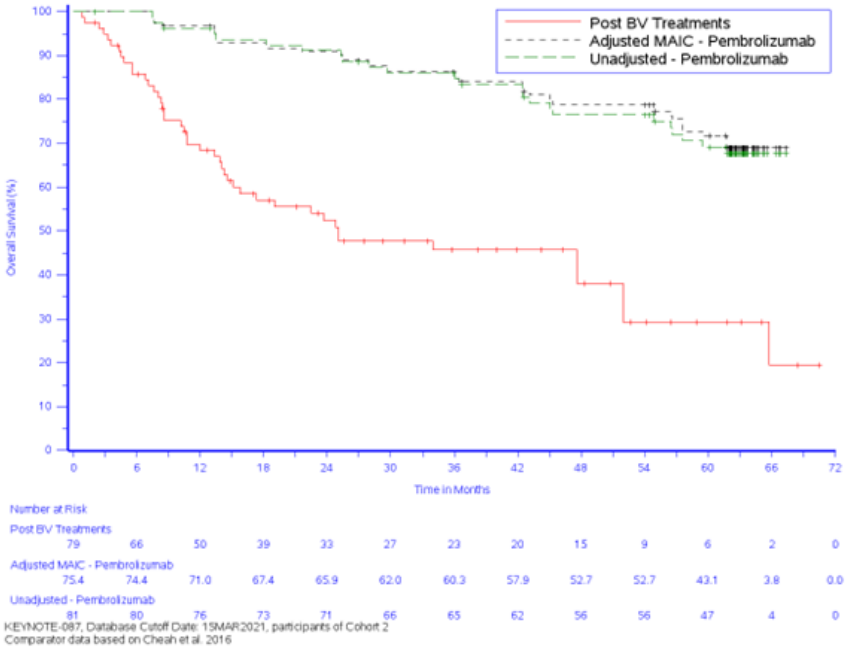
Table 3.29: Summary of OS estimates derived from indirect comparisons of pembrolizumab versus SoC based on Cheah (2016)

Dataset	Comparison	Sample size/effective sample size, n	Pembrolizumab Events, n (%)	Post BV[a] (N=89) Events, n (%)	HR (95% CI),[b] p value[b,c]
KEYNOTE-087 cohort 2	Unadjusted	81	24 (29.6)	46 (58.2)	0.25 (0.15 to 0.41) <0.001
KEYNOTE-087	MAIC	75.4[d]	21 (27.9)	46 (58.2)	0.24 (0.14 to 0.40) <0.001

Dataset	Comparison	Sample size/effective sample size, n	Pembrolizumab Events, n (%)	Post BV[a] (N=89) Events, n (%)	HR (95% CI),[b] p value[b,c]
SACT[e]	Unadjusted	215	68 (31.6)	46 (58.2)	0.59 (0.40 to 0.86) 0.006

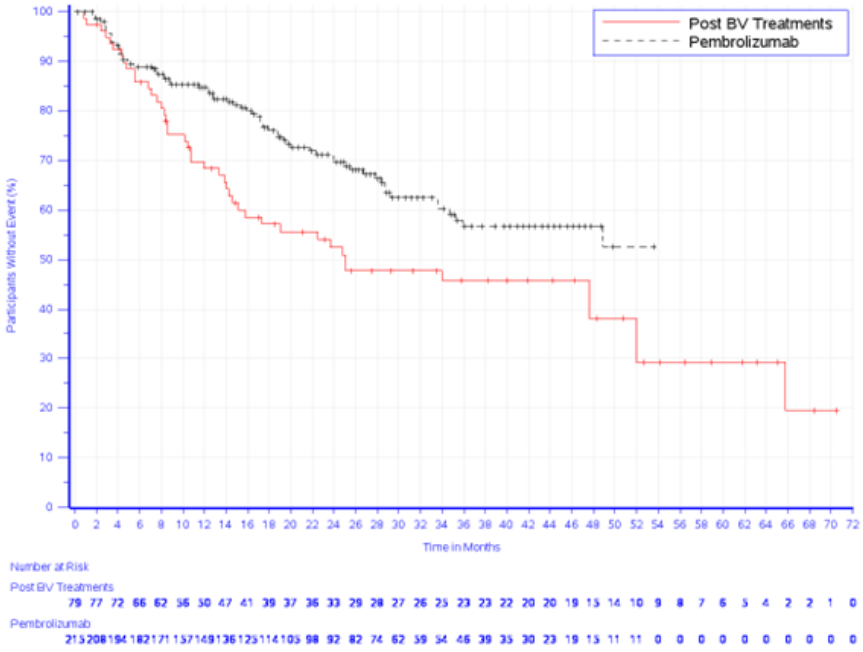
Based on Table 20 of the CS[4]
[a] Based on mAPaT population from Cheah (2016).
[b] Based on Cox regression model with treatment as a single covariate.
[c] Two-sided p-value using Wald test (Score test in case of zero event in one treatment group).
[d] Effective sample size computed as sum of weights.
[e] Based on Real World Evidence source (SACT database) in participants relapsing after treatment with or failing to respond to BV and ineligible for SCT, mAPaT.
BV = brentuximab vedotin; CI = confidence interval; CS = company submission; HR = hazard ratio; MAIC = matching adjusted indirect comparison; mAPaT = defined as analysis populations used to report comparator study results; OS = overall survival; SACT = systemic anti-cancer therapy; SoC = standard of care

Figure 3.10: Kaplan-Meier curve for OS for pembrolizumab versus SoC based on Cheah (2016) with pembrolizumab data derived from Cohort 2 from KEYNOTE-087



Based on Figure 9 of the CS[4]
BV = brentuximab vedotin; CS = company submission; MAIC = matching-adjusted indirect treatment comparison; OS = overall survival; SoC = standard of care;

Figure 3.11: Kaplan-Meier curve for OS for pembrolizumab versus SoC based on Cheah (2016) with pembrolizumab data derived from the SACT dataset



Based on Figure 10 of the CS[4]
BV = brentuximab vedotin; CS = company submission; OS = overall survival; SACT = systemic anti-cancer therapy; SoC = standard of care

3.4.2 Results of the ITC based on Eyre (2017)

A summary to the OS estimates derived from the ITC of pembrolizumab versus SoC based on Eyre (2017)[5] is presented in Table 3.30. The two ITCs, naïve and MAIC, based on KEYNOTE-087, cohort 2 had similar results, with HRs of 0.23 and 0.21, respectively, favouring pembrolizumab (Table 3.30); with statistically significant results (p <0.001). On the other hand, the naïve ITC for SACT versus SoC based on Eyre (2017)[5] resulted in a higher HR of 0.66, while the difference between the treatments was statistically significant (p=0.040) (Table 3.24). A comparison of baseline characteristics between KEYNOTE-087, cohort 2 and SACT cohort with Eyre (2017)[5] is presented in Tables 19 and 23 of Appendix D of the CS[17], respectively.

Table 3.30: Summary of OS estimates derived from indirect comparisons of pembrolizumab versus SoC based on Eyre (2017)

Dataset	Comparison	Sample size/ effective sample size, n	Pembrolizumab Events, n (%)	BV[a] (N=99) Events, n (%)	HR (95% CI),[b] p value[b,c]
KEYNOTE-087 cohort 2	Unadjusted	81	24 (29.6)	37 (37.4)	0.23 (0.12 to 0.42) <0.001[d]
KEYNOTE-087	MAIC	77.7	21 (27.0)	37 (37.4)	0.21 (0.12 to 0.37) <0.001
SACT	Unadjusted	215	68 (31.6)	37 (37.4)	0.66 (0.44 to 0.98) 0.040

Dataset	Comparison	Sample size/ effective sample size, n	Pembrolizumab Events, n (%)	BV[a] (N=99) Events, n (%)	HR (95% CI),[b] p value[b,c]
Based on Tables 20, 22, 24 of the CS appendices[17] [a] Based on mAPaT population from Eyre (2017)[5] [b] Based on Cox regression model with treatment as a single covariate [c] Two-sided p-value using Wald test (Score test in case of zero event in one treatment group) [d] This result is reported for MAIC in Table 33 and Table 3 in the response to request for clarification[9], but as unadjusted in Table 20 and Table 22 of the Appendices[17] BV = brentuximab vedotin; CI = confidence interval; CS = company submission; HR = hazard ratio; MAIC = matching-adjusted indirect comparison; mAPaT = defined as analysis populations used to report comparator study results; OS = overall survival; SACT = systemic anti-cancer therapy; SoC = standard of care					

3.4.3 Results from further ITC analyses

A subgroup of patients from KEYNOTE-204 who were SCT-naïve at baseline was used to extract an HR for OS for pembrolizumab, and the data presented in TA524 were used to estimate an HR for OS for BV versus SoC. The Bucher comparison produced an HR for OS of [redacted] in favour of pembrolizumab. The Company noted that although the cohort of patients informing the Bucher ITC were not receiving 4 line treatment after BV, they considered that the analysis supports the clinical effectiveness of pembrolizumab in R/R cHL.

According to the Company, a further ITC was conducted on clinical effectiveness of pembrolizumab versus BV through MAIC using data on clinical outcomes from KEYNOTE-204 and Eyre (2017)[5], evaluating BV in SCT-naïve patients. The Company considers this population perhaps most analogous to those forming cohort 2 of KEYNOTE-087. The Company acknowledges that BV is not a comparator of interest specified in the decision problem but consider that the MAICs versus BV are informative for decision-making.

3.4.4 Summary of ITCs

A summary of the OS HR estimates of the ITCs executed by the Company are presented in Table 3.30. The Company used estimate 1 (Bucher ITC of KEYNOTE-204 and TA524) as their base analysis. They based their decision on KEYNOTE-204 being the only relevant source of randomised evidence, and a comparison that had been validated in TA524[36]. In addition, the Company stated "*We felt that, since both these sources of evidence are anchored in some way and not obviously biased in favour of one comparator. Estimate 1 using the Bucher ITC is therefore potentially more reliable than the various unanchored ITCs we conducted. The key limitation is that it relates to third line rather than fourth line patients.*" (page 104[4]). The EAG does not consider KEYNOTE-204 directly relevant to this STA as is examines pembrolizumab compared with BV in addition TA524 which examined BV did not use comparators relevant to this STA. Further details are provided in Sections 3.3.3 and 3.5.

The EAG considering the limitations of each estimate concluded that estimate 5 (ITC SACT versus Cheah (2016)[3]) was the most appropriate choice, albeit with its own limitations, and also explored estimate 7 (MAIC KN087 versus Cheah (2016)[3]) in a scenario analysis. Cheah (2016)[3] remains the only source of evidence known to the EAG that reports outcomes for the most relevant comparators and at the most relevant line of therapy. More details on the limitations of the ITC analyses are reported in Section 3.4.5.

Table 3.31: Summary of OS HR estimates

Estimate Number	Comparison	HR (CI)	Key limitations noted by the Company
1	Bucher ITC (KEYNOTE 204 and TA524)	[redacted]	Two third line studies, assumed s.e. from TA524
2	KEYNOTE-204 OS	[redacted]	Third line study, control arm is BV
3	Bucher ITC (SACT versus Eyre (2017)[5] and TA524)	0.41 (0.22 - 0.77)	100% patients fit for transplant in Eyre (2017)[5] study, assumed s.e. from TA524
4	ITC SACT versus Eyre (2017)[5]	0.66 (0.44 - 0.98)	Eyre (2017)[5] is third line BV study, 100% fit for transplant
5	ITC SACT versus Cheah (2016)[3]	0.59 (0.4 - 0.86)	71% had prior transplant, 30% received investigational agents
6	MAIC KEYNOTE 087 versus Eyre (2017)[5]	*Before matching:* 0.23 (0.12 - 0.42) *After matching:* 0.21 (0.12 to 0.37)	4 plus KEYNOTE 087 applicability concerns. Comparator BV
7	MAIC KEYNOTE 087 versus Cheah (2016)[3]	0.24 (0.14 - 0.4)	5 plus KEYNOTE 087 UK applicability concerns

Based on Table 33 of the CS[4] and Table 22 of Appendix D[17]
* Note: these are parameterised as reciprocal HRs in the economic model because OS data must be anchored to the pembrolizumab arm.
BV = brentuximab vedotin; CI = confidence interval; CS = company submission; HR = hazard ratio; ITC = indirect treatment comparison; MAIC = matching adjusted indirect comparison; OS = overall survival; SACT = systemic anti-cancer therapy; TA = Technology Appraisal; UK = United Kingdom

3.4.5 Indirect treatment comparison limitations

The Company has highlighted a series of limitations in the ITCs presented in the CS:

- The comparative analyses were limited to unanchored ITC due to the use of single-arm and observational studies.
- Only naïve-unadjusted comparison was executed for pembrolizumab data from the SACT cohort versus SoC, due to the lack of IPD.
- The naïve-unadjusted ITCs should be interpreted with caution due to the differences across studies in terms of study design:
    - prospective versus retrospective
    - single-arm versus observational
- The naïve-unadjusted ITCs do not match patient populations.
- There are differences in key patient baseline characteristics which could be prognostic factors or treatment effect modifiers. Patients in KEYNOTE-087, cohort 2 were older but fitter than

those in Cheah (2016)[3], in contrast to those forming the SACT dataset, who were considerably older and less fit than patients in Cheah (2016)[3]:
    - Proportion of people aged >45 years: 43.2% in cohort 2 versus ≥55% in the SACT dataset versus 14% in Cheah (2016)[3]
    - ECOG performance status 0: 54.3% in cohort 2 versus 27% in SACT dataset versus 41% in Cheah (2016)[3];
    - ECOG performance status 1: 45.7% in cohort 2 versus 40% in SACT dataset versus 54% in Cheah (2016)[3].
- Patients in KEYNOTE-087, cohort 2 and in SACT dataset were SCT-naïve, whereas 72% of patients in Cheah (2016)[3] had undergone SCT.
- According to NICE DSU[39] in unanchored comparisons it is necessary to assume that all effect modifiers and prognostic factors have been accounted for. In the MAIC for pembrolizumab versus SoC this was impossible to do.
- The variables matched in the MAIC for pembrolizumab versus SoC were determined by the characteristics reported in Cheah (2016)[3]. These characteristics are unlikely to be all relevant prognostic variables and treatment effect modifiers for OS. It is likely that the MAIC results contain systematic error, the extend of which could not be quantified.
- Cheah (2016)[3] acknowledges selection bias in their study which limits the generalisability of their findings. The authors note that outcomes among patients with BV-resistant disease outside academic centres may be less favourable.
- The diversity observed in post-progression treatment regimens does not reflect clinical practice in England.

Nevertheless, the Company reiterated that due to their inability to identify new relevant studies within their SLR, Cheah (2016)[3] remains the most relevant study to inform estimates of comparative clinical effectiveness for this STA. The Company acknowledged that the limitations of the ITCs lead to uncertainty in the comparative effect estimates for OS but note that all results favour pembrolizumab and reach statistical significance.

EAG comment:

- The EAG agrees with the Company's list of key limitations associated to the ITCs presented in Section 3.4.5. Further discussion of the ITC outcomes is included in the cost effectiveness Section.
- In Section B.2.9.2, the Company describes a MAIC of pembrolizumab versus BV using data from KEYNOTE 204 and Eyre (2017)[5]. Nevertheless, in Table 33 of the CS where the OS HR estimates from the ITCs are reported, this ITC analysis is not included. Instead, a Bucher ITC of SACT versus Eyre (2017)[5] and TA524 is reported, which is not mentioned in this Section.
- There is an inconsistency in one of the ITC results. In Table 33 of the CS[4] and in Table 3 of the response to the request for clarification[9], the Company reports ITC #6 as: MAIC KEYNOTE-087 versus Eyre (2017)[5] with an OS HR of 0.23 (95% CI, 0.12 - 0.42). On the other hand, in Table 20 of the CS[4] and Table 22 of Appendix D[17], they report an OS HR of 0.23 (95% CI, 0.12 - 0.42) as the result before matching and an OS HR of 0.21 (95% CI, 0.12 to 0.37) as the result after matching. It is likely that the unadjusted values were used instead of the matched.
- The EAG noted that only three effect modifiers were selected for the MAIC process: age, sex and ECOG performance status, while other prognostic factors and possible effect modifiers were noted by the Company. These included: time to initial relapse after high-dose chemotherapy and autoSCT, time to relapse and history of primary refractory disease. The EAG

requested that further characteristics should be included in the analysis. The Company responded that "*The factors highlighted by the EAG as omitted from the MAIC were not reported in all sources used to inform the MAIC and so adjustment was not possible*" (page 19[9]).

- The MAIC analysis is based on the population characteristics reported in Cheah (2016)[3]. These characteristics may not fully represent UK clinical practice.
- The EAG inquired how the treatments received by patients in Cheah (2016)[3] compared to the treatments received by patients in KEYNOTE-087, cohort 2 and the SACT cohort. The Company replied that "*MSD have no details on prior lines of treatment for the SACT cohort, except that they must have satisfied the mentioned criteria to have been treated with pembrolizumab through the CDF*." (page 20[9]). The lack of information of the specific prior treatments leads to further uncertainty regarding comparability of the populations.
- The EAG requested that the methods for the ITC analysis involving KEYNOTE-204 would be provided. The Company responded "*Overall survival data from the subgroup of patients who were treated at third line is reported in Appendix P. The hazard ratio (HR) from this analysis was used in a simple Bucher indirect treatment comparison with the HR from NICE TA524 (see response to A27c below). As detailed in section B3.3.1.3 of the CS, the HR for pembrolizumab vs SoC was calculated by multiplying the HR for pembrolizumab vs BV and the HR for BV vs SoC together. The standard error of the log HR was obtained by taking the square root of the sum of the two variances. The upper and lower confidence limits on the natural scale were the exponentials of the limits on the log scale.*" (page 20-21[9]). The EAG is satisfied with this response.
- The EAG requested that the Company would provide the full methods used and the full analysis results for the Markov trace from the BV versus SoC model used in TA524. The Company has replied that "*The method is already described in full in section B.3.3.1.3 (page 101) of the CS. Briefly, we digitised the OS Markov trace to 4 years, created a new curve that applied a constant HR to the SoC Markov trace and varied this constant HR until the visual fit to the BV Markov trace was optimised. We observed a close visual fit across the 4-year time horizon using a HR of 0.62 and because of this, the proportional hazards assumption appeared to be reasonable. We then confirmed with clinicians at the advisory board that an HR of 0.62 was reasonable in their experience. Of course, it had already been considered reasonable by implication by the NICE Committee assessing BV and recommending BV in TA524.*" (page 21[9]) Given the novelty of using a Markov trace as an input to a further model, the EAG requested that the Company would justify this methodology/methods by citing relevant references or guidance. The Company replied "*We were unable to find references establishing a precedent for this method but, conceptually, it is sensible. In the absence of actual data on the SoC (a bundled comparator mostly consisting of generic chemotherapies), the Markov trace from TA524 represents the best available evidence on the treatment effect of BV on OS vs. SoC in R/R cHL. The approach has the advantage that, in an evidence light area, the comparator arm has been drawn from a model based on assumptions that have previously been accepted by NICE. Furthermore, clinicians at an advisory board confirmed that the estimated HR for the comparison between BV and SoC was plausible. MSD acknowledge that this approach is associated with uncertainty, therefore, several alternative scenarios were presented, the strengths and limitations of which are elaborated on further in the CS and our response to B12 below*." (page 21[9]). The use of non-peer reviewed and verified methods is adding to the already high uncertainty of the ITC analyses presented by the Company.

- The Company has used evidence of BV treatment throughout the CS and in the ITCs. As BV is not a comparator defined either in NICE scope or in the Company's decision problem, but in fact is the treatment in the previous line of therapy (third line of treatment in the UK setting) the EAG suggested that the Company would remove the BV evidence input from any analysis as it is not relevant to this submission. The Company responded "*BV is relevant to this submission in that, prior to the approval of pembrolizumab, it was the standard of care in R/R cHL and comprises one arm of the only source of randomised evidence in this setting. It is also relevant in that "3L+" and "4L+" trials in R/R cHL are not tightly prescriptive by treatment line; for example, 29% of patients in Eyre 2017 were 4L+, Cheah 2016 has a median of 3 but a range of 0 to 9 prior therapies listed, in KEYNOTE-204 the percentage who were 3L vs 4L+ was 21% vs 16%, in KEYNOTE-087 the proportion of patients who were 4L vs 5L+ was 46% vs 54%. Consequently, marketing authorisations in the area typically cover multiple lines. It is also relevant in that BV is considered to be a more effective treatment than standard chemotherapy in R/R cHL.*

  If pembrolizumab is more effective than BV then it must, by implication, be at least that much more effective than SoC. This would only not be true if failure on BV would alter the patient characteristics such that the treatment effect of pembrolizumab would diminish but the opposite is more likely to be true. In preparing our responses, we asked a clinician to comment on the direction of bias in using the effectiveness of pembrolizumab vs BV in the 3L setting as a surrogate for effectiveness of pembrolizumab vs. SoC in the 4L (BV failed) population and he commented that the observed effectiveness would likely be greater. His reasoning was that a similar proportion of patients would respond well to pembrolizumab whereas a much lower proportion would respond to SoC than had responded to BV. This is because the 4L population is effectively an even more chemotherapy-insensitive population (BV is a chemotherapy based regimen). He commented that failure to respond to BV is unlikely to meaningfully affect a patient's ability to respond to pembrolizumab because it is a different mechanism of action. This is supported by the data from the trials; very similar CR and PR rates to pembrolizumab are observed in the no-prior SCT groups of KN204 and KN087 (CR = 27% and 26%, respectively and PR = 35% and 38%, respectively).

  The treatment effect of pembrolizumab vs BV is useful information to incorporate in the economic model because it allows decision makers to examine how various levels of plausible effectiveness affect the cost-effectiveness results using the effectiveness of pembrolizumab vs BV as a reference. KEYNOTE-204 was a 3L+ trial and this setting is 4L+ among patients who have failed BV, this is a source of indirectness of population which biases against pembrolizumab (for the reasons discussed above) rather than total lack of applicability of this evidence." (pages 22-23[9]). The EAG would like to point out that the marketing authorisation of the treatments are not entirely relevant to NICE guidelines as STAs can regularly have a narrower scope. The Company is suggesting that if pembrolizumab is more effective than BV then it must, by implication, be at least that much more effective than SoC. Although this may sound plausible, such a statement should absolutely be supported by hard evidence, which the Company has not provided. A single clinical expert's opinion, while respected, cannot trump the lack of evidence. The potential cost effectiveness implications is discussed in the cost effectiveness Section of this report.
  The cumulative uncertainties in the ITC analyses are a key issue. The naïve ITCs, MAIC and Bucher ITCs have major limitations thus neither can be considered fully reliable for decision making.

3.5 Additional work on clinical effectiveness undertaken by the EAG

No further additional work was undertaken by the EAG.

3.6 Conclusions of the clinical effectiveness Section

The CS and response to the request for clarification provided sufficient details for the EAG to appraise the literature searches conducted to identify relevant evidence on the efficacy and safety of treatments for patients with R/R cHL[9, 40]. Searches were conducted in May 2023. The EAG noted that the previous search strategies utilised in the 2017 SLR were not rerun for this update which the Company explained was due to a change in scope, however the new strategies contained a number of limitations that the EAG consider could have affected the overall recall of results. Key areas for concern were raised in the request for clarification and the EAG asked that the searches be rerun and expanded with the afore mentioned points in mind. The Company declined to carry out any additional searches stating their reasons for each point. The EAG does not accept their rationale for not doing so and remains concerned that relevant studies may have been missed. Further limitations were identified in the methods of the SLR related to study identification and eligibility criteria. The cumulative limitations speak to the overall quality of the SLR. In this CS, missing potentially relevant records/studies is even more important as the underlying basis of the analysis is that no new evidence have been published since TA540.

One study of the efficacy and safety of pembrolizumab was identified (KEYNOTE-087) and this formed the basis of the submission. KEYNOTE-087 (NCT02453594) is a well-executed, ongoing, phase II, multicentre, multi-cohort, single arm, trial of pembrolizumab in patients with R/R cHL. Cohort 2 of the study was only included in the CS, which was aligned to the decision problem of this CS. In cohort 2, 81 patients were enrolled, 10 of which came from the three study sites located in the UK. There are concerns around the generalisability of the outcomes to the UK setting. The evidence in this CS[4] are based on a 5-year data cutoff (data cutoff date 15 March 2021) with a median duration of follow-up at 62.2 months.

Further clinical data were obtained from a SACT database. After publication of TA540[1], data were collected prospectively for pembrolizumab on specific outcomes during its time in the CDF. Nevertheless, the recommendations for data collection published in the FAD were not entirely met. In addition, only data on pembrolizumab were collected and not on any comparators, which would have provided a robust evidence base for this STA. Efficacy, HRQoL and safety outcomes were provided from KEYNOTE-087 but only efficacy outcomes from the SACT database.

In KEYNOTE-087, 52 (64.2%) and 55 (67.9%) of 81 patients achieved either CR or PR as per IWG and Lugano criteria, respectively. Median time to response by IWG criteria for participants achieving CR or PR (n=52) was 2.8 months (2.2 to 11.0 months), while median DOR was 11.1 months (0.0+ to 59.0+ months). Regarding the proportion of people receiving SCT, KEYNOTE-087, cohort 2 and SACT are comparable (29.6% and a 30.2%, respectively), with the difference that the majority of patients received autoSCT in KEYNOTE-087, cohort 2 (58.3%), while the majority of patients in SACT received alloSCT (64.6%).

There is also a notable difference in the estimated mean time to SCT between cohort 2 of KEYNOTE-087 which was 30.3 months and the SACT dataset, where the median time to SCT of patients eligible for SCT was 17.5 months. Time to alloSCT, the outcome defined in the NICE final scope[2] and by the FAD of TA540[1] was not reported for either cohort.

Mean OS for KEYNOTE-087, cohort 2 was 53.7 months (SE 1.8 months), median OS was not reached. At a median follow-up of 62.2 months, 24 out of the 81 patients (29.6%) had died. In the SACT dataset, at a median follow-up of 19.2 months, 73 out of the 215 patients (34%) had died. Median OS was also not reached and mean OS was not available. There is a notable difference between the proportion of patients that were still alive at the set time points. In KEYNOTE-087, cohort 2, out of the 24 patients who received SCT, 19 (79.2%) remained alive at the last follow-up and in the SACT cohort, of the 65 patients, 59 (91%) remained alive at the time of data cutoff. In KEYNOTE-087, cohort 2, 57 patients (70.4%) experienced an event and median PFS was 11.1 months (95% CI: 7.5 to 13.7). No subgroup analysis was reported for any characteristics as according to the Company the number of patients in the subgroups were too small.

The differences observed in the outcomes are largely attributed by the Company to the differences between baseline characteristics, as patients in KEYNOTE-087 were younger and fitter than those in the SACT dataset and to the differences in clinical practise (worldwide versus UK only). Regarding AEs, very high rates of AEs and drug-related AEs were reported in cohort 2 of KEYNOTE-087. No AEs evidence for the SACT dataset was presented in the CS. In addition, no AEs of comparators was presented in the clinical effectiveness Section. Therefore, a meaningful comparison and critique by the EAG was not feasible.

The comparators defined in the decision problem as well as the treatment basket are not aligned with the NICE final scope. Four additional sources were used in the ITC analyses, two retrospective observational studies: Cheah (2016)[3] and Eyre (2017)[5], KEYNOTE-204 and TA524[36]. Only OS was explored in the ITCs. The Company performed three types of analyses: naïve indirect comparison, MAICs and Bucher ITCs proposing ultimately seven different ITCs. The Company has acknowledged a series of key limitations in all the ITC analyses. Further key limitations identified by the EAG expose a high level of uncertainty in the presented OS HR of pembrolizumab versus SoC. The only publication which reports comparator data for treatments subsequent to BV is Cheah (2016)[3], the key limitation of this study is that 72% of the patients were not SCT-naïve. The Company suggested that the SACT dataset is a reliable source of information especially regarding generalisability to the UK setting. The EAG would agree that between KEYNOTE-087, cohort 2 and the SACT dataset, the latter is the more generalizable to the UK setting.

4. COST EFFECTIVENESS

4.1 EAG comment on Company's review of cost effectiveness evidence

Three SLRs were performed to identify and select relevant 1) cost effectiveness studies relevant to the R/R cHL population (CS, Appendix G); 2) relevant studies reporting utility values for the R/R cHL population in the UK setting (CS, Appendix H); and 3) studies reporting cost and healthcare resource use data for the R/R cHL population (CS, Appendix I).

4.1.1 Searches performed for cost effectiveness Section

The following paragraphs contain summaries and critiques of all searches related to cost effectiveness presented in the CS. The CADTH evidence-based checklist for the PRESS, was used to inform this critique.[15, 16] The EAG has presented only the major limitations of each search strategy in the report.

The Company provided a single search combining facets designed to retrieve economic evaluations, costs and resource utilisation outcomes, and HRQoL data for patients with R/R cHL. These searches were performed on 20 February 2023.

A summary of the sources searched is provided in Table 4.1

Table 4.1: Data sources searched for Appendix G: Published cost effectiveness studies (as reported in the CS)

Resource	Host/Source	Date Ranges	Date searched
Electronic databases			
Embase & MEDLINE	EMBASE.com	2017/07-2023/02/20	20/2/23
CS = company submission			

EAG comment:

- The CS reported that an update search was conducted on 23 February 2023 from 2017 onwards, using the search terms, methodology, and inclusion and exclusion criteria as per TA540. The search was transparent and reproducible, and appropriate study design filters were used. The Company reported that "Due to the lack of new emerging evidence in the small population of interest, hand-searching of additional publications from conferences and grey literature was unlikely to identify new evidence and therefore was not conducted".[4, 17]
- In the original 2017 submission the EAG asked the Company to clarify whether the reported MEDLINE/Embase strategy 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".[1] The EAG 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. Although simultaneous searching of Embase.com should automatically identify and search for equivalent MEDLINE medical subject heading (MeSH) terms, as the EAG did not have access to Embase.com for testing it was unclear if this was the case for all potentially useful MeSH terms. Given the potential limitations of this approach, the EAG considered it preferable to search each database separately, or at least to ensure inclusion of both Emtree and MeSH terms in the search strategy. The Company appears to have taken the same approach in the 2023

update, and as there were no other grey literature searches undertaken to ameliorate any loss of recall it is unclear if any relevant papers may have been missed.

- The EAG queried an unusual date limit in lines #76: (#22 AND #74 AND [01-07-2017]/sd NOT [13-01-2023]/sd) and #147 (#22 AND #145 AND [01-07-2017]/sd NOT [13-01-2023]/sd. Given that the searches were run on 20 February 2023, both lines would appear to discard results added to the database since 13 January 2023. At clarification the Company confirmed that this was a reporting error and provided screen shots of the correct strategy.

4.1.2 Inclusion/exclusion criteria

In- and exclusion criteria for the review on cost effectiveness studies, HRQoL studies and costs and resource use studies are presented in Table 4.2.

Table 4.2: Eligibility criteria for the SLRs

	Inclusion criteria	Exclusion criteria
Patient population	Adult (age ≥18 years) patients with R/R cHL, irrespective of age or gender	Patients under the age of 18 Disease other than R/R cHL
Intervention and 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)
Outcomes(s) 1 (Published economic evaluations)	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
Outcomes(s) 2 (HRQoL studies)	Health state utility data that are considered of interest to the review	Cost-evaluation studies comparing an impact of intervention on cost Cost and resource use studies
Outcomes(s) 3 (Cost/resource use studies)	Studies reporting the following outcomes: Total, direct, or indirect cost associated with disease Costs of absenteeism Costs of end-of-life care Volume of resource use	Cost-evaluation studies as such studies compare impact of intervention on cost Utility data only
Study design 1 (CEA studies)	CEA Cost utility analysis Cost benefit analysis Cost minimisation analysis Budget impact models	Other study designs: Epidemiology studies Clinical studies Pharmacokinetic/Pharmacodynamic (Animal/in-vitro) study

	Inclusion criteria	Exclusion criteria
	Cost consequence studies	General QoL studies
Study design 2 (HRQoL studies)	Studies that report utilities using one of the following instruments: EQ-5D questionnaire EORTC QLQ-C30 SF-36 HUI VAS Time trade off Standard gamble	Other study designs: Epidemiology studies Clinical studies Pharmacokinetic or pharmacodynamic (animal/in-vitro) study General QoL studies
Study design 3 (Cost/resource use studies)	Cost studies, surveys, or analysis Burden of illness studies Resource use studies	Other study designs: Epidemiology studies Clinical studies Pharmacokinetic or pharmacodynamic (animal/in-vitro) study General QoL studies
Based on Table 35 from Appendix G, Table 46 from Appendix H, Table 50 from Appendix I CEA = cost effectiveness analysis; EORTC QLQ-C30 = European Organization for Research and Treatment of Cancer Quality of Life Core 30; EQ-5D = European quality of life-5 dimensions; HUI = Health Utility Index; ICER = incremental cost-effectiveness ratio; LYG = life years gained; QALYs = quality-adjusted life years; QoL = quality of life; R/R cHL = relapsed or refractory classic Hodgkin lymphoma; SLRs = systematic literature reviews; VAS = visual analogue scale		

EAG comment: The EAG agrees that the eligibility criteria are suitable to fulfil the Company's objective to identify cost effectiveness studies.

The EAG questioned why despite that the inclusion and exclusion criteria expanded to include studies from US, Canada, Europe, Germany, or Denmark, in case of limited UK evidence as indicated in Table 50 in Appendix I, 1.1, the eight relevant studies that were identified from the US and France were excluded. In their response to clarification question B2.3 the Company explained that cost data from the US cannot be generalisable to the UK setting and therefore no cost data extracted were utilised in the economic model. Furthermore, they indicated that the data or resource use from the US studies were either too granular to be used in the economic model or were not specific to the line of treatment. Therefore, the Company relied on expert opinion for resource use in the UK context with regards to the population of interest. The one single-centre study in France was excluded as no resource use and only cost data specific to nivolumab treatment in the French hospital setting were reported, which were not considered to be generalisable for the UK setting. Furthermore, the EAG questioned why cost evaluation studies were excluded for cost and healthcare resource use review. The Company explained in their response to clarification question B2.3 that cost evaluation and costs and resource use were looked at in isolation, and in case they were included, no cost evaluation studies were identified in the 2017 search, with only one study identified in the 2023 literature search[41], which is only available as an abstract and no details on costs or resource could be retrieved. The rationales for excluding cost effectiveness studies after full paper reviewing are considered appropriate given the defined in- and exclusion criteria.

4.1.3 Conclusions of the cost effectiveness review

The CS Appendices G, H and I provides an overview of the included cost effectiveness, HRQoL and resource use and costs studies, but no specific conclusion was formulated.

EAG comment: The CS and response to clarification provided sufficient details for the EAG to appraise the single search of MEDLINE and Embase via Embase.com conducted to identify economic, HRQoL and cost data from the published literature on patients with R/R cHL.[9, 40] The CS reported that an update search was conducted on 23 February 2023 from 2017 onwards, using the search terms, methodology, and inclusion and exclusion criteria as per TA540. The search was transparent and reproducible, and appropriate study design filters were used. The EAG has some concerns about the limitations of searching MEDLINE and Embase together. The Company also reported that "Due to the lack of new emerging evidence in the small population of interest, hand-searching of additional publications from conferences and grey literature was unlikely to identify new evidence and therefore was not conducted". This approach may have resulted in some relevant papers being missed, but without the time to rerun and screen the searches, the EAG is unable to say what the overall impact on recall of results may have been.

4.2 Summary and critique of Company's submitted economic evaluation by the EAG

4.2.1 NICE reference case checklist

Table 4.3: NICE reference case checklist

Element of HTA	Reference case	EAG comment on CS
Perspective on outcomes	All direct health effects, whether for patients or, when relevant, carers	Consistent with NICE reference case
Perspective on costs	NHS and PSS	Consistent with NICE reference case
Type of economic evaluation	Cost utility analysis with fully incremental analysis	Consistent with NICE reference case
Time horizon	Long enough to reflect all important differences in costs or outcomes between the technologies being compared	Consistent with NICE reference case
Synthesis of evidence on health effects	Based on systematic review	Partly consistent with NICE reference case
Measuring and valuing health effects	Health effects should be expressed in QALYs. The EQ-5D is the preferred measure of HRQoL in adults	Consistent with NICE reference case
Source of data for measurement of HRQoL	Reported directly by patients and/or carers	Consistent with NICE reference case
Source of preference data for valuation of changes in HRQoL	Representative sample of the UK population	Consistent with NICE reference case
Equity considerations	An additional QALY has the same weight regardless of the other characteristics of the individuals receiving the health benefit	Consistent with NICE reference case

Element of HTA	Reference case	EAG comment on CS
Evidence on resource use and costs	Costs should relate to NHS and PSS resources and should be valued using the prices relevant to the NHS and PSS	Consistent with NICE reference case
Discounting	The same annual rate for both costs and health effects (currently 3.5%)	Consistent with NICE reference case

CS = company submission; EAG = Evidence Assessment Group; ED-5D = European quality of life-5 dimensions questionnaire; HTA = Health Technology Assessment; HRQoL = health-related quality of life; NHS = National Health Service; NICE = National Institute for Health and Care Excellence; PSS = Prescribed Specialised Services; QALY = quality-adjusted life year; UK = United Kingdom

4.2.2 Model structure

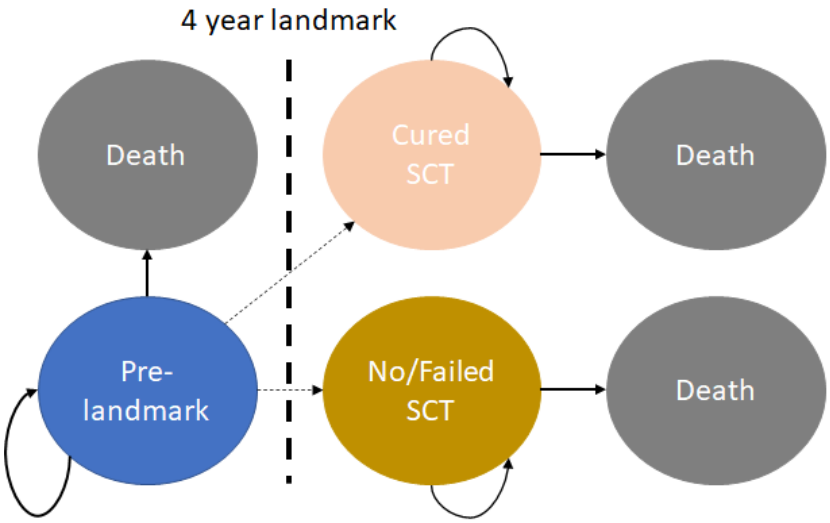
The Company deviated from the original TA540 model structure. In the new model structure, the Company sought to minimise complexity, model assumptions, number of health states and transitions while still capturing the most important outcomes.

The new model structure (Figure 4.1) entails a landmark point (at 48 months) after which patients are subdivided according to SCT status and consists of four health states:

- Alive pre-landmark
- Alive post-landmark; no or failed SCT
- Alive post-landmark; successful SCT
- Death

The Company assumed that pembrolizumab would improve the proportion of patients with (successful) SCT as well as improve both survival and HRQoL, both pre-landmark (all patients) and post-landmark (in case of no/failed SCT).

Figure 4.1: Model structure



Source: Based on Figure 5 of the clarification responses[9]
SCT = stem cell transplant

EAG comment: The main concerns of the EAG relate to: a) model structure that deviates from best modelling practices; and b) pembrolizumab OS and HRQoL benefits.

a) In the CS it is stated *"Allowing the OS curves to continue to four years has the major advantage that the SCT related events, their treatment effects and short term outcomes do not have to be estimated explicitly but rather are implicit parts of the OS curves". This statement indicates that the current model structure is potentially inconsistent with good modelling practices. Firstly, it is suboptimal from a transparency perspective as outcomes are not estimated explicitly. Secondly, it violates the homogeneity within health states assumption: "states need to be homogeneous with respect to both observed and unobserved (i.e., not known by the decision maker) characteristics that affect transition probabilities."*[42] In response to clarification question B4 the Company acknowledged this concerns but stated that some degree of within-state heterogeneity is almost always present (e.g. with the standard 3-state partitioned survival models) and that the Company does not expect this limitation to have a decision-important effect on the ICER. Although the EAG agrees that that some degree of within-state heterogeneity is almost always present, the current model structure, pre-landmark consists of an alive state including patients that had no, failed, and successful SCT. Given SCT is an important (if not the main) mechanism through which pembrolizumab affects outcomes as well as in the disease pathway, the EAG is concerned that the Company's deviation from best modelling practices (i.e., the within-state heterogeneity in the Company's model structure) might produce substantially biased results. This is particularly problematic given alternative model structures are available that do not increase complexity (i.e., adhering to the principles of parsimony), increase face validity as well as transparency and allow the explicit modelling of (time to) SCT. Additionally, the Company's response to clarification question B4 suggests that choices regarding the model structure were data-driven, as reported in the International Society for Pharmacoeconomics and Outcomes Research-Society for Medical Decision Making (ISPOR-SMDM) Modelling Good Research Practices Task Force paper *"Although data are essential to a model, the conceptual structure should be driven by the decision problem or research question and not determined by data availability"*, i.e. the model structure should not be driven by the availability of data.[43] Given the above, the EAG has serious concerns about the Company's model structure and therefore suggested an alternative model structure in clarification question B6d. This model structure consists of three health states:
    i. no/failed SCT (without distinction before/after an arbitrary landmark)
    ii. successful SCT
    iii. death

    This model structure is simple, transparent, adheres to best modelling practices and does not require tunnel states. Moreover, compared with the Company's model structure, it would not necessitate using an arbitrary landmark point and allow including the probability of transitioning to successful SCT every cycle. The Company indicated they have "have no objections to the suggested structure and we were initially interested in implementing something like this, which is why we requested the "time to SCT or death" data from SACT". However, the Company responded to B6 that it did not adopt this model structure for the following reasons; i) the Company believes it will make very little difference to the decision and; ii) it would require the estimation of time to SCT or death which the Company considered to be a difficult composite end-point for clinicians to help estimate. These justifications were not compelling to the EAG, particularly given the original TA540 FAD data collection recommendations were mainly focused on (time to) SCT as well as the importance of SCT in the disease pathway and mechanism through which pembrolizumab affects outcomes, it would

be very informative to estimate "time to SCT or death" and incorporate this explicitly in the economic model. Hence, the EAG believes an alternative model structure should be adopted to address the current decision problem.

b) In the original TA540 CS it is stated that, "it is expected that pembrolizumab monotherapy will be used as a "bridge" to alloSCT". Similarly, in the current CS it is stated that *"pembrolizumab is a bridge to SCT"* and *"Clinicians from the MSD UK Advisory Board stated among patients who received prior BV and are autoSCT ineligible, pembrolizumab would be the preferred treatment option to help achieve a better or durable response to bridge them to SCT"*. In addition, the recommendations for data collection (TA540 FAD) were mainly focused on (time to) SCT. Given the above, the main mechanism through which pembrolizumab affects patient outcomes is presumably through increasing the probability of (curative) SCT. However, the Company additionally assumed that pembrolizumab would also improve both OS and HRQoL, both pre-landmark (all patients) and post-landmark (in case of no/failed SCT). In clarification response B5a, the Company highlighted the OS HR of [redacted] from KEYNOTE-204 for patients who never had an SCT (also reported in CS, Table 33). Moreover, KEYNOTE-204 HRQoL data submitted in response to clarification question B19 indicated that there might be a HRQoL benefit of pembrolizumab (see also Section 4.2.8 of this report). The transferability (given the different population and comparator in KEYNOTE-204) as well as the duration of these benefits is nevertheless questionable. This supports the plausibility of pre-landmark OS and HRQoL benefits. However, the inclusion of post-landmark OS and HRQoL benefits is questionable, also given the 2-year pembrolizumab stopping rule. In response to clarification question B5g, the Company provide a scenario assuming no independent OS and HRQoL benefit post-landmark, this increase the ICER by [redacted]. This scenario highlights that the impact of assuming post-landmark OS and HRQoL benefits for pembrolizumab might be substantial. Given the 2-year pembrolizumab stopping rule and given the lack of evidence to support post-landmark OS and HRQoL benefits, the EAG preferred assuming no post-landmark OS and HRQoL benefits for pembrolizumab in its base-case (see also Sections 4.2.6 and 4.2.8).

4.2.3 Population

As per the NICE final scope, the patient group considered in the CS was patients with R/R cHL who have BV and cannot have autoSCT. This population is narrower than the marketing authorisation for pembrolizumab, which includes all patients with R/R cHL who have failed autoSCT or following at least two prior therapies when autoSCT is an unsuitable treatment option. Treatment with pembrolizumab for patients who have failed autoSCT or who cannot have autoSCT and who have not been treated with BV was already considered and recommended in TA772. Treatment with pembrolizumab was not recommended for cohort 1 in the previous appraisal TA540, i.e., patients with R/R cHL in adults who have had BV and autoSCT.

The modelled baseline patient characteristics were presented in Table 29 of the CS. These were based on the baseline characteristics of participants in cohort 2 from KEYNOTE-087 and the cohort forming the SACT dataset.

EAG comment: The population in the CS is as per the NICE final scope.

The EAG questioned why the baseline characteristics for the population model were derived from two different sources (SACT database and KEYNOTE-087, cohort 2) and whether this would cause issues with generalisability. The Company explained in their response to clarification question B4.9 that the SACT database was the preferred source of inputs for the economic model, where data are available, as it better reflected the R/R cHL population of interest in the UK real-world clinical setting and therefore

'Baseline age' and 'Proportion female' were derived from the SACT data. However, data on 'Weight' and 'Body surface area (BSA)' are not reported in the provided SACT data and instead were taken from the KEYNOTE-087, cohort 2 population. They further explained when all the baseline characteristics in the model were derived from KEYNOTE-087, it had a little impact on the ICER that was favourable to the intervention. The EAG considers this as resolved.

4.2.4 Interventions and comparators

The intervention considered in the CS was pembrolizumab. Consistent with the license, pembrolizumab was implemented in the economic model at a dose of 200 mg over 30 minutes Q3W to a maximum of 35 treatment cycles. A scenario analyses assessed the impact of pembrolizumab at a fixed dose of 400 mg Q6W as per the license.

The comparator considered in this review is SoC. The NICE final scope listed single or combination chemotherapy including drugs such as gemcitabine, vinblastine and cisplatin and BSC as comparators. In the TA540 submission the comparator composition was based on the Cheah (2016)[3] study. The listed comparators from Cheah were amended to reflect clinical practice at that time according to the 'guideline on the management of primary resistant and relapsed classical Hodgkin Lymphoma' as published in the British Journal of Haematology Board (2014) and from a previous NICE R/R cHL appraisal, TA462. The final comparator composition for TA540 comprised chemotherapy, bendamustine and investigational agents. In this review, the Company created a blended comparator based on Cheah (2016)[3], Eyre (2017)[5] and expert opinion to inform the comparator arm in the economic model, which consists of bendamustine, ICE, weekly chemotherapy (PMitCEBO), gemcitabine-based, oral chemotherapy, radiotherapy and mini-BEAM. The Company justified the selection of the blended comparator by stating that patients with R/R cHL who are not suitable for SCT after BV have few treatment options, and no new evidence evaluating the comparators specified by NICE on treatment post BV was identified through the SLR. The proportions of all treatments in the SoC arm were assumed to be equal as summarised in Table 4.3 and was varied in the sensitivity analysis. According to the CS, this assumption was followed as the Advisory Board could not give confident estimates for proportions of SoC regimens in fourth line. The Company did not consider BSC to be an active treatment and thus it was excluded as a comparator.

Table 4.4: Composition of SoC

Treatments	Proportion in the CS of TA540 based on Cheah (2016)[3]	Proportion in this CS	Source for the treatments in this CS
Bendamustine	18.5%	14.29%	Cheah (2016)[3]
Chemotherapy	38.5% (3.2% per regimen)	14.29%	Advisory Board
Investigational agents	43.1% 0.77	-	-
ICE	-	14.29%	Cheah et al. (2016), Eyre et al (2017) and MSD Advisory Board
Gemcitabine-based (IGEV, GEM-P, GDP, GVD)	-	14.29%	Cheah et al. (2016), Eyre et al (2017) and MSD Advisory Board

Oral chemotherapy (DECC)	-	14.29%	Cheah et al. (2016), Eyre et al (2017) and MSD Advisory Board
Radiotherapy	-	14.29%	Eyre (2017)[5]
Mini-BEAM	-	14.29%	Cheah et al. (2016), Eyre et al (2017) and MSD Advisory Board

Based on CS Section B.3.2.5 and Table 31[4]
CS = company submission; DECC = dexamethasone, etoposide, chlorambucil, lomustine; ICE = ifosfamide, carboplatin, etoposide; GDP = gemcitabine, dexamethasone, cisplatin, GEM-P = gemcitabine, cisplatin, methylprednisolone; GVD = gemcitabine, vinorelbine, doxorubicin; IGEV = ifosfamide, gemcitabine, vinorelbine; mini-BEAM = carmustine, etoposide, cytarabine, melphalan; TA = Technology Appraisal

EAG comment: The main concern of the EAG relates to the uncertainty about what should be included in SoC and the assumption of equal proportions in all the treatments that constitute the SoC.

In the NICE final scope, the comparator is single or combination chemotherapy including drugs such as gemcitabine, vinblastine and cisplatin, in addition to BSC. The EAG wishes first to highlight that BSC was not incorporated in the CS base-case as they considered the term BSC to refer to *"no active treatment"*. The Company also deviated from the NICE final scope by creating a blended comparator which is SoC that constitute various treatments that were informed by different sources (based on Cheah (2016)[3], Eyre (2017)[5] and expert opinion). This approach is different from the approach followed in TA540 where SoC was based on the composition and proportions of Cheah (2016) and amended in accordance with the British Journal of Haematology Board and TA462. Furthermore, the Company assumed equal proportions in all the of the treatments in the SoC arm. In their response to clarification question B4.8, the Company stated that the blended comparator reflects the interventions used in UK clinical practice and that adjusting the composition of the comparator will not affect health outcomes in the model but only SoC treatment costs. Furthermore, they explored assumptions about the composition of SoC within sensitivity analyses. The Company explained that the composition of TA540 was not considered as it is not reflective to current clinical practice for R/R cHL. They also added that the Advisory Board was not able to provide estimates for the proportions of patients that would receive each of the remaining regimens as there is no agreed SoC due to the availability of pembrolizumab in fourth line R/R cHL, patient heterogeneity, heterogeneity of chemotherapy options in use in difference centres and the rarity of the population in question. Therefore, the approach of equal proportions of all the treatments was assumed for simplicity. The Company explained further that they tested alternative scenarios where bendamustine was set to 100% in the SoC arm and a second scenario where the SoC treatment costs were halved as outlined in the CS. The EAG acknowledges that there is a lack of empirical evidence on what constitutes SoC and what would be the proportions for the included treatments in SoC. Therefore, it considers that there is a remaining uncertainty about the current composition and assuming equal proportions of treatments in the SoC arm in the CS. Therefore, the EAG used a different composition and percentage in its base-case. The proportion in the EAG base-case was as follows: bendamustine 23%, mini-BEAM 23%, gemcitabine 12%, radiotherapy 12%, chemotherapy, oral chemotherapy, and ICE 10% each. Radiotherapy and gemcitabine proportions were lowered as reported in the Eyre (2017)[5] study, while bendamustine and mini-BEAM proportions were increased. The proportion for the remaining treatments were adjusted to reach a sum of 100%.

4.2.5 Perspective, time horizon and discounting

The analysis in this review is performed, in accordance with the NICE reference case, from National Health Service (NHS) and Personal Social Services (PSS) perspective. Discount rates of 3.5% are applied to both costs and utilities. The model cycle length and lifetime horizon were consistent with TA540, which is a 1-week cycle with a lifetime horizon of 40 years.

EAG comment: Perspective, time horizon and discounting are as per the NICE final scope.

4.2.6 Treatment effectiveness and extrapolation

4.2.6.1 Overview

To inform the economic model, the Company estimated the following (treatment) effectiveness parameters using the following sources of evidence in the base-case:

- OS up to the landmark with pembrolizumab – SACT
- OS HR up to the landmark – KEYNOTE-204 for comparison with brentuximab vedotin combined with TA524 for indirect comparison (Bucher) with SoC
- Probability that a patient gets SCT on pembrolizumab – SACT
- Probability that a patient gets SCT on SoC – structured expert elicitation
- Ratio of autoSCT and alloSCT on pembrolizumab and SoC – SACT
- Probability that SCT is curative on pembrolizumab and SoC – structured expert elicitation
- OS after the landmark when cured with SCT after pembrolizumab and SoC – general population mortality
- OS after the landmark with no/failed SCT after pembrolizumab – SACT (subset of patients who never got an SCT)
- OS HR after the landmark with no/failed SCT after SoC – KEYNOTE-204 (subset of patients who never got an SCT)

4.2.6.2 Overall survival up to the landmark with pembrolizumab

Standard parametric survival models were fit to the KM data obtained from SACT. Data up to the 48-months landmark were available. No significant differences between the different survival curves were observed until the 4-year landmark. The Company selected the log-logistic curve because it had the property of declining hazards. It was also among the curves with the best statistical fit (CS, Table 32)[4]. The Company selected the SACT data over the KEYNOTE-087 data, given *"that the SACT dataset is much larger (N=215 vs N=81) and represents how pembrolizumab is used in the UK NHS (i.e. as a bridge to transplant where possible)"* (CS, Section B3.3.1.1)[4], but note that *"... outcomes in the Real World Data are worse than in the KEYNOTE-087 trial, likely due to higher age, worse patient fitness and presence of comorbidities"* (CS, Section B3.3.1.1)[4].

4.2.6.3 Overall survival hazard ratio up to the landmark

Overall survival on the SoC up to the 4-year landmark was calculated by applying a constant HR to the pembrolizumab arm. The Company considered different data sources, namely:

- Cheah (2016)[3]
- Eyre (2017)[5]
- KEYNOTE-204 pembrolizumab versus BV (unpublished)
- Markov trace from BV versus SoC used in TA524[36]

In the base-case the Company used the KEYNOTE-204 (in the third-line population) and TA524 (in the third-line CD30-positive population) data assuming that the treatment effect of pembrolizumab versus SoC can be composed of the sum of the treatment effects for pembrolizumab versus BV and BV versus SoC (Bucher method).

To estimate the treatment effect of BV versus SoC, the Company used a schematic detailing the health state membership over time in the Company's economic model from the committee papers of TA524. The Company digitized the schematic to obtain curves for the proportion alive in both model arms during the first 4 years (CS, Figure 14). Hazard ratio was derived by approximating the BV curve as closely as possible, by varying the HR until the new curve fit the original BV trace as closely as possible based on visual assessment. The Company considered *"that the proportional hazards assumption is not obviously violated within the first four years of model time"* (CS, B.3.3.1.3) and deemed the constant HR appropriate. The standard error was assumed large to include 1 in the CI.

The Company performed Bucher ITCs, calculating the HR for pembrolizumab versus SoC ([redacted]) by multiplying the HR ([redacted]) for pembrolizumab versus BV and the HR for BV versus SoC (0.62) together.

Different HRs were estimated using the sources presented in Table 4.5.

Table 4.5: Hazard ratios derived from different sources

Estimate number	Comparison	HR (CI)	Use in model by Company
1	Bucher ITC (KEYNOTE-204 and TA524) pembrolizumab versus BV versus SoC in TA524	[redacted]	Base-case
2	KEYNOTE-204 pembrolizumab versus BV	[redacted]	Scenario
3	Bucher ITC (SACT versus Eyre (2017)[5] and TA524)	0.41 (0.22 - 0.77)	Scenario
4	ITC SACT versus Eyre (2017)[5]	0.66 (0.44 - 0.98)	Scenario
5	ITC SACT versus Cheah (2016)[3]	0.59 (0.4 - 0.86)	Scenario
6	MAIC KEYNOTE-087 versus Eyre (2017)[5]	0.23 (0.12 - 0.42)	Scenario
7	MAIC KEYNOTE-087 versus Cheah (2016)[3]	0.24 (0.14 - 0.4)	Scenario

Based on CS Table 33[4]
BV = brentuximab vedotin; CI = confidence interval; CS = company submission; HR = hazard ratio; ITC = indirect treatment comparison; MAIC = matching-adjusted indirect comparison; SACT = systemic anti-cancer therapy; SoC – standard of care

4.2.6.4 Probability that a patient gets auto/alloSCT on pembrolizumab

The probability of patients treated with pembrolizumab receiving an SCT in the base-case was informed using the SACT data (30.2%). The proportions of people receiving auto versus alloSCT were also based on the SACT data (35.4% versus 64.6%) and applied to both the pembrolizumab and SoC arms. The Company note the discrepancy in proportions of auto versus alloSCT between SACT and KEYNOTE-087 where these probabilities were reversed. All probabilities are shown in Table 35 in the CS[4].

4.2.6.5 Structured expert elicitation: probability of SCT and probability that SCT is curative for SoC

The Company performed structured expert elicitation (SEE) to inform the following parameters:

- Proportion of people on SoC having a SCT
- Proportion of people on SoC who have an SCT that are cured
- Proportion of people on checkpoint inhibitors who have an SCT that are cured
- Proportion of people on SoC alive after 4 years (used for validation purposes)

The Company followed the recommended MRC protocol[44]. Six experts were invited, and five provided responses to the SEE questions (one had IT issues). The individual elicitation exercises were performed remotely, using a Microsoft Excel chip and bin template. A training slide deck was presented prior to the exercise. Pooled results were presented to all experts for discussion in a group meeting that also included the sixth invitee. The STEER R Shiny app was used for mathematical aggregation of experts' beliefs. The clinical validity and clarity of questions was assessed by one of the clinical experts prior to the meeting and the questions refined for clarity based on this feedback.

Proportion of people on standard of care having a SCT

The final distribution (with a mean of 8.17%) included responses from three of five experts, as the expert group concluded that this group of responses most closely aligned with an interpretation of the question that matched the patient population of interest. Distributions including all experts' opinions were used in scenario analysis.

Proportion of people on standard of care who have a SCT that are cured

The final distribution (with a mean of 37.18%) included all responses. This resulted in a total curative SCT proportion of 3.04% (based on model, approximately the 8.17% × 37.18% as per above) for SoC.

Proportion of people on checkpoint inhibitors who have an SCT that are cured

The final distribution (with a mean of 52.89%) included all responses. The experts highlighted that the improved cure rate was not solely due to improved responses with checkpoint inhibitor treatment. They noted that treatment with checkpoint inhibitors could chemo sensitise patients, increasing the chances that subsequent lines of chemotherapy would produce stronger responses, which would lead to higher cure rates. This resulted in a total curative SCT proportion of 15.99% (based on model, approximately 30.2% × 52.89% as per above) for pembrolizumab.

Proportion of patients on standard of care who are alive after 4 years

The final distribution (with a mean of 20.52%) included all responses.

4.2.6.6 Overall survival after the landmark when cured with SCT

Survival for the cured SCT group was assumed to be treatment independent and equal to the (age-matched) general population using the 2019-2020 national life tables for England (Office for National Statistics (ONS)).[45] At model start, the baseline age was 50.96. 60% of modelled patients were male. At 5 years, for the population of this age, the average survival would be at 98.27%. At 10 years, it would be 95.66%.

Clinical experts stated that average OS might be lower than in the general population but were not able to state by how much. The Company explored this in a scenario analysis where the standard mortality

ratio was set to a factor of 1.2 and the ICER increased to [redacted] per quality-adjusted life year (QALY) gained.

4.2.6.7 Overall survival after the landmark with no/failed SCT

For OS after pembrolizumab with no/failed SCT, in the absence of directly applicable data, the closest data available were KM curves on patients who had never received an SCT from the SACT data. It is unknown whether this is generalisable to the patient group modelled here, which includes both people who never received an SCT and those with a failed SCT. Standard parametric models were fitted. The Company selected the log-logistic in their base-case as it exhibited decreasing hazards. The Company noted that using the log-logistic, the OS of patients treated with pembrolizumab among the no/failed SCT group might be optimistic (CS, B.3.14.1). The exponential, which had the best statistical fit, was tested in scenario analysis.

For OS on SoC with no/failed SCT, the Company applied a HR to the SACT-derived transition probabilities. The rationale for why there would still be a difference in OS in the two groups ≥2 years after pembrolizumab was stopped was: *"Clinicians at the UK advisory board confirmed that there would be a treatment effect for some years after stopping pembrolizumab but were unable to say how long this would last"*(CS, B.3.3.1.6).

The HR of [redacted] (pembrolizumab versus BV) was obtained from third line patients in KEYNOTE-204 that never had an SCT within the study and was assumed applicable to pembrolizumab versus Soc for fourth line patients with no/failed SCT.

4.2.6.8 Treatment effect waning

No treatment effect waning was assumed in the Company base-case. In a scenario, treatment effect waning was applied only to the post-landmark OS in patient with no/failed SCT. In this scenario, hazards in the pembrolizumab arm were gradually waned to become equal to the calculated hazards in the SoC arm over time, equivalent to 3–5 years post cessation of pembrolizumab.

EAG comment: The main concerns of the EAG relate to: a) the new OS evidence is less favourable for pembrolizumab and uncertain; b) uncertain comparative effectiveness regarding OS pre-landmark; c) assumption of prolonged treatment effect beyond treatment discontinuation post-SCT; d) assumption of prolonged treatment effect beyond treatment discontinuation pre-SCT; e) using the log-logistic for modelling OS post-landmark in the no/failed SCT state; f) assumption of general population mortality in the cured post-SCT health state; and g) uncertain comparative effectiveness regarding probability of SCT.

a) The new OS evidence from SACT is less favourable for pembrolizumab compared with the KEYNOTE-087 evidence. The collected OS data from the SACT show a significantly lower survival rate in SACT versus KEYNOTE-087 at every timepoint (see Table 19, CS). The EAG notes that despite these less favourable data, the Company's new model results in a much more favourable ICER, driven by substantially increased incremental QALYs. The EAG was wondering what the impact of these data would be in the original Company model. In response to clarification question B1, the Company did not incorporate the new data in the original model as they deemed that model as not suitable, but the Company explained that the old model in fact under-estimated OS compared to both SACT data and KEYNOTE-087. The Company stated that the old and new models resulted in relatively similar mean life years gained (LYG) at 4 years (2.8 and 2.7 life years (LYs) respectively). The differences in ICERs ([redacted] in the

new model versus £36,950 for the 24-week and £55,628 for the 12-week original model according to the FAD[1], according to the Company, result from the following:

- A significant reduction in the cost of SCT
- A significant change in PAS discount for pembrolizumab
- In the old model, there were structural model assumptions that lead to patients spending a long time in the progressed disease state in combination with a utility estimate for progressed disease that may be unrealistically low
- A lack of a cured state in which patients do not experience relapse in the old model
- The use of the severity modifier in the new model

The EAG notes that the change in ICERs seems to be mostly driven by incremental QALYs instead of costs. The change in QALYs appears mainly to be attributable to utilities, since LYs appear similar. This would support the Company's reflection on the impact of the progressed disease state and its low utility value as well as the inclusion of a cured state. This is critiqued in Section 4.2.2. Whilst the collected OS data from SACT appear to have a relatively small impact on the ICER in the current model, this may be conditional on other assumptions and the model structure. The EAG notes that when changing the data source for pembrolizumab OS to KEYNOTE-087, 77% of people in the model are alive at the 4-year landmark in the pembrolizumab arm (compared to 52% in the base-case model), and 33% in the SoC arm (compared to 8% in the base-case), keeping all other assumptions constant. With this change, the incremental QALYs change by [redacted], reducing the Company's base-case ICER by approximately [redacted] per QALY gained, showing that there is some impact of the OS data on model results. This could potentially be amplified when taken together with other model assumptions. The uncertainty about the baseline OS together with other assumptions, for example, the relative treatment effect and included health states, are therefore a key issue. Related to this, there is remaining uncertainty about the choice of distribution for pre-landmark OS. The EAG tests the impact of using the exponential (which has the best statistical fit) in a scenario.

b) One of the main uncertainties in TA540 had been the comparative effectiveness. Ideally new comparative effectiveness evidence would be available in this review, but because in the CDF data were only collected on patients treated with pembrolizumab, this could not be provided. The EAG in the clarification questions suggested that historical SACT data be used to inform effectiveness of SoC as this, while potentially exhibiting historical bias, would overcome most of the limitations associated with the indirect comparisons used in the CS. The Company responded that they did not have access to any historical data from SACT, and that they were not aware that any had been published. As a result, the CS presents seven indirect comparisons, all of which are associated with significant limitations (Table 3 in clarification response B12[9]). Of the seven indirect comparisons, the EAG prefers the ITC of SACT versus Cheah (2016)[3]. The reasons for this are detailed in Section 3.4, including limitations surrounding the MAICs, and the lack of generalisability to patients in KEYNOTE-204 (third line) and TA524 (where SoC was only single-agent chemotherapy). The ITC of SACT versus Cheah (2016)[3] is adopted in the EAG base-case for the estimation of the pre-landmark OS HR. An additional option would be the use of the KEYNOTE-087 for baseline OS and the MAIC of KEYNOTE-087 versus Cheah (2016)[3], which is explored in an EAG scenario. In addition, there remains uncertainty about the appropriateness of the proportional hazards assumption. The Company provided diagnostics, i.e., log-cumulative hazard and Schoenfeld residual plots, upon request. According to the Company, Figure 3 in an addendum to question B11 of the clarification letter, which shows KEYNOTE-087, cohort 2 compared with BV data from Eyre (2017)[5], indicates

only little evidence for violation of the proportional hazard assumption before matching, and the observations were similar after matching. The caveat about the differences in the two populations applies. The EAG accepts the proportional hazards assumption.

c) The Company assumed a prolonged treatment effect beyond treatment discontinuation both in the estimation of pre- and post-landmark OS (in the no/failed SCT state post-landmark). A compelling justification for a treatment effect in the no/failed SCT post-landmark state was not provided, as the justification given was that *"Clinicians at the UK advisory board confirmed that there would be a treatment effect for some years after stopping pembrolizumab but were unable to say how long this would last"*(CS, B.3.3.1.6). The data used to inform the HR was not directly applicable to this population, as it was pembrolizumab versus BV in the third line population, presumably not collected after treatment cessation. Given that the prognosis of patients on SoC is unknown and that the Company's model underestimates OS in this population compared with expert opinion, the EAG considers the assumption of a modelled treatment effect post-landmark to be questionable and disables this in its base-case.

d) The EAG considers that pre-landmark OS treatment effect may also start to wane after treatment discontinuation. Especially given the lack of comparative evidence it is difficult to know when and how quickly this would occur. The EAG had recommended to explore the effect of this in scenarios, but in response to question B14, the Company stated that this would be contrary to established precedent in previous NICE appraisals of pembrolizumab, where hazards equalise from 3 years after treatment cessation to 5 years (i.e., years 5-7 in model time). The Company provided Schoenfeld residual plots and smoothed HRs which appeared to indicate treatment waning (as the HR approached 1 over time). However, the Company stated that this is not necessarily linked to treatment waning as about 30% of patients in all comparator populations used (studies from Cheah (2016)[3] and Eyre (2017)[5]) received SCT which may be potentially curative. The EAG notes that patients in the modelled control arm would also receive SCT, however, the Company states that the proportion would be far smaller than in the comparator studies and indeed their elicitation exercise resulted in only approximately 8% of patients in the comparator arm that would receive SCT. It therefore remains unclear whether there is treatment effect waning and how fast this occurs after treatment cessation.
   The EAG recommends that the impact of treatment effect waning in the pre-landmark model be explored in scenarios. The EAG also explored a scenario in which the treatment effect on OS was disabled in the pre-landmark model (during treatment and after treatment cessation), which is to be interpreted as an extreme scenario exploring the impact of assuming that pembrolizumab is solely a bridge to SCT with benefits in terms of HRQoL but not in reduction of mortality.

e) The justification for the choice of the log-logistic distribution in the post-landmark survival (no/failed SCT) state was not entirely convincing as the Company cited decreasing hazards as the reason for it being more plausible. However, the EAG considers it unclear whether this far into the time horizon and after treatment cessation, the hazards would indeed still be decreasing. The Company also changed their preference in response to clarification question B11 and gave preference for the exponential, to be in line with expert opinion on close to zero patients being alive at 10 years. The EAG agrees with this assessment and uses the exponential distribution in its base-case.

f) The Company's assumption of general population mortality applying to the cured SCT state is likely optimistic, as was supported by the experts in the Company's Advisory Board. In addition, it is unclear from the submission whether the Company used gender-matched life tables, accounting for a larger proportion of males (60%). Hence, the EAG explores a standard

mortality ratio of 1.5 in scenarios. Similar to the assumption that the mortality would be equal to that of the general population, this is also arbitrary, as no data could be identified.

g) The Company had to rely on SEE for the proportion of people on SoC having a SCT, the proportion of people on SoC who have an SCT that are cured and the proportion of people on checkpoint inhibitors who have an SCT that are cured. Whilst the SEE was performed in accordance with methodological guidance (the MRC protocol), the lack of data on patients in the SoC arm remains an important uncertainty. Probabilities of having an SCT on SoC and probabilities of cure in both arms were tested in the Company's scenarios and found to have a relatively minor impact. The EAG explores a scenario in which both probabilities of having SCT on SoC and probabilities of cure were set equal at the same time.

4.2.7 Adverse events

Adverse events were applied as a one-time cost and disutility in the first cycle of the model. The Company included the same AEs as in the original TA540[1]. AE rates were based on cohort 2 of KEYNOTE-087 and are presented in Table 38 of the CS. For SoC, the same AEs were assumed to be of interest. The Company updated the composition of SoC, which lead to changes in AE rates for SoC. The individual AE incidences were calculated by dividing the number of events per AE by the overall sample size per study to give a percentage. These percentages were then weighted by the proportion of patients receiving each regimen to calculate a total weighted AE incidence for the SoC arm for use in the model. The final AE rates for SoC are shown in Table 40 of the CS.

EAG comment: The main concerns of the EAG relate to the selection of comparators to estimate AE rates in the SoC arm, which results in much higher AE rates in the SoC arm than the pembrolizumab arm. There is general uncertainty about whether these AE rates are generalisable to the population of interest in this appraisal. It is particularly questionable whether weekly chemotherapy would be used as much as other treatments, given that it is associated with such a high incidence of neutropenia. The EAG explores the impact of SoC AE rates set equal to pembrolizumab AE rates in a scenario.

4.2.8 Health-related quality of life

The utility values were estimated for the model health states (see Section 4.2.2) as well as disutility values related to AEs and the SCT procedure.

42.8.1 Health-related quality of life data identified in the review

According to the CS, the SLR identified two studies reporting utility values for R/R cHL. Out of these, the Company stated that, originally for TA540, the study by Swinburn (2015)[46] was used to inform the utility decrement for PD patients. The Company stated that the other study (Ramsey (2016)[47] was not used given that "*the study did not provide utility data by responses status as per TA540 model structure*". No new published studies were identified in the SLR update (conducted in February 2023). The two studies identified in the original SLR where not used in the CS base-case, no justification was provided in CS Section B.3.4. why the utility values provided by Swinburn (2015)[46] and Ramsey (2016)[47] were not considered relevant for the updated model structure.

4.2.8.2 Health-related quality of life data identified in KEYNOTE-087 and KEYNOTE-204

In addition to studies identified in the SLR, relevant utility (EQ-5D-3L) data were available from KEYNOTE-087 (CS, Table 42) and KEYNOTE-204 (CS, Table 43). KEYNOTE-087 is a single arm, phase II, study of pembrolizumab in patients with R/R cHL. KEYNOTE-204, a randomised, phase III study, evaluated the clinical efficacy of pembrolizumab against BV in people with R/R cHL, and

comprised a mixture of those who had received prior autoSCT (37%) and those who were ineligible for autoSCT (63%).

For KEYNOTE-087, the cohort 2 patient-reported outcome (PRO) FAS population was analysed, i.e., patients who are SCT-naïve and have relapsed after treatment with, or failed to respond to, BV (cohort 2; fourth line) and that have at least 1 PRO assessment available and have received at least 1 dose of the study medication (PRO FAS). For KEYNOTE-204, the third line PRO FAS population was analysed, i.e., patients who are SCT-naïve (third line) and that have at least 1 PRO assessment available and have received at least 1 dose of the study medication (PRO FAS).

The Company decided not to use the KEYNOTE-087 utility data for the base-case analysis; rather the KEYNOTE-204 utility data were used. The following arguments were used for informing this decision:

- The Company had generalisability concerns related to KEYNOTE-087 as the OS is much longer than in the SACT dataset, suggesting a fitter patient group.
- KEYNOTE-204 provides utility data for pembrolizumab and a chemotherapy-based regimen (BV) and it is considered good practice to use the same data source for utility values across model arms.
- The sample size for whom EQ-5D data are available is larger for KEYNOTE-204 (N=134 in the third line cohort without prior SCT) than for KEYNOTE-087 (N=81 in cohort 2).
- The mean utility is very similar between KEYNOTE-087 and KEYNOTE-204 (0.834 versus 0.837 respectively), supporting the generalisability of KEYNOTE-204.

4.2.8.3 Health state utility values

The health state utility values, estimated based on KEYNOTE-204, reported in CS, Table 43 were used in the Company base-case. The estimated utility for third line BV was assumed to be applicable to fourth line SoC, similarly the estimated utility for third line pembrolizumab was assumed to be applicable to fourth line pembrolizumab. The Company justified the use of treatment specific utilities by stating that "*it was established during NICE TA772 and TA540 that pembrolizumab has a treatment effect on utility as well as disease progression and because the assumption of a persistent utility treatment effect was validated by clinicians at the UK clinical advisory board*". The pembrolizumab utility increment compared with SoC was 0.095 pre-landmark and 0.136 post-landmark for patients that had no successful SCT. In absence of health state utility values from KEYNOTE-204, for the "alive post-landmark; no or failed SCT" health state, the Company assumed progressed disease health state utility values for this health state. KEYNOTE-204 utility data were collected pre-dose at Cycle 1 (baseline), Cycle 3 (Week 6), Cycle 5 (Week 12), Cycle 7 (Week 18), and Cycle 9 (Week 24) and every 12 weeks thereafter until PD or up to 1 year while patients receive treatment, i.e., collected before the 48-month landmark point.

General population utility values were assumed for patients after successful SCT, i.e., assuming that the consequences of prior treatments and the health condition on HRQoL are fully reversible.

A summary of the health state utility values used in the CEA is provided in Table 4.6.

Table 4.6: Health state utility values based on KEYNOTE-204

Health state	Utility value (SE)		
Company base-case	**Pembrolizumab**	**SoC**	**Difference**
Alive pre-landmark	0.837 (0.012)	0.742 (0.016)	0.095
Alive post-landmark; no or failed SCT	0.807 (0.026)	0.671 (0.033)	0.136

Health state	Utility value (SE)		
Alive post-landmark; successful SCT	General population utility values (see CS, Table 44)		
EAG base-case (see EAG comment)	**Pembrolizumab**	**SoC**	**Difference**
Alive pre-landmark	0.816	0.730 (0.022)	0.085 (0.031)
Alive post-landmark; no or failed SCT	0.730 (0.022)	0.730 (0.022)	0.000
Alive post-landmark; successful SCT	0.770	0.770	0.000
Based on CS Table 43 and Section 4.6 of the "KEYNOTE-204 Utility Analysis" document[4] CS = company submission; EAG = Evidence Assessment Group; SCT = stem cell therapy; SE = standard error; SoC = standard of care			

4.2.8.4 Disutility values

The Company incorporated disutility values related to Grade ≥3 AE experienced whilst on the initial treatment as well as related to SCT.

Disutility values and duration for AE reported in CS, Tables 46 and 47 and were mainly retrieved from TA306 and TA476 (AE duration) as well as TA462 (AE disutility values). This approach was consistent with TA540. This resulted in a total AE related disutility of [redacted] for pembrolizumab and SoC respectively. The Company indicated that the treatment specific utilities, theoretically, already accounted for the AE related disutility and thus the AE related disutility was removed in a scenario analysis.

SCT related disutility was based on TA524 (BV for treating CD30-positive HL) and was dependent on time since SCT:

- 0.40 for 0-14 days after SCT;
- 0.22 for 14 days to 3 months after SCT;
- 0.05 for 3-24 months after SCT.

This resulted in a total QALY decrement of 0.150 (disutility value retrieved from the economic model). When multiplying this with the proportion of patients that was estimated to receive SCT (30.2% and 8.2%), this resulted in SCT related disutility of [redacted] for pembrolizumab and SoC respectively (implemented as a one-off utility in the first model cycle).

EAG comment: The main concerns of the EAG relate to: a) the use of KEYNOTE-204 instead of the pivotal trial (KEYNOTE-087) to inform health state utility values; b) estimation of health state utility values; c) maintaining pembrolizumab utility increment post-landmark; d) assuming general population utility values for patients that had a successful SCT; e) SCT related QALY decrement of 0.150 and; f) not using studies identified in SLR without justification.

a) The Company used KEYNOTE-204 to estimate the health state utilities. There are multiple limitations regarding the use of KEYNOTE-204, as it does not reflect the target population (third line, not fourth line), nor the appropriate comparator (BV, not Soc), assumptions are required to estimate the impact of SCT (progressed disease utility values are used for the "alive post-landmark; no or failed SCT" health state) and all utility data were collected before the 48-month landmark point. Additionally, KEYNOTE-087 is the pivotal study (according to CS, Section B2, clinical effectiveness). The Company "consider that KEYNOTE-087 and the SACT dataset represent the most robust evidence on the clinical effectiveness of pembrolizumab relevant to the decision problem" while KEYNOTE-204 is not mentioned as relevant study in the clinical effectiveness chapter of the CS (i.e., Sections B.2.1 to B.2.6).

Ideally the utility values are also retrieved from, what the Company considers "most robust evidence … relevant to the decision problem". However, the Company used KEYNOTE-204 utility data as this study provides a comparison of pembrolizumab versus BV and indicated that this can be considered an appropriate proxy for pembrolizumab versus SoC in the fourth line (clarification response B17). The EAG believes it is informative to provide the utility increment of pembrolizumab from KEYNOTE-204 but would have preferred that the Company would also provide detailed analyses of utility data retrieved from the pivotal trial (KEYNOTE-087).

b) Based on clarification responses B18 and B19, it became clear that the utility values reported in CS, Tables 42 and 43 were "simple naïve means", i.e., averaging the longitudinal utility data without considering the missing data (i.e., complete case analysis). Unfortunately, in response to clarification question B18b, the Company did not provide utility values estimated based on the pivotal trial (KEYNOTE-087) using a mixed effects model as requested. However, the Company did provide utility values estimated based on KEYNOTE-204 using a mixed effects model as requested in clarification question B19. This analysis (specifically Section 4.6 of the "KN204 Utility Analysis" document) indicated that the utility increment for pembrolizumab versus BV would be 0.08532. Potentially this increment could have been applied to KEYNOTE-087 utilities that were estimated using a mixed effects model as requested. However, given the Company did not provide the results of these analyses, the EAG adopted the KEYNOTE-204 BV utility of 0.73041 for SoC (from Section 4.6 of the "KN204 Utility Analysis" document) and 0.81573 for pembrolizumab (adding the abovementioned utility increment) in its base-case analysis, noting the limitations related to KEYNOTE-204 discussed above. For SoC, the EAG did not distinguish between pre- and post-landmark utilities given: i) the arbitrariness of this timepoint, ii) that all KEYNOTE-204 utility data were collected pre-landmark; iii) required assumptions to obtain post-landmark utility values, i.e., assuming progressed disease utilities post-landmark; and iv) the Company indicated (clarification response B6d) that it is reasonable to combine patients with no SCT and failed SCT.

c) The Company assumed that the pembrolizumab utility increment is maintained post-landmark (i.e., after 48 months) in the no/failed SCT health state. This assumption is questionable, given the 2-year pembrolizumab stopping rule and given that the data to inform this utility increment (KEYNOTE-204) were collected pre-landmark. No compelling evidence or arguments for including a post-landmark pembrolizumab utility increment were provided in clarification responses B5e or B20, hence the post-landmark pembrolizumab utility increment is removed in the EAG base-case. This is also consistent with TA772 where the committee *"concluded that the post-progression health-state utility values for pembrolizumab are uncertain but that it was unlikely that the health-state utility values estimated in KEYNOTE-204 would persist for the whole period of progression"*.

d) The Company assumed general population utility values for patients that had a successful SCT. This implicitly assumes the consequences of the treatments and health condition (fourth line R/R cHL) on HRQoL are fully reversed to a health state similar to that of the age-matched general population. The justification for this assumption provided by the Company in response to clarification question B24 is not very compelling to the EAG. The EAG noted that in TA524, a utility of 0.77 was adopted for patients being progression free after SCT (see also CS Table 48), for TA813 and TA451 this was 0.71 (≥3 months post SCT). Consistent with the Company's source for SCT disutility calculations, the EAG adopted the 0.77 from TA524 for patients being progression free after SCT. According to clarification response B25, the Company considers that the current decision problem "represents a sub-set of patients from TA524, R/R cHL patients post-BV and ineligible for SCT and therefore MSD considers the population to be transferable".

e) The Company, based on TA524, assumed a SCT related QALY decrement of 0.150. The EAG noted that in TA567, a QALY decrement of 0.300 related to SCT was adopted. This would likely double the abovementioned SCT related disutility of [redacted] for pembrolizumab and SoC respectively. To explore the impact of the SCT related QALY decrement, the TA567 SCT related QALY decrement was adopted by the EAG in a scenario analysis.
f) The Company did identify two potentially relevant HRQoL studies in the SLR and described why these were originally not used for TA540. However, no justification was provided why these HRQoL studies (Swinburn (2015)[46] and Ramsey (2016)[47]) were not used in the updated model structure.

4.2.9 Resources and costs

The cost categories included in the model were treatment acquisition costs, treatment administration, health state costs and resource use, subsequent therapy costs, autoSCT and alloSCT, costs of managing AEs and terminal costs.

Unit prices were based on the NHS reference prices[48], British National Formulary (BNF)[49], the electronic market information (eMIT)[50], UK Advisory Board, and the Nuffield Trust[51].

4.2.9.1 Resource use and costs data identified in the review

According to the CS, the SLR conducted originally for TA540 until 12 July 2017, was updated on 20 February 2023, and identified eight studies. None of these studies were UK-specific, therefore, no cost or resource use data from the SLR was used in the economic analysis.

4.2.9.2 Treatment acquisition costs

Pembrolizumab was given in accordance with the marketing authorisation, in 21-day treatment cycles for a maximum of 35 cycles, which included a fixed dose of 200 mg Q3W at a list price of £5,260 per cycle. The cost per cycle for each patient on treatment was inclusive to a Patient Access Scheme (PAS) discount and the administration cost. There were no specific NHS or payment-by-results (PbR) tariffs specific for costing pembrolizumab. Pembrolizumab administration cost was based on NHS reference cost code SB12Z[52] for the "administration of a simple chemotherapy", with an infusion lasting 30 minutes.

Due to the lack of relative dose intensity (RDI) data from SACT, the Company assumed 100% RDI and a hard stop at 35 cycles for pembrolizumab. A scenario analysis assessed the impact of using pembrolizumab in a fixed dose of 400 mg Q6W.

A blended SoC was created for this review for the comparator arm in the model, which was sourced from Cheah (2016)[3], Eyre (2017)[5] and the Clinical Advisory Board conducted by Merck Sharp and Dohme (MSD). The proportions of treatments were assumed to be equal. This approach departed from the CS in the original TA540, where SoC comprised of chemotherapy (38.46%), bendamustine (18.46%) and investigational agents (43.1%). The dosing and cycle details for each treatment in the SoC are summarised in CS Table 51. The model included a maximum of four vial/pack size for each SoC component, with the lowest cost combination of vials to make up the required dosage for the average patient. Vial sharing was applied and with 0% drug wastage assumed. Cost per unit was calculated by dividing the pack cost by units per pack. Costs for each SoC treatment are summarised in CS Table 52. Treatment costs per cycle for SoC treatments are summarised in CS Table 54.

Each treatment's cost in the SoC arm is applied in the model at the start of the treatment's cycle and up until it is maximum treatment duration. The maximum treatment cycles per each treatment were informed by TA540[4], Collins (2014)[13], Northern Cancer Alliance (2016)[53] and Lymphoma Group (2022)[54] as summarised in CS, Table 51. According to the CS, no data was found on the mean treatment duration for each SoC treatment, therefore, similarly to the approach in TA540, the total cost of SoC was down-weighted by using time on treatment (ToT) curve from the BV arm of KEYNOTE-204, which resulted in treatment costs of around ~90% of the maximum (when maximum number of treatment cycles were used). A sensitivity analysis examined a 70% reduction in the cost of SoC arm.

Radiotherapy was incorporated in the SoC composition as a one-time cost in the model (£5,340.59), following the clinical experts' advice. Due to the limited data, the costing approach of radiotherapy was based on published NICE Guidelines for Lung Cancer (NG122) in non-small cell lung cancer (NSCLC)[55] and NHS reference costs were used. The total radiotherapy cost was calculated using a weighted average of the multipliers (number of resource units) and cost per Healthcare Resource Group (HRG) radiotherapy components that were summarised in the CS Table 56. The number of fractions and admissions were reduced in accordance with the Royal College of Radiologists (RCR) guidelines[56], to ensure the cost reflects R/R cHL.

4.2.9.3 Administration costs

Pembrolizumab is administered as an intravenous infusion over 30 minutes. The administration cost is £286.71 which was sourced from 2021/22 NHS reference costs and was applied to every treatment cycle of pembrolizumab.

SoC administration costs per cycle were calculated by multiplying each component's cost per administration by the respective frequency in each cycle, as summarised in CS Table 53. The HRG codes SB14Z and SB15Z[52] were applied to delivering complex chemotherapy at first attendance and delivering subsequent elements of a chemotherapy cycle respectively as applied in TA540. All treatment administration and acquisition costs are summarised in Table 4.7.

Table 4.7: Treatment administration and acquisition costs per cycle

Treatment	Cycle length (days)	Maximum number of cycles	Administration cost per cycle	Acquisition cost per cycle
Pembrolizumab arm	21	35.0	£286.71	£5,260
SoC arm				
Bendamustine	28	6.0	£573.42	£69.00
ICE	14	3.0	£1,211.82	£1,379.91
IGEV	21	4.0	£1,580.26	£2,351.67
GEM-P	28	3.0	£1,211.82	£13.77
GDP	21	2.0	£843.38	£58.60
GVD	21	2.0	£843.38	£736.17
PMitCEBO	14	8.0	£843.38	£1,928.36
DECC	42	6.0	NA	£195.20
Mini-BEAM	28	3.0	£1,948.70	£9,694.81
Radiotherapy	N/A	N/A	NA	£5,340.59
Based on CS Tables 49, 53, 54[4]				

CS = company submission; DECC = dexamethasone, etoposide, chlorambucil, lomustine; GDP = gemcitabine, dexamethasone, cisplatin, GEM-P = gemcitabine, cisplatin, methylprednisolone; GVD = gemcitabine, vinorelbine, doxorubicin; ICE = ifosfamide, carboplatin, etoposide; IGEV = ifosfamide, gemcitabine, vinorelbine; mini-BEAM = carmustine, etoposide, cytarabine, melphalan; PMitCEBO = prednisolone, mitoxantrone, cyclophosphamide, etoposide, bleomycin and oncovin; SoC = standard of care

4.2.9.4 Health state costs

Health resource use estimates used in the model were obtained from TA524 as cited in TA540 and TA772 and validated by the Advisory Board. Costs were sourced from NHS reference costs. The clinical Advisory Board reviewed the high resource user (HRU) estimates from TA540 and agreed on omitting computed tomography (CT) scans and adjusting the frequency of positron emission tomography (PET) to every 3 to 4 months instead of 6 months. Resources for patients post 4-years who had not had or relapsed after SCT were costed yearly. No health resource costs were considered for patients cured with SCT, as related costs were assumed to be covered either by SCT costs or general pre-landmark costs. Costs and resource use pre-landmark and post-landmark for patients with no SCT or relapsed after SCT are summarised in Table 4.8.

Table 4.8: Costs and resource and cost use pre-landmarks for all patients and post-landmark for patients with no SCT or relapsed after SCT

Treatment	Healthcare resources							
	Outpatient attendance		Biochemistry		Cell blood count		PET scan	
	Weekly usage	Health state cost	Weekly usage	Health state cost	Weekly usage	Health state cost	Weekly usage	Health state cost
Pembrolizumab: pre-landmark	0.04	8.05	0.02	0.03	0.02	0.06	0.02	17.84
Pembrolizumab: post-landmark (no/relapsed SCT)	0.23	48.33	0.23	0.36	0.23	0.68	0.08	71.37
SoC: pre-landmark	0.04	8.05	0.02	0.03	0.02	0.06	0.02	17.84
SoC: post-landmark (no/relapsed SCT)	0.23	48.33	0.23	0.36	0.23	0.68	0.08	71.37

Based on CS economic model, health states costs Table[4]
CS = company submission; PET = positron emission tomography; SCT = stem cell transplant; SoC = standard of care

4.2.9.5 Subsequent therapy cost

According to the CS, and based on expert feedback from the Advisory Board, some patients receive good disease control for many years after pembrolizumab and may not need a subsequent treatment. Additionally, some patients would have died before subsequent treatment is considered. Therefore, subsequent treatment use was assumed to be lower in the pembrolizumab arm. The proportion of patients receiving subsequent treatments in KEYNOTE-204, which was 50.8% and 69.1% for pembrolizumab, and BV arms respectively, was used to represent subsequent treatment use in the pembrolizumab and SoC arms in the model, with the BV arm representing the SoC arm.

The subsequent treatment composition followed the approach of TA540 where the composition of subsequent treatments was based on the approach in TA306 and TA462, along with the SACT data. Bendamustine and radiotherapy were added to the composition in both pembrolizumab and SoC arm

while gemcitabine, ifosfamide, mesna, prednisolone, rituximab, vinorelbine (RVIG) was removed. Nivolumab was included only in the SoC arm as a subsequent treatment based on TA462. Costs and data on each regimen included for subsequent therapy are summarised in the CS, Tables 59-62. Subsequent treatment composition, distribution and administration costs are summarised in Table 4.9.

The acquisition and administration cost per cycle for each component of the subsequent treatment regimens were multiplied by the expected duration and usage to give a one-off weighted average cost. Subsequent treatment costs for both arms were calculated in the model by multiplying the weighted average subsequent treatment cost by the number of newly discontinued patients at each cycle.

Table 4.9: Subsequent treatment composition, distribution and administration costs

Therapies in the subsequent treatment arm	Treatment duration (cycles)*	Administration cost per cycle (£)	Source	Treatment distribution (%) for pembrolizumab	Treatment distribution (%) for SoC
Bendamustine	2.0	£573.42	BNF (2023)	35.56%	48.37%
Gemcitabine monotherapy	4.0	£860.13	eMIT (2023)	2.54%	3.34%
DHAP	2.0	£474.94	eMIT (2023)	2.54%	3.34%
CHOP	6.0	£474.94	eMIT (2023)	2.54%	3.34%
IVAC	3.5	£1,948.70	eMIT 2023)/ BNF (2023)	2.54%	3.34%
PMitCEBO	7.0	£843.38	eMIT 2023)/ BNF (2023)	2.54%	3.34%
Radiotherapy	1.0	£5,340.59	NHS reference costs	2.54%	3.34%
Nivolumab	36.5	£286.71	BNF (2023)	0.00%	0.07%
No active treatment	N/A	N/A	N/A	49.20%	31.52%
Weighted total subsequent treatment cost				£1,624.51	£2,230.43

Based on CS model, Subsequent Treatment Costs Table[4]

BNF = British National Formulary; CHOP = cyclophosphamide, doxorubicin, prednisolone, vincristine; CS = company submission; DHAP = dexamethasone, cytarabine, cisplatin; eMIT = electronic market information tool; IVAC = cytrabine, etoposide, ifosfamide, mesna; NHS = National Health Service; PMitCEBO = bleomycin, cyclophosphamide, etoposide, mitoxantrone, prednisolone, vincristine; SoC = standard of care

4.2.9.6 Autologous stem cell transplant and allogeneic stem cell transplant

In the original submission TA540, alloSCT cost was taken from the Radford study; a retrospective analysis and follow-up of 40 cHL patients who had relapsed after autoSCT, with 15 patients subsequently receiving alloSCT. The Radford study was considered not directly applicable to the population of interest in this review.

AutoSCT and alloSCT were micro costed and validated by clinicians following approaches taken in recent haematology appraisals (TA567[57] and TA813[58]) by splitting the process into stem cell harvesting; transplant procedure; and follow-up. Costs of stem cell harvesting, and transplant procedure were taken from the NHS reference costs. The alloSCT procedure cost was calculated using a weighted average of NHS reference costs (2021/22) for stem cell harvesting and transplant. A 24-month follow-up cost that

was sourced from the UK Stem Cell Strategy Oversight Committee report (2014)[59] and then inflated to 2021/2022 cost year using Personal and Social Services Research Unit (PSSRU) index. AutoSCT follow-up cost was calculated as a proportion of the follow-up costs for alloSCT using the relative costs from Blommestein (2021)[60]. The proportion of SCTs that were autoSCT or alloSCT (35%/65%) was taken from the SACT report. Costs of SCT were not discounted. According to the CS, ~90% of events occurred in the first year when inspecting the time to SCT curves from SACT, with the remainder occurring soon afterwards. Therefore, SCT costs were not discounted as it will only be reduced by tenth of the discount rate, or 0.35%. Table 66 in the CS summarises the total costs of autoSCT and alloSCT.

4.2.9.7 Adverse events costs

A one-time AE cost of £228.79 and £1,343.60 in the first cycle of the model was added to pembrolizumab and SoC arm respectively. The Company included the same AEs as in the original TA540. Adverse event costs were sourced from NHS reference costs using HRG codes from various NICE appraisals consistent with the TA540 appraisal. Adverse event unit costs were estimated by calculating the total cost per HRG code and dividing by the activity in that respective code. The cost per AE for each arm is reported in Table 4.10.

Table 4.10: Costs per AE for pembrolizumab and SoC arm

AE	Cost per AE	Weighted AE incidence for pembrolizumab	AE cost for pembrolizumab arm (£)	Weighted AE incidence for SoC	AE cost for SoC arm (£)
Thrombocytopenia	£993	0.04	36.75	0.12	117.81
Vomiting	£1,847	0.00	0.00	0.16	287.97
Anaemia	£941	0.04	34.83	0.08	70.97
Diarrhoea	£1,847	0.01	22.17	0.01	9.79
Dyspnoea	£863	0.00	0.00	0.11	95.07
Fatigue	£2,015	0.03	50.38	0.01	21.97
Leukopenia	£1,366	0.03	34.14	0.03	38.10
Nausea	£1,030	0.00	0.00	0.16	162.56
Neutropenia	£1,366	0.04	50.52	0.29	398.59
Pyrexia	£1,322	0.00	0.00	0.11	140.77

Based on CS Tables 38, 40, and 64[4]
AE = adverse event; CS = company submission; SoC = standard of care

4.2.9.8 Terminal care costs

A one-off terminal care cost of £8,752.32 was applied to all patients in the pembrolizumab or SoC arm who died prior to the landmark and who were in the no/failed SCT state after the landmark. A lower one-off terminal cost of £7,224.46 was applied to all patients in the curative SCT state after the landmark. Costs were sourced from the Nuffield Trust (2014)[51] research report and were inflated to the 2021/22 cost year using the PSSRU index[61].

EAG comment: The main concerns of the EAG relate to: a) assumptions about subsequent therapy proportions; b) the higher AE costs for the SoC arm; and c) SoC treatment duration.

a) The Company assumed a lower subsequent therapy proportion in the pembrolizumab arm based on feedback from experts in the Advisory Board, which was justified with patients having a

good disease management after pembrolizumab, and a lower number needing subsequent therapy. The proportion of patients who receive subsequent therapy in both arms was informed using the proportions in KEYNOTE-204, with the BV arm representing the SoC arm in this submission. Furthermore, the treatment composition in the subsequent therapy was based on TA306 along with data derived from SACT and clinical experts and not the composition from KEYNOTE-204. Nivolumab was also included only in the SoC arm as a subsequent treatment, as it was recommended in TA462. In response to clarification question B28, the Company justified the composition of subsequent treatments by stating that they followed the methodology of costing subsequent therapies as those that are available and in use in the UK NHS. The EAG questions if the proportions of subsequent therapies in KEYNOTE-204 apply to the different population in this submission and notes general uncertainty about the use of subsequent therapies in the population of interest. Therefore, the EAG explores a scenario in which costs of subsequent therapies are excluded.

b) Adverse event costs are higher in the SoC arm in comparison to pembrolizumab arm. This is a result of the higher AE rates that were assumed in the SoC arm compared to the pembrolizumab arm, especially in the estimation of rates for neutropenia, vomiting, and nausea, which had higher AE rates in the SoC arm in comparison to the incidence rate of 0 in pembrolizumab arm. The EAG questions whether these treatments with very high AE rates are used in practice to the extent assumed in this model and will explore a scenario on the impact of SoC AE rates when set equal to pembrolizumab AE rates in a scenario (see critique point in Section 4.2.4).

c) According to the CS, not all patients were expected to get the maximum treatment cycles in all SoC regimens. However, according to the Company, data on the mean treatment duration for each SoC treatment was not found, therefore, the total cost of SoC was down-weighted by using ToT curves from the BV arm of KEYNOTE-204, which resulted in treatment costs of around ~90% of its maximum costs, in an approach that is similar to TA540. The Company stated that this might be an overestimation of the SoC costs and thus, they examined a 70% reduction in the cost of SoC arm in a sensitivity analysis. The EAG agrees that this remains a point of uncertainty as the BV ToT curve in KEYNOTE-204 may be an overestimation of the actual ToT for SoC treatments and preferred if ToT was informed by other sources.

4.2.10 Severity

The Company used the online severity modifier tool (https://shiny.york.ac.uk/shortfall/) developed by Sheffield Centre for Health and Related Research (ScHARR) (University of Sheffield) to calculate the absolute shortfall. The Company assumed a baseline age of 51 years and that 40% of the target population would be female (CS, Table 67), resulting in expected QALYs of 15.6 for the (age and gender matched) general population. The estimated QALYs for people living with the condition (CS, Table 68), treated with SoC was [redacted] Resulting in absolute QALY shortfalls of [redacted]. Consequently, the absolute QALY shortfall was estimated to be between 12 and 18, resulting in a 1.2 QALY weight, which the Company applied in its base-case.

EAG comment: The EAG replicated the shortfall analysis reported in the CS, Section B.3.7. The reported absolute QALY shortfall (CS, Table 68) and the 1.2 QALY weight were successfully reproduced. Based on the abovementioned QALYs of 15.6 for the (age and gender matched) general population, the estimated QALYs for SoC should be >3.6 to result in an absolute QALY shortfall of less than 12, implying a 1.0 QALY weight.

4.2.11 Uncertainty

The main uncertainties listed in the CS were: small patient population, lack of RCTs and limited data on patient outcomes on SoC. The Company stated that this uncertainty is reduced by the collection of SACT data on patients treated with pembrolizumab.

EAG comment: The main uncertainty, i.e., lack of appropriate comparator data (which stems from the lack of RCTs) has not been addressed. The resulting uncertainty manifests in a choice of potential comparator data used in the model which suffers from indirectness (differences in populations, e.g., third line versus fourth line or the rates of SCT), the methods for analysing these data (unanchored comparisons), and the identification of treatments with their dosing regimens, costs, and AEs. Similarly important, there is now significant uncertainty surrounding the model structure, which may introduce bias in the comparison of pembrolizumab versus SoC.

5. COST EFFECTIVENESS RESULTS

5.1 Company's cost effectiveness results

The CS base-case cost effectiveness results as summarised in Table 5.1 indicated that pembrolizumab is both more effective (QALYs of [redacted]) and more costly (cost of [redacted]) when compared to current care amounting to an ICER of [redacted] per QALY gained. Pembrolizumab improved the OS on treatment, increased the number of patients spending time in the "after landmark" and "up to landmark" health state, increased the number of patients eligible for curative SCT, and increased the probability of having a curative SCT. All incremental QALY gains were multiplied by the 1.2 severity modifier. Life years were not discounted. The base-case disaggregated results in QALY gains and costs are summarised in Table 5.2 and Table 5.3.

Table 5.1: Base-case results - aggregated

Intervention	Total costs (£)	Total LYG	Total QALYs	Inc costs (£)	Inc LYG	Inc QALYs (including severity modifier)	ICER (£/QALY)
SoC	[redacted]	[redacted]	[redacted]				
Pembrolizumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Based on CS Table 70[4]
CS = company submission; ICER = incremental cost-effectiveness ratio; LYG = life years gained; QALYs = quality-adjusted life years; SoC = standard of care

Table 5.2: Base-case results in QALY gain (by age disutility, health state, SCT and AE disutility (discounted) (with severity modifier 1.2) - disaggregated

Treatment	Age disutility	QALYs by health state				SCT disutility	AE disutility	Absolute QALYs
		Up to landmark	After landmark					
			No curative SCT	Curative SCT	Overall QALYs			
Pembrolizumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
SoC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
***Inc QALYs**	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
% absolute increment*	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Based on CS Appendix J, Table 58[17]
* Severity modifier of 1.2 applied, the % absolute increment has been adjusted to include the absolute QALYs with the 1.2 severity modifier
AE = adverse event; CS = company submission; N/A = not applicable; QALY = quality-adjusted life-year; SCT = stem cell transplant; SoC = standard of care

Table 5.3: Base-case results in costs - disaggregated

Treatment	Drug cost	Health state	Terminal care	AE costs	Subsequent Treatment	Total cost
Pembrolizumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
SoC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Inc costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
% absolute increment	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Source: CS Appendix J, Table 60[17]
AE = adverse event; CS = company submission; SoC = standard of care

Overall, the technology is modelled to affect QALYs by:

- Increased QALYs for pembrolizumab, by impacting HRQoL pre-landmark, when cured after SCT and when not cured after SCT; and AE and SCT disutilities (QALY increased by [redacted] compared with SoC without the severity modifier).
- Increased LYG for pembrolizumab up to the landmark but especially after the landmark (LYG increased by [redacted] compared with SoC overall).

Overall, the technology is modelled to affect costs by:

- Increased drug costs (additional costs of [redacted] compared with SoC).
- The higher health state costs (additional costs of [redacted] compared with SoC).
- Cost-savings in terminal care, AE and subsequent treatment costs (cost-saving of [redacted] compared with SoC).

EAG comment: The main concern of the EAG relates to the significant difference between treatment arms in clinical outcomes in the Company's disaggregated results. In the Company's model, the treatment effect is reflected in different outcomes including the OS on treatment, number of patients who are eligible for SCT, the probability of having a curative SCT, and OS after the treatment has stopped even with failed SCT or SCT not received. The EAG questions whether there may be double-counting of treatment effect as alluded to in other Sections and wonders whether these results are truly reflective of the treatment effect of pembrolizumab on patients.

5.2 Company's sensitivity analyses

The Company performed and presented the results of probabilistic sensitivity analyses (PSA), deterministic sensitivity analyses (DSA) as well as scenario analyses.

A PSA was run with 1,000 iterations, varying the uncertain parameters. According to the CS, the ICER and all disaggregated results were found to be stable with pembrolizumab having a greater than [redacted] probability of being cost effective versus a threshold of £30,000/QALY gained (including a 1.2 severity modifier to incremental QALYs).

The Company stated that DSAs were conducted for varying key inputs. Parameters were varied within their 5% and 95% CIs where possible, in the case of certain parameters with no associated probability distribution, to arbitrarily decided values e.g., the location of the landmark was varied to 3 and 5 years. According to the CS, the model outcomes were robust to all these changes, with all ICERs below

[redacted]. The following parameters were identified as most influential on the cost effectiveness of pembrolizumab versus SoC:

- OS HR up to landmark: SoC versus pembrolizumab
- Applying treatment waning (years 5 to 7)
- Probability that SCT will be curative: pembrolizumab

Based on the Company's scenario analyses, modelling assumptions that have the greatest effect on the ICER were:

- All SoC and subsequent treatment costs were removed
- Exponential survival curve after landmark
- Treatment waning effect on no/failed SCT OS

EAG comment: No comments.

5.3 Model validation and face validity check

5.3.1 Face validity assessment

The Company provided figures showing the health state membership over time that would allow a face validity check. However, the Company was not able to clinically validate these. No further details of a systematic face validity assessment (including on the model structure) were provided in CS Section B.3.14.

5.3.2 Technical verification

No details of any technical verification were provided in CS Section B.3.14.

5.3.3 Comparisons with other Technology Appraisals

Cross validity assessments were performed for certain parameter inputs and assumptions detailed in CS, Table 30. Comparisons of aggregate costs, LYs and QALYs with previous TAs were provided in response to clarification question B32.

5.3.4 Comparison with external data used to develop the economic model

The modelled OS estimates were compared with those from the data used to develop the model. The Company noted that the 4-year OS estimated for SoC (~8%) in the Company base-case (i.e. using the Bucher ITC with data derived from KEYNOTE-204 and TA524) was lower than had been estimated by advisors at the UK Advisory Board (~20%), but that it was similar to the SoC survival that had been estimated and accepted as plausible for third line SoC during TA524 (~10%). For pembrolizumab, the Company estimated 4-year OS of ~52%. The Company also extrapolated the SACT OS data fully, i.e., not stopping at the 4-year landmark (CS, Figure 27), to compare the estimates derived with the model to these extrapolations. This showed that the model's predictions about long term survival were within the wide range of predictions produced by standard extrapolation methods.

5.3.5 Comparison with external data not used to develop the economic model

The modelled OS estimates for people treated with SoC were compared with estimates from alternative data sources used in scenario analyses and concluded that modelled OS estimates were approximately in line with SoC estimates for TA524 in the third line setting (modelled 8% versus 10% in TA524).

EAG comment: The main concerns of the EAG relate to: a) technical verification; b) limited face validity assessment of the model structure; and c) external validity checks.

a) No technical verification was reported in the CS. In response to clarification questions, the Company provided their internal technical verification checklist. This appears to be appropriate and no specific doubts remain about the internal validity of the model.
b) No or limited (i.e., implicit) face validity assessment was performed of the model structure. The EAG remains in doubt over whether the present model structure exhibits face validity.
c) External validity checks were performed. Their results do cast doubt over whether the modelling of SoC is appropriate as the modelled OS at 4 years is below that expected by experts for this population. The EAG did not find the similarity to TA524 particularly reassuring as SoC in that third line setting is not comparable to SoC in this fourth line population. The Company also mentioned a comparison of OS between patients who failed BV and did not get an SCT in the Eyre (2017)[5] study at 20 months (25%) and the model (28%), which appears more in line. Overall, there remains uncertainty about the modelled OS in the SoC arm.

6. EVIDENCE ASSESSMENT GROUP'S ADDITIONAL ANALYSES

6.1 Exploratory and sensitivity analyses undertaken by the EAG

Table 6.1 summarises the key issues related to the cost effectiveness categorised according to the sources of uncertainty as defined by Grimm (2020)[62]:

- Transparency (e.g., lack of clarity in presentation, description, or justification).
- Methods (e.g., violation of best research practices, existing guidelines, or the reference case).
- Imprecision (e.g., particularly wide CIs, small sample sizes, or immaturity of data).
- Bias & indirectness (e.g., there is a mismatch between the decision problem and evidence used to inform it in terms of population, intervention/comparator and/or outcomes considered).
- Unavailability (e.g., lack of data or insight).

Identifying the source of uncertainty can help determine what course of action can be taken (i.e., whether additional clarifications, evidence and/or analyses might help to resolve the key issue). Moreover, Table 6.1 lists suggested alternative approaches, expected effects on the cost effectiveness, whether it is reflected in the EAG base-case as well as additional evidence or analyses that might help to resolve the key issues.

Based on all considerations in the preceding Sections of this EAG report, the EAG 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 EAG form the EAG base-case and were subdivided into three categories (derived from Kaltenthaler (2016))[63]:

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

6.1.1 EAG base-case

Adjustments made by the EAG, to derive the EAG base-case (using the CS base-case as a starting point) are listed below. Table 6.2 shows how individual adjustments impact the results plus the combined effect of all abovementioned adjustments simultaneously, resulting in the EAG base-case. The 'fixing error' adjustments were combined and the other EAG analyses were performed also incorporating these 'fixing error' adjustments given the EAG considered that the 'fixing error' adjustments corrected unequivocally wrong issues.

6.1.1.1 Fixing errors

One error was identified: for ITC 6, the values before matching were used in the model, when it should be the values after matching. This does not impact the Company or EAG base-case.

6.1.1.2 Fixing violations

No violations were identified.

6.1.1.3 Matters of judgement

1. Amended comparator proportions for cost calculations (Section 4.2.4)
2. Exponential in no/failed SCT OS (instead of log-logistic) (Section 4.2.6)

3. Set no/failed SCT HR to 1 (Section 4.2.6)
4. ITC 5 (SACT versus Cheah (2016)[3]) for pre-landmark OS HR (Section 4.2.6)
5. Pre-landmark utilities from KEYNOTE-204 utility analysis (Section 4.2.8)
6. Set post-landmark HRQoL (pembrolizumab and SoC) equal to SoC pre-landmark (Section 4.2.8)
7. Utility of 0.77 for successful SCT (Section 4.2.8)

6.1.2 EAG exploratory scenario analyses

The EAG performed the following exploratory scenario analyses to explore the impact of alternative assumptions conditional on the EAG base-case.

6.1.2.1 Exploratory scenario analyses

1. KEYNOTE-087 and naïve ITC versus Cheah (2016)[3] (Section 4.2.6)
2. Set subsequent treatment costs to 0 (Section 4.2.9)
3. Set SoC arm prob of SCT and being cured by SCT equal to pembrolizumab arm (Section 4.2.6)
4. Apply SMR of 1.5 to cured SCT state (Section 4.2.6)
5. QALY decrement of 0.3 for SCT (Section 4.2.8)
6. Set equal AE rates on SoC to pembrolizumab (Section 4.2.7)
7. Exponential for pre-landmark OS (Section 4.2.6)
8. Landmark of 2 years (Section 4.2.2)
9. No OS gain pre-landmark (Section 4.2.6)
10. All EAG changes except scenarios 1 and 9
11. Pessimistic scenario, all EAG changes except EAG scenario 1

6.1.3 EAG subgroup analyses

No subgroup analyses were performed by the EAG.

Table 6.1: Overview of key issues related to the cost effectiveness

Key issue	Section	Source of uncertainty	Alternative approaches	Expected impact on ICER[a]	Resolved in EAG base-case[b]	Required additional evidence or analyses
5) The Company model structure is inconsistent with good modelling practices	4.2.2	Methods	Alternative model structure	+/-	No	Alternative model structure
6) The composition and proportions of the SoC in the comparator arm	4.2.4	Unavailability	Scenarios	+/-	Explored	Real world evidence on treatment use
7) Uncertain comparative effectiveness	4.2.6	Indirectness	Amendments to model assumptions/scenarios	+	Explored	Historical SACT data
8) Uncertain duration of treatment effect	4.2.6	Indirectness	Amendments to model assumption/scenarios	+	Explored	Historical SACT data
9) KEYNOTE-087 utilities estimated through a mixed effects model	4.2.8	Methods	Provide analysis	+/-	No	Provide analysis
10) Uncertainty about subsequent therapies	4.2.9	Indirectness	Scenarios	+/-	Explored	Real world evidence on treatment use

[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 EAG and '+' indicates that the EAG believes this issue likely induces bias in favour of the intervention versus at least one comparator
[b] Explored
EAG = Evidence Assessment Group; ICER = incremental cost-effectiveness ratio; SACT = systemic anti-cancer therapy; SoC – standard of care

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

In Section 6.1 the EAG base-case was presented, which was based on various changes compared to the Company base-case. Tables 6.2 and 6.3 show how individual changes impact the results plus the combined effect of all changes simultaneously. The exploratory scenario analyses are presented in Tables 6.4 and 6.5. These are all conditional on the EAG base-case. The analyses numbers in Tables 6.2 to 6.5 correspond to the numbers reported in Section 6.1. The submitted model file contains technical details on the analyses performed by the EAG (e.g., the "ERG" sheet provides an overview of the cells that were altered for each adjustment).

Table 6.2: Deterministic EAG base-case, no severity modifier

Technologies	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER (£/QALY)
CS base-case					
Pembrolizumab	[redacted]	[redacted]			
Standard of care	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
EAG1 Comparator proportions amended					
Pembrolizumab	[redacted]	[redacted]			
Standard of care	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
EAG2 Exponential in no/failed SCT OS (instead of log-logistic)					
Pembrolizumab	[redacted]	[redacted]			
Standard of care	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
EAG3 Set no/failed SCT OS HR to 1					
Pembrolizumab	[redacted]	[redacted]			
Standard of care	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
EAG4 ITC 5 for pre-landmark OS HR					
Pembrolizumab	[redacted]	[redacted]			
Standard of care	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
EAG5 Pre-landmark utilities from KEYNOTE-204 utility analysis: SoC 0.73041 and pembrolizumab 0.81573					
Pembrolizumab	[redacted]	[redacted]			
Standard of care	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
EAG6 Set post-landmark HRQoL (pembrolizumab and SoC) equal to SoC pre-landmark					
Pembrolizumab	[redacted]	[redacted]			
Standard of care	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
EAG7 Utility of 0.77 for successful SCT					
Pembrolizumab	[redacted]	[redacted]			
Standard of care	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
EAG base-case					
Pembrolizumab	[redacted]	[redacted]			
Standard of care	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
EAG base-case (probabilistic)					
Pembrolizumab	[redacted]	[redacted]			
Standard of care	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Results deterministic unless indicated.					

Technologies	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER (£/QALY)
CS = company submission; EAG = Evidence Assessment Group; HR = hazard ratio; HRQoL = health-related Quality of Life; ICER = incremental cost-effectiveness ratio; ITC = indirect treatment comparison; LYG = life years gained; OS = overall survival; QALYs = quality-adjusted life years; SCT = stem cell transplant; SoC = standard of care					

Table 6.3: Deterministic EAG base-case, with severity modifier

Technologies	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER (£/QALY)
CS base-case with severity modifier					
Pembrolizumab	[redacted]	[redacted]			
Standard of care	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
EAG1 Comparator proportions amended					
Pembrolizumab	[redacted]	[redacted]			
Standard of care	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
EAG2 Exponential in no/failed SCT OS (instead of log-logistic)					
Pembrolizumab	[redacted]	[redacted]			
Standard of care	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
EAG3 Set no/failed SCT OS HR to 1					
Pembrolizumab	[redacted]	[redacted]			
Standard of care	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
EAG4 ITC 5 for pre-landmark OS HR					
Pembrolizumab	[redacted]	[redacted]			
Standard of care	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
EAG5 Pre-landmark utilities from KEYNOTE-204 utility analysis: SoC 0.73041 and pembrolizumab 0.81573					
Pembrolizumab	[redacted]	[redacted]			
Standard of care	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
EAG6 Set post-landmark HRQoL (pembrolizumab and SoC) equal to SoC pre-landmark					
Pembrolizumab	[redacted]	[redacted]			
Standard of care	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
EAG7 Utility of 0.77 for successful SCT					
Pembrolizumab	[redacted]	[redacted]			
Standard of care	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
EAG base-case					
Pembrolizumab	[redacted]	[redacted]			
Standard of care	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
EAG base-case (probabilistic)					
Pembrolizumab	[redacted]	[redacted]			
Standard of care	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Results deterministic unless indicated. CS = company submission; EAG = Evidence Assessment Group; HR = hazard ratio; HRQoL = health-related quality of life; ICER = incremental cost-effectiveness ratio; ITC = indirect treatment comparison; LYG = life years gained; OS = overall survival; QALYs = quality-adjusted life years; SCT = stem cell transplant; SoC = standard of care					

Table 6.4: EAG scenarios without severity modifier

Technologies	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER (£/QALY)
EAG base-case					
Pembrolizumab	[redacted]	[redacted]			
Standard of care	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
EAG scenario 1 ITC 7 KEYNOTE-087 versus Cheah (2016)[3] for pre-landmark OS HR					
Pembrolizumab	[redacted]	[redacted]			
Standard of care	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
EAG scenario 2 Set subsequent treatment costs to 0					
Pembrolizumab	[redacted]	[redacted]			
Standard of care	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
EAG scenario 3 Set equal in both arms prob of SCT and being cured by SCT					
Pembrolizumab	[redacted]	[redacted]			
Standard of care	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
EAG scenario 4 Apply SMR of 1.5 to cured SCT state					
Pembrolizumab	[redacted]	[redacted]			
Standard of care	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
EAG scenario 5 QALY decrement of 0.3 for SCT					
Pembrolizumab	[redacted]	[redacted]			
Standard of care	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
EAG scenario 6 Equal AE rates on SoC					
Pembrolizumab	[redacted]	[redacted]			
Standard of care	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
EAG scenario 7 Exponential for pre-landmark OS					
Pembrolizumab	[redacted]	[redacted]			
Standard of care	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
EAG scenario 8 Landmark of 2 years					
Pembrolizumab	[redacted]	[redacted]			
Standard of care	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
EAG scenario 9 No OS gain pre-landmark					
Pembrolizumab	[redacted]	[redacted]			
Standard of care	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
EAG scenario 10 All EAG changes except EAG scenarios 1 and 9					
Pembrolizumab	[redacted]	[redacted]			
Standard of care	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
EAG scenario 11 Pessimistic scenario, all changes except EAG scenario 1					
Pembrolizumab	[redacted]	[redacted]			
Standard of care	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Results deterministic unless indicated.
AE = adverse event; EAG = Evidence Assessment Group; HR = hazard ratio; ICER = incremental cost-effectiveness ratio; ITC = indirect treatment comparison; OS = overall survival; QALYs = quality-adjusted life years; SCT = stem cell transplant; SoC = standard of care

Table 6.5: EAG scenarios with severity modifier

Technologies	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER (£/QALY)
EAG base-case					
Pembrolizumab	[redacted]	[redacted]			
Standard of care	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
EAG scenario 1 ITC 7 KEYNOTE-087 versus Cheah (2016)[3] for pre-landmark OS HR					
Pembrolizumab	[redacted]	[redacted]			
Standard of care	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
EAG scenario 2 Set subsequent treatment costs to 0					
Pembrolizumab	[redacted]	[redacted]			
Standard of care	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
EAG scenario 3 Set equal in both arms prob of SCT and being cured by SCT					
Pembrolizumab	[redacted]	[redacted]			
Standard of care	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
EAG scenario 4 Apply SMR of 1.5 to cured SCT state					
Pembrolizumab	[redacted]	[redacted]			
Standard of care	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
EAG scenario 5 QALY decrement of 0.3 for SCT					
Pembrolizumab	[redacted]	[redacted]			
Standard of care	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
EAG scenario 6 Equal AE rates on SoC					
Pembrolizumab	[redacted]	[redacted]			
Standard of care	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
EAG scenario 7 Exponential for pre-landmark OS					
Pembrolizumab	[redacted]	[redacted]			
Standard of care	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
EAG scenario 8 Landmark of 2 years					
Pembrolizumab	[redacted]	[redacted]			
Standard of care	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
EAG scenario 9 No OS gain pre-landmark					
Pembrolizumab	[redacted]	[redacted]			
Standard of care	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
EAG scenario 10 All EAG changes except EAG scenarios 1 and 9					
Pembrolizumab	[redacted]	[redacted]			
Standard of care	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
EAG scenario 11 Pessimistic scenario, all changes except EAG scenario 1					
Pembrolizumab	[redacted]	[redacted]			
Standard of care	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Results deterministic unless indicated.
* No severity modifier applied because SoC QALY gains are above the threshold
AE = adverse event; EAG = Evidence Assessment Group; HR = hazard ratio; ICER = incremental cost-effectiveness ratio; ITC = indirect treatment comparison; OS = overall survival; QALYs = quality-adjusted life years; SCT = stem cell transplant; SoC = standard of care

6.3 EAG's preferred assumptions

The estimated EAG base-case ICER (probabilistic), based on some of the EAG preferred assumptions highlighted in Section 6.1, was without and with severity modifier [redacted] or [redacted] per QALY gained. It is important to note that this ICER is not fully reflective of EAG preferences as uncertainty about the model structure and relative treatment effectiveness could not be resolved. With the severity modifier included, the probabilistic EAG base-case analyses indicated a cost effectiveness probability of [redacted] at a willingness-to-pay threshold of £20,000 per QALY gained. The most influential adjustments were using the exponential in no/failed SCT OS (instead of log-logistic) and using ITC 5 instead of ITC 1. The ICER increased most in the scenario analysis assuming no OS gain pre-landmark (an extreme scenario) and using equal probabilities of SCT and cure after SCT in both arms. The ICER only decreased in one scenario where the KEYNOTE-087 data was used to inform baseline OS and the relative treatment effect was informed using the MAIC of KEYNOTE-087 versus Cheah (2016)[3]. It is noteworthy that most scenarios individually appeared to have a modest impact, but used together significantly increased the ICER.

6.4 Conclusions of the cost effectiveness section

The CS adheres mostly to the NICE reference case, with the caveat that the evidence synthesis only partly adheres to the NICE reference case. It is a review that shows a significantly lower ICER ([redacted] without and [redacted] with the severity modifier) than in the original TA540, where ICERs were £36,950 (24-Week model) and £55,628 (12-Week model) per QALY gained as per the TA540 FAD[1]. The most prominent issues from the EAG perspective were: 1) the difference in results (compared with the original TA540) were not driven by the evidence collected during the CDF; 2) the Company's updated model structure; 3) pembrolizumab OS and HRQoL benefits; and 4) assumptions related to the comparator costs, utility values adopted pre-and post-landmark.

Firstly, it is important to mention that this improved ICER is not due to the evidence collected during the CDF, which showed generally worse survival than what was observed in KEYNOTE-087. In fact, the Company explained that the collected OS data from SACT had little impact on the model results. Instead, a combination of factors appears to explain the differences, including the omission of a progressed disease state with low utility values and the inclusion of a cured state with high utility values. In addition, it appears that the Company's new model structure, in particular using a landmark of 4 years to divide costs and effects into before and after SCT, also affects the model results.

Secondly, the Company's updated model structure is inconsistent with good modelling practices as it is suboptimal from a transparency perspective as outcomes are not estimated explicitly and it violates the homogeneity within health states assumption. The EAG is concerned that this might produce substantially biased results and thus proposed an alternative model structure with three health states that is simple, transparent, adheres to best modelling practices and does not require tunnel states. Compared with the Company's model structure, it would not necessitate using an arbitrary landmark point and allow including the probability of transitioning to successful SCT every cycle. Moreover, this is more consistent with a) the original TA540 FAD data collection recommendations that mainly focused on (time to) SCT as well as b) the importance of SCT in the disease pathway and mechanism through which pembrolizumab affects outcomes. Hence, it would be very informative to estimate "time to SCT or death" and incorporate this explicitly in the economic model.

Thirdly, there is still a lack of comparative effectiveness data, which leads to significant uncertainty about relative OS and HRQoL estimates. Uncertainty also remains about the duration of the pembrolizumab OS and HRQoL benefits, and probabilities of SCT and cure after SCT in the comparator arm. Importantly, the Company models treatment effect in multiple ways and not all of these were

evidence-based: in addition to a pre-landmark OS gain the Company also modelled a post-landmark OS gain for people with no/failed SCT. The EAG questioned that the evidence presented to inform this relative treatment effect was relevant to this setting.

Fourthly, the EAG questioned numerous model assumptions, including those related to the comparator costs; utility values adopted pre- and post-landmark and subsequent costs; and resolved some of these in the EAG base-case.

With the majority of changes made by the EAG, the ICERs increased. While most of the individual changes only had a relatively modest impact on the ICER, when most EAG scenarios were included in one analysis, the ICER went up substantially. Hence, there is substantial uncertainty about the most appropriate ICER and there is remaining uncertainty about the model structure and its impact that could not be fully resolved.

In conclusion, large uncertainty remains about the cost effectiveness of pembrolizumab versus SoC in R/R cHL, which can be at least partly resolved by the Company by conducting further analyses. The EAG considers the current model structure (both in the CS and EAG base-case) flawed and this could conceivably change the ICER. Therefore, the EAG believes that neither the CS nor the EAG report contain an unbiased ICER of pembrolizumab compared with SoC. Further data on the comparator, for example by obtaining access to historical SACT data (if available), could be helpful in informing relative treatment effectiveness and treatment use.

7. REFERENCES

[1] National Institute for Health and Care Excellence. *Pembrolizumab for treating relapsed or refractory classical Hodgkin lymphoma. NICE Technology appraisal guidance TA540 [Internet]*. London: NICE, 2018 [accessed 1.11.23] Available from: https://www.nice.org.uk/guidance/ta540/

[2] National Institute for Health and Care Excellence. *Pembrolizumab for treating relapsed or refractory classical Hodgkin lymphoma after brentuximab vedotin [review of TA540]: Final scope*. London: NICE, 2023 [accessed 12.7.23]

[3] 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

[4] Merck Sharp & Dohme (UK) Limited. *Pembrolizumab for treating relapsed or refractory classical Hodgkin lymphoma [review of TA540]: Submission to National Institute of Health and Care Excellence. Single technology appraisal (STA): Document B - Company evidence submission*, 2023 [accessed 25.9.23]. 193p.

[5] Eyre TA, Phillips EH, Linton KM, Arumainathan A, Kassam S, Gibb A, et al. Results of a multicentre UK-wide retrospective study evaluating the efficacy of brentuximab vedotin in relapsed, refractory classical Hodgkin lymphoma in the transplant naive setting. Br J Haematol 2017; 179(3):471-9

[6] MSD Uk Ltd. *Classical Hodgkin Lymphoma Advisory Board Meeting, London, Monday 10 July 2023. MSD data on file: 2023 [As referenced in the CS]*.

[7] European Medicines Agency (EMA). *Keytruda: assessment report [Internet]*, 2021 [accessed 20.9.23] Available from: https://www.ema.europa.eu/en/documents/variation-report/keytruda-h-c-3820-ii-0090-epar-assessment-report-variation_en.pdf

[8] National Institute for Health and Care Excellence. *Pembrolizumab for treating relapsed or refractory classical Hodgkin's lymphoma after brentuximab vedotin [review of TA540][ID5084]: Clarification letter*. London: NICE, 2023

[9] National Institute for Health and Care Excellence. *Pembrolizumab for treating relapsed or refractory classical Hodgkin's lymphoma after brentuximab vedotin [review of TA540] [ID5084]: Response to request for clarification from the ERG*, 2023 [accessed 31.10.23]

[10] European Medicines Agency (EMA). *Keytruda 25 mg/ml concentrate for solution for infusion: summary of product characteristics (SmPC) [PDF provided by the company]*, 2022 [accessed 20.9.23]

[11] 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 1.11.23] Available from: https://www.nice.org.uk/guidance/ta462/

[12] Follows GA, Ardeshna KM, Barrington SF, Culligan DJ, Hoskin PJ, Linch D, et al. Guidelines for the first line management of classical Hodgkin lymphoma. Br J Haematol 2014; 166(1):34-49

[13] Collins GP, Parker AN, Pocock C, Kayani I, Sureda A, Illidge T, et al. Guideline on the management of primary resistant and relapsed classical Hodgkin lymphoma. Br J Haematol 2014; 164(1):39-52

[14] Follows GA, Barrington SF, Bhuller KS, Culligan DJ, Cutter DJ, Gallop-Evans E, et al. Guideline for the first-line management of Classical Hodgkin Lymphoma - A British Society for Haematology guideline. Br J Haematol 2022; 197(5):558-72

[15] McGowan J, Sampson M, Salzwedel DM, Cogo E, Foerster V, Lefebvre C. PRESS Peer Review of Electronic Search Strategies: 2015 Guideline statement. J Clin Epidemiol 2016; 75:40-6

[16] Sampson M, McGowan J, Cogo E, Grimshaw J, Moher D, Lefebvre C. An evidence-based practice guideline for the peer review of electronic search strategies. J Clin Epidemiol 2009; 62(9):944-52

[17] Merck Sharp & Dohme (UK) Limited. *Pembrolizumab for treating relapsed or refractory classical Hodgkin lymphoma [review of TA540]: Submission to National Institute of Health and Care Excellence. Single technology appraisal (STA): Document B: Appendices C-P - Company evidence submission*, 2023 [accessed 25.9.23]. 256p.

[18] Morrison A, Polisena J, Husereau D, Moulton K, Clark M, Fiander M, et al. The effect of English-language restriction on systematic review-based meta-analyses: a systematic review of empirical studies. Int J Technol Assess Health Care 2012; 28(2):138-44

[19] Egger M, Zellweger-Zähner T, Schneider M, Junker C, Lengeler C, Antes G. Language bias in randomised controlled trials published in English and German. Lancet 1997; 350(9074):326-9

[20] Lefebvre C, Glanville J, Briscoe S, Featherstone R, Littlewood A, Marshall C, et al. Chapter 4: Searching for and selecting studies. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al., editors. *Cochrane Handbook for Systematic Reviews of Interventions version 6.3 (updated February 2022)*: Cochrane, 2022. Available from: http://www.training.cochrane.org/handbook

[21] Higgins J, Thomas J, Chandler J, Cumpston M, Li T, Page M, et al. *Cochrane handbook for systematic reviews of interventions version 6.3 (updated February 2022) [Internet]*: Cochrane, 2022 [accessed 4.3.22] Available from: https://training.cochrane.org/handbook

[22] Wells GA, O'Connell D, Peterson J, Welch V, Losos M, Tugwell P. *The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses [Internet]*, n.d. [accessed 9.8.23] Available from: https://www.ohri.ca/programs/clinical_epidemiology/oxford.asp

[23] Chen R, Zinzani PL, Fanale MA, Armand P, Johnson NA, Brice P, et al. Phase II study of the efficacy and safety of pembrolizumab for relapsed/refractory classic hodgkin lymphoma. J Clin Oncol 2017; 35(19):2125-32

[24] Chen R, Zinzani PL, Lee HJ, Armand P, Johnson NA, Brice P, et al. Pembrolizumab in relapsed or refractory Hodgkin lymphoma: 2-year follow-up of KEYNOTE-087. Blood 2019; 134(14):1144-53

[25] Merck Sharp Dohme. *KEYNOTE-087 clinical study report [as referenced in the CS]*.

[26] Moskowitz C, Zinzani PL, Fanale M, Armand P, Johnson N, Ribrag V, et al. Multicohort Phase 2 study of pembrolizumab for relapsed/refractory classical Hodgkin Lymphoma (R/R CHL): KEYNOTE-087. Haematologica 2016; 101(Suppl 1):319

[27] Moskowitz C, Zinzani PL, Fanale MA, Armand P, Johnson N, Ribrag V, et al. Pembrolizumab for relapsed/refractory classical Hodgkin Lymphoma: multicohort, Phase 2 KEYNOTE-087 study. Haematologica 2016; 101(Suppl 5):44

[28] Moskowitz CH, Zinzani PL, Fanale M, Armand P, Johnson NA, Radford JA, et al. *Pembrolizumab in Relapsed/Refractory Classical Hodgkin Lymphoma: Primary End Point Analysis of the Phase 2 Keynote-087 Study [Internet]*, 2016 [accessed 4.9.23] Available from: https://ashpublications.org/blood/article/128/22/1107/96203

[29] von Tresckow B, Fanale M, Ardeshna KM, Chen R, Meissner J, Morschhauser F, et al. Patient-reported outcomes in KEYNOTE-087, a phase 2 study of pembrolizumab in patients with classical Hodgkin lymphoma. Leuk Lymphoma 2019; 60(11):2705-11

[30] Zinzani PL, Chen R, Lee HJ, Armand P, Johnson NA, Brice P, et al. *Two-year follow-up of Keynote-087 study: Pembrolizumab monotherapy in relapsed/refractory classic hodgkin lymphoma [Internet]*, 2018 [accessed 4.9.3] Available from: https://ashpublications.org/blood/article/132/Supplement%201/2900/263697

[31] Zinzani PL, Lee HJ, Armand P, Johnson N, Brice P, Radford J, et al. *Three-year follow-up of Keynote-087: Pembrolizumab monotherapy in relapsed/refractory classic hodgkin lymphoma [Internet]*, 2019 [accessed 4.9.23] Available from: https://ashpublications.org/blood/article/134/Supplement_1/240/426193

[32] 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):3733-9

[33] Hanel W, Shindiapina P, Bond DA, Sawalha Y, Epperla N, Voorhees T, et al. A phase 2 trial of ibrutinib and nivolumab in patients with relapsed or refractory classical hodgkin's lymphoma. Cancers (Basel) 2023; 15(5)

[34] 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

[35] 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

[36] National Institute for Health and Care Excellence. *Brentuximab vedotin for treating CD30-positive Hodgkin lymphoma. NICE Technology appraisal guidance TA524 [Internet]*. London: NICE, 2018 [accessed 17.5.23] Available from: https://www.nice.org.uk/guidance/ta524/

[37] Signorovitch JE, Wu EQ, Yu AP, Gerrits CM, Kantor E, Bao Y, et al. Comparative effectiveness without head-to-head trials: a method for matching-adjusted indirect comparisons applied to psoriasis treatment with adalimumab or etanercept. Pharmacoeconomics 2010; 28(10):935-45

[38] Signorovitch JE, Sikirica V, Erder MH, Xie J, Lu M, Hodgkins PS, et al. Matching-adjusted indirect comparisons: a new tool for timely comparative effectiveness research. Value Health 2012; 15(6):940-7

[39] Phillippo DM, Ades AE, Dias S, S P, Abrams KR, Nj W. NICE DSU Technical support document 18: Methods for population-adjusted indirect comparisons in submissions to NICE. 2016;

[40] Drummond MF, Sculpher MJ, Torrance GW, O'Brien BJ, Stoddart GL. *Methods for the economic evaluation of health care programmes*. 3rd ed. Oxford: Oxford University Press, 2005.

[41] Jones B, Ward T, Harrison JP, Hurst M, Tyas D, McEwan P, et al. The cost-effectiveness of nivolumab for the treatment of people with relapsed or refractory classical Hodgkin lymphoma following autologous stem cell transplant and brentuximab vedotin. Value Health 2017; 20:A433

[42] Siebert U, Alagoz O, Bayoumi AM, Jahn B, Owens DK, Cohen DJ, et al. State-transition modeling: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force--3. Value Health 2012; 15(6):812-20

[43] Roberts M, Russell LB, Paltiel AD, Chambers M, McEwan P, Krahn M. Conceptualizing a model: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force-2. Med Decis Making 2012; 32(5):678-89

[44] Bojke L, Soares M, Claxton K, Colson A, Fox A, Jackson C, et al. Developing a reference protocol for structured expert elicitation in health-care decision-making: a mixed-methods study. Health Technol Assess 2021; 25(37):1-124

[45] Office for National Statistics. *National life tables: England [Internet]*, 2021 [accessed 25.8.23] Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/lifeexpectancies/datasets/nationallifetablesenglandreferencetables/current

[46] 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

[47] 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):860-7

[48] NHS England. *2021/22 National Cost Collection Data Publication [Internet]*, 2023 [accessed 20.11.23] Available from: https://www.england.nhs.uk/publication/2021-22-national-cost-collection-data-publication/

[49] BNF. *British National Formulary (BNF) [Internet]*, 2023 [accessed 29.8.23] Available from: https://bnf.nice.org.uk/

[50] Department of Health and Social Care. *Drugs and pharmaceutical electronic market information tool (eMIT) [Internet]*, 2023 [accessed 29.8.23] Available from: https://www.gov.uk/government/publications/drugs-and-pharmaceutical-electronic-market-information-emit

[51] Georghiou T, Bardsley M. *Exploring the cost of care at the end of life: research report [Internet]*. London: Nuffield Trust, 2014 [accessed 20.09.23] Available from: https://www.nuffieldtrust.org.uk/sites/default/files/2017-01/end-of-life-care-web-final.pdf

[52] NHS England. *National Schedule of NHS Costs - Year 2020-21 - NHS trusts and NHS foundation trusts [Internet]*, 2023 [accessed 25.8.23] Available from: https://www.england.nhs.uk/costing-in-the-nhs/national-cost-collection/

[53] NHS England. *DECC: Relapsed and progressive Hodgkin and Non Hodgkin Lymphoma*, 2016 [accessed 29.8.23] Available from: https://www.northerncanceralliance.nhs.uk/wp-content/uploads/2018/11/DECC-NECN-protocol-CRP08-H0062.pdf

[54] Lymphoma Group. *Mini-BEAM (modified) [Internet]*: Thames Valley Strategic Clinical Network, 2022 [accessed 29.8.23] Available from: https://nssg.oxford-haematology.org.uk/lymphoma/documents/lymphoma-chemo-protocols/L-30-mini-beam-modified.pdf

[55] National Institute for Health and Care Excellence. *Lung cancer: diagnosis and management. NICE guideline NG122 [Internet]*. London: NICE, 2019 [accessed 25.8.23] Available from: www.nice.org.uk/guidance/ng122

[56] Royal College of Radiologists. *Radiotherapy dose fractionation (third edition)[Internet]*. London: RCR, 2019 [accessed 25.8.23] Available from: https://www.rcr.ac.uk/publication/radiotherapy-dose-fractionation-third-edition

[57] National Institute for Health and Care Excellence. *Tisagenlecleucel for treating relapsed or refractory diffuse large B-cell lymphoma after 2 or more systemic therapies. NICE Technology appraisal guidance TA567 [Internet]*. London: NICE, 2019 [accessed 25.8.23] Available from: https://www.nice.org.uk/guidance/ta567/

[58] National Institute for Health and Care Excellence. *Asciminib for treating chronic myeloid leukaemia after 2 or more tyrosine kinase inhibitors. NICE Technology appraisal guidance TA813 [Internet]*. London: NICE, 2022 [accessed 25.8.23] Available from: https://www.nice.org.uk/guidance/TA813/

[59] NHS Blood and Transplant. *Unrelated donor: stem cell transplantation in the UK. A report from the UK Stem Cell Strategy Oversight Committee November 2014 [Internet]*, 2014 Available from: https://nhsbtdbe.blob.core.windows.net/umbraco-assets-corp/29047/unrelated-donor-stem-cell-transplantation-in-the-uk-2014.pdf

[60] Blommestein HM, Verelst SG, Huijgens PC, Blijlevens NM, Cornelissen JJ, Uyl-de Groot CA. Real-world costs of autologous and allogeneic stem cell transplantations for haematological diseases: a multicentre study. Ann Hematol 2012; 91(12):1945-52

[61] Personal Social Services Research Unit. *Unit Costs of Health and Social Care programme (2022 – 2027)*, 2022 [accessed 30.8.23] Available from: https://www.pssru.ac.uk/unitcostsreport/

[62] Grimm SE, Pouwels X, Ramaekers BLT, Wijnen B, Knies S, Grutters J, et al. Development and Validation of the TRansparent Uncertainty ASsessmenT (TRUST) Tool for assessing uncertainties in health economic decision models. Pharmacoeconomics 2020; 38(2):205-216

[63] 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

Single Technology Appraisal
Pembrolizumab for treating relapsed or refractory classical Hodgkin's lymphoma after brentuximab vedotin [review of TA540]
ID5084

EAG report – factual accuracy check and confidential information check

"Data owners may be asked to check that confidential information is correctly marked in documents created by others in the evaluation before release." (Section 5.4.9, NICE health technology evaluations: the manual).

You are asked to check the EAG report to ensure there are no factual inaccuracies or errors in the marking of confidential information contained within it. The document should act as a method of detailing any inaccuracies found and how they should be corrected.

If you do identify any factual inaccuracies or errors in the marking of confidential information, you must inform NICE by **5pm on 4 December 2023** using the below comments table.

All factual errors will be highlighted in a report and presented to the appraisal committee and will subsequently be published on the NICE website with the committee papers.

Please underline all confidential information, and information that is submitted as [redacted] should be highlighted in turquoise and all information submitted as [redacted] in pink.

Issue 1 **Wording of the description of data collection for the SACT dataset**

Description of problem	Description of proposed amendment	Justification for amendment	EAG response
Pages 12 (Table 1.1), 14 (Section 1.4), and 35. One of the EAG's key issues relates to the collection of data for the SACT dataset. Issue 3 is described in Table 1 as: "The recommended data collection from the TA540 has not been fulfilled as the majority of the presented outcomes of the SACT dataset do not differentiate between alloSCT and autoSCT".	For clarity, MSD suggest amending the wording to read: " The reported outcome of time to SCT as captured from the SACT dataset does not differentiate between timing of alloSCT and autoSCT".	MSD consider the EAG's text describing the data collection from pembrolizumab's time in the Cancer Drugs Fund (CDF) as "most outcomes not fulfilling the recommendations from TA540" to be misleading. As noted by the EAG, the Final Appraisal Determination (FAD) for TA540 proposed data collection on: • proportion of people having pembrolizumab who have an allogenic stem cell transplant (allo-SCT) • time to allo-SCT; • duration of treatment with pembrolizumab before allo-SCT; • long-term follow-up of people having pembrolizumab with or without subsequent allo-	Not a factual inaccuracy. As outlined in Section 3.2.1.2, not all of the data have been presented in a way meeting the requirements stipulated in the FAD.

		SCT (in particular, collection of data on overall survival [OS]). Two of the four outcomes captured from the SACT dataset align with the recommendations and are reported in the company's submission (CS) are: • overall survival (OS); and • proportion of patients who receive allo-SCT; The proportion of patients who underwent both auto-SCT and allo-SCT was reported in SACT, as described by the EAG in Table 1.4 (page 15). Therefore, of SCT-related outcomes reported in the CS, only time to SCT is not reported separately for auto-SCT versus allo-SCT. MSD note that the data collection agreement between MSD UK Ltd and Public Health England specified that data would be captured on overall	

		survival, duration of treatment, the proportion of patients who received a SCT, the time from treatment to transplant and intention to transplant: differentiation of auto-SCT from allo-SCT was not stipulated. MSD consider that data were collected by NHS England as agreed.	

Issue 2 Description of outcomes collected from the SACT dataset

Description of problem	Description of proposed amendment	Justification for amendment	EAG comment
Page 15 (Table 1.4) The EAG states: "In the FAD for TA540 recommendations for data collection the proposals for further data collection included: time to alloSCT; duration of treatment with pembrolizumab before alloSCT; long-term follow-up of people having pembrolizumab with or without subsequent alloSCT	MSD suggest amending the last two sentences to read: "Not all data were not available from the SACT dataset. Data for autoSCT and alloSCT combined were presented for the outcome time to SCT".	MSD suggest that the EAG's original text is misleading. Data were not available rather than were not presented by MSD. Additionally, data on OS were reported.	The wording in Table 1.4 has been amended to clarify this point. The underlying issue, whether and how the recommended data collection from TA540 was performed, remains.

(in particular, collection of data on overall survival). These data were not presented in the CS. Instead, data for autoSCT and alloSCT combined were presented for the outcome time to SCT".			

Issue 3 Description of population relevant to the decision problem

Description of problem	Description of proposed amendment	Justification for amendment	EAG comment
Page 15, EAG comment: There is a misalignment between the population defined by a) the National Institute for Health and Care Excellence (NICE) final scope[ref], b) the decision problem addressed in the CS[ref], c) the marketing authorisation[ref], d) the recommendation from TA540 [ref] and e) the inclusion criteria for the systemic anti-cancer therapy (SACT) data	Amend the statement to read: The population for which the company submitted evidence on clinical and cost effectiveness is narrower than the population defined in the final scope issued by NICE. The scope defined the population of interest as people with relapsed or refractory classical Hodgkin lymphoma (R/R cHL) who have had brentuximab vedotin (BV) and cannot have autologous stem cell transplant auto(SCT), whereas the company submitted evidence on patients who were SCT-naïve at time of treatment.	MSD consider the EAG's statement to capture incorrectly the status of NICE recommendations and the marketing authorisation of pembrolizumab for the treatment of relapsed/refractory classical Hodgkin lymphoma (R/R cHL). Additionally, the EAG comments on page 87 that the modelled population in the CS is as per the NICE final scope. Data for the modelled population are derived from	Not a factual inaccuracy. NB: This issue is discussed in Section 2.1 of the EAG report.

	The narrower population falls within the marketing authorisation for pembrolizumab and reflects the patient population eligible for treatment with pembrolizumab through the CDF.	sources described in the clinical effectiveness section. In brief, MSD consider the population in cohort 2 of TA540/KEYNOTE-087, those in SACT, those in the scope and those in our submission to be, in effect, the same patients, albeit described with slightly different wording. In TA540, NICE recommended pembrolizumab for use in the CDF as an option for treating R/R in adults who have had brentuximab vedotin (BV) and cannot have auto-SCT. MSD note that the recommendation restricted use to adults for the specific population, and the population is encompassed within the marketing authorisation for pembrolizumab in R/R cHL. MSD appreciate that it is not clear from the wording of the recommendation that the relevant patients are those who are SCT-naïve, but consider that, taken in the context of the	

		other recommendation from TA540 (described below), patients must be SCT-naïve to be eligible for treatment. As part of TA540, NICE did not recommend pembrolizumab as an option for R/R in adults who had undergone auto-SCT and treatment with BV. As part of the scoping process, MSD confirmed that we would not be submitting additional evidence from KEYNOTE-087 on those who had received auto-SCT and prior BV. Clinical and cost effectiveness of pembrolizumab in those who had not received prior brentuximab vedotin (BV) was evaluated through the technology appraisal process – TA772. TA772 recommended pembrolizumab as a treatment option for people aged 3 and older who have had an auto-SCT that has not worked or they have had at least 2	

		previous therapies and an auto-SCT is not an option, and only if: • they have not had brentuximab vedotin and; • the company provides pembrolizumab according to the commercial arrangement". Thus, patients who have received auto-SCT and no BV are not relevant to the review of TA540, but are captured within the marketing authorisation for pembrolizumab in the management of R/R cHL. Based on NICE recommendation outlined above, MSD consider that the population relevant to the decision problem is those who are SCT-naïve and have failed on two or more prior therapies, including BV. The inclusion criteria for access to pembrolizumab through the	

		CDF reflect the population for which NICE recommended pembrolizumab's entry into the CDF.	

Issue 4 Description of company's search for ongoing studies

Description of problem	Description of proposed amendment	Justification for amendment	EAG comment
Page 34. The EAG states: It is not clear why the Company stated that they have conducted separate searches to identify ongoing studies or why this was done. In addition, the Company's statement regarding the intervention's suitability is not entirely pertinent due to the lack of relevant data for this CS.	MSD suggest deletion of the text.	The statement is incorrect. As the EAG note in the same paragraph on page 34, and as noted in MSD's response to clarification, the search for ongoing studies was conducted to fulfil the request detailed in Section B.2.11 of the NICE STA template to identify ongoing studies. The study identified is ongoing and results have not been published.	Part of the text has been deleted, however, the issue remains.

Issue 5 Reporting of number of people from UK in Cohort 2 of KEYNOTE-087

Description of problem	Description of proposed amendment	Justification for amendment	EAG comment
Pages 34 and 79 In two instances on page 34 and one instance on page 79, the EAG state that 14 people in Cohort 2 were enrolled from the UK	Amend 14 to 10.	Typographical error. Page 54 of the CS indicates that 10 people from the UK were enrolled into Cohort 2; 14 people from the UK were enrolled into KEYNOTE-087, with 4 in Cohort 1 and 10 in Cohort 2.	Changed accordingly

Issue 6 Description of proportion of people receiving SCT

Description of problem	Description of proposed amendment	Justification for amendment	EAG comment
Page 49 The EAG states: “There is a notable difference between the proportion of people receiving SCT between the two cohorts (cohort 2 of	Amend “SCT” to “allo-SCT”. Alternatively, amend the text to read that the proportion of people receiving SCT is similar but there is a notable difference in the proportion of SCTs that were allo-SCT.	Typographical error. The proportion of people undergoing a SCT was similar for cohort 2 and the SACT dataset, with 29.6% and 30.2% of people receiving SCT, respectively. However, the proportion of SCTs that were allo-SCT was	Not a factual inaccuracy. The text on page 49 of the EAG report provides relevant context in support of the statement.

KEYNOTE-087 versus SACT dataset)”		37.5% for Cohort 2 and 64.6% for the SACT dataset.	

Issue 7 Description of statistical significance of a result

Description of problem	Description of proposed amendment	Justification for amendment	EAG comment
Page 73 The EAG states: “On the other hand, the naïve ITC for SACT versus SoC based on Eyre (2017)[ref] resulted in a higher HR of 0.66, while the difference between the treatments was not statistically significant (p=0.040) (Table 3.24)”.	Amend to read: “…of 0.66, and the difference between treatments was statistically…”	Typographical error. The difference between treatments is statistically significant. The reported p value is 0.04 and the 95% Confidence Interval (CI) reported in Table 3.24 (0.44 to 0.98) does not cross the line of no effect.	Wording amended (“not” removed”).

Issue 8 Comparators

Description of problem	Description of proposed amendment	Justification for amendment	EAG comment
Page 88 The EAG states: "In TA540 submission the comparator composition was solely based on the Cheah (2016) study which consisted of chemotherapy, bendamustine and investigational agents".	Amend to read: "In the TA540 submission the comparator composition was based on the Cheah (2016) study. The listed comparators from Cheah were amended to reflect clinical practice at that time according to the 'guideline on the management of primary resistant and relapsed classical Hodgkin Lymphoma' as published in the British Journal of Haematology Board (2014) and from a previous NICE R/R cHL appraisal, TA462. The final comparator composition for TA540 comprised chemotherapy, bendamustine and investigational agents."	MSD in the TA540 submission ensured the comparators were in line with clinical guidelines at that time and in accordance with previous NICE appraisals (i.e. TA462).	This has been amended.
Page 88 the EAG states the reference for ICE, Gemcitabine-based (IGEV, GEM-P, GDP, GVD) as: "Advisory Board".	Amend to read: "Cheah et al. (2016), Eyre et al (2017) and MSD Advisory Board"	Typographical error	Amended

Page 88 the EAG states the reference for oral chemotherapy (DECC), radiotherapy and Mini-BEAM as: "Eyre et al. 2017".	Amend to read: "Cheah et al. (2016), Eyre et al (2017) and MSD Advisory Board"	Typographical error	Amended
Page 89 the EAG states: "This approach is different from the approach followed in TA540 where SoC was based on the composition and proportions of Cheah (2016)"	Amend to read: "This approach is different from the approach followed in TA540 where SoC was based on the composition and proportions of Cheah (2016) and amended in accordance with the British Journal of Haematology Board and TA462".	MSD in the TA540 submission ensured the comparators were in line with clinical guidelines at that time and in accordance with precedence (i.e. TA462).	Amended
Page 89, the EAG states: "The Company explained that the composition of TA540 was not considered as it is not reflective to current clinical practice for R/R cHL".	Amend to read: "The Company explained that the composition of TA540 was not considered as the comparators from Eyre et al. (2017) were also listed (in addition to Cheah et al. 2016)) and subsequently validated by the Clinical Advisory Board. The conclusions from that Advisory Board suggested the comparator composition from TA540 was no longer reflective of current clinical practice for R/R cHL".	MSD suggest the EAG's current wording is misleading. The blended comparator for this submission deviated from the TA540 submission for the reasons given in the CS.	Not a factual inaccuracy.

Issue 9 Health-related quality of life

Description of problem	Description of proposed amendment	Justification for amendment	EAG comment
On page 97, the EAG states: "In absence of health state utility values from KEYNOTE-204, for the "alive post-landmark; no or failed SCT" health state, the Company assumed progressed disease health state utility values for this health state."	Amend to read: "In the absence of health state utility values from KEYNOTE-204, for the "alive post-landmark; no or failed SCT" health state, the Company assumed progressed disease health state utility values for this health state based on their rationale that the data were statistically significant and the clinical experts from the Advisory Board confirmed that QoL benefit may persist for many years after receiving treatment".	MSD suggest the EAG's current wording is misleading. The rationale for progressed disease utility value was stated in the CS the data from KEYNOTE-204 were statistically significant. Furthermore, clinical justification was made as to why no SCT/failed SCT were assigned the progressed disease health state utility value as evidenced by the Clinical Advisory Board.	Not a factual inaccuracy.

Issue 10 Description of rationale for modelling approach

Description of problem	Description of proposed amendment	Justification for amendment	EAG comment
Page 86: The EAG states that "Additionally, the Company's response to clarification question B4 suggests that	MSD suggest the following text be deleted: "Additionally, the Company's response to clarification question B4 suggests that choices regarding the model structure were data-driven".	MSD suggest that the EAG's statements are misleading as they do not fully capture the rationales provided by the	Not a factual inaccuracy.

choices regarding the model structure were data-driven".		company in response to clarification. The response to B4 is not solely based on data availability, but also indicates that alternative approaches may introduce additional structural uncertainty and potentially lack transparency.	
The EAG states that: "However, the Company responded to B6d that it did not adopt this model structure for the following reasons; i) the Company believes it will make very little difference to the decision and; ii) it would require the estimation of time to SCT or death which the Company considered to be a difficult composite end-point for clinicians to help estimate."	Amend text on alternative model structure to read: "However, the Company responded to B6f that it did not adopt this model structure for the following reasons; i) the Company believes it will make little difference to the decision, ii) it would require the estimation of an additional uncertain treatment effect or some sort of model calibration to estimate outcomes in the SoC arm and; iii) it would require the estimation of time to SCT or death which the Company considered to be a difficult composite end-point for clinicians to help estimate". Include text discussing the likely impact of the EAG's proposed model	The description of the response to B6 is incomplete. Typographical correction: The model structure is discussed in question B6f.	This typographical error was corrected.

	structure on the ICER (see company response to clarification question B6c and further details in justification for amendment)		
The company's position that the EAG's proposed model structure would slightly reduce the ICER for pembrolizumab, while introducing the need for additional calibration to model OS, has not been reflected.	Include text discussing the company's position (see company response to clarification question B6c and response to justification for amendment).	It would be extremely helpful for all stakeholders if the EAG could clarify that the expectation is that their desired model structure (which consists of either a partitioned survival model or state transition model where the "alive and without [successful] SCT" health state is dictated by a "time to SCT or death" curve or similar) would slightly reduce the ICER for pembrolizumab, were it possible to implement. This is because the model would be exactly the same after the landmark, because OS prior to the landmark would be exactly the same and therefore the only difference would be that alive patients prior to the landmark would be divided into those with successful SCT and	Not a factual inaccuracy. The impact is unknown as it depends on how the model will be parameterised and the related assumptions.

		those without. Given there are more successful SCT patients with lower health state resource use and higher utility in the pembrolizumab arm and all other costs would remain the same, the only effect of implementing this model would be to slightly lower the ICER. The data also do not exist to model this accurately, particularly for the SoC arm, and therefore calibration exercises would be needed, which are also methodologically suboptimal vs. using observed outcomes directly. Overall, while we sympathise with the EAG's view that such a model structure, were the data available, would provide a slightly more nuanced way to capture the outcomes of interest, it is critical that stakeholders understand that doing this could not plausibly have a meaningful influence on the ICER.	

Issue 11 Description of the landmark point

Description of problem	Description of proposed amendment	Justification for amendment	EAG comment
Page 86 the EAG states that: "Moreover, compared with the Company's model structure, it would not necessitate using an arbitrary landmark point and allow including the probability of transitioning to successful SCT every cycle."	Amend to read: "Moreover, compared with the Company's model structure, it would not necessitate using a landmark point and allow including the probability of transitioning to successful SCT every cycle". MSD also propose that the evidence used to inform the choice of landmark timepoint be discussed or cross-referenced by the EAG.	The description of the landmark point as "arbitrary" is not factually accurate. As described in section B3.2.4.3 of the company submission the choice of landmark timepoint was based on: • Data from the SACT to inform when all SCT events had resolved; • Evidence on SCT outcomes from the literature indicating a timeframe for when SCT events of interest would have occurred. These assumptions were validated by experts at an MSD Advisory Board. MSD acknowledge that the exact positioning of the	Not a factual inaccuracy. The EAG notes that a landmark point is almost always arbitrary, as discussed in NICE TSD 21: "Landmark time-points may be arbitrary, and may importantly influence the results of the analysis. An early landmark time-point may miss delayed responses, whereas a late landmark time-point may result in less meaningful categorisation as a proportion of patients (likely to be non-responders) may die before the landmark point is reached. This will have important

		landmark timepoint could potentially influence results and explored alternative landmark timepoints in scenario analyses. MSD consider that is important that the evidence used to inform the choice of landmark position is reflected fully in the EAG report.	consequences in terms of the estimation of uncertainty around the estimates also."

Issue 12 Description of evidence for OS and HRQoL benefits post-landmark

Description of problem	Description of proposed amendment	Justification for amendment	EAG comment
Page 87: The EAG states that "Given the 2-year pembrolizumab stopping rule and given the lack of evidence to support post-landmark OS and HRQoL benefits, the EAG preferred assuming no post-landmark OS and HRQoL benefits for pembrolizumab in its base-case".	MSD propose deletion of the phrase "lack of evidence" on page 87 and note that supportive evidence (such as clinical expert opinion and KEYNOTE-204 data) has been (partially) discussed elsewhere (e.g., on page 92 of the EAG report). MSD note that it is routine for pembrolizumab trials to report OS benefit beyond treatment cessation at 2 years. The debate at NICE committee meetings typically centres on when such	MSD understand the EAG's exploration of uncertainty in OS and HRQoL evidence post landmark and have provided scenario analyses excluding these treatment effects to support this exploration, while noting these scenarios are conservative.	Not a factual inaccuracy. There is a lack of evidence to support post-landmark OS and HRQoL benefits.

	a treatment effect will wear off with the most recent decisions being 3–5 years post cessation, which is why this was implemented in the model in sensitivity analysis. The addition of discussion on data provided in the KEYNOTE-204 CQ addendum would be valuable. **KEYNOTE-204 CQ addendum** • OS results are consistent at third line and fourth line plus groups, with outcomes on pembrolizumab unlikely to be meaningfully worse for 4L+ patients and outcomes on BV potentially slightly worse. This indicates that the results from KEYNOTE-204 are likely to be generalisable (and potentially conservative) when used as a surrogate for the KEYNOTE-087 population.	However, MSD consider the statements “lack of evidence” and “The only rationale” are misleading and notes that several sources of evidence appear not to have been fully considered by the EAG.	
Page 92: The EAG states that “The only rationale for why there would still be a difference in OS in the two groups ≥2 years after pembrolizumab	MSD propose deletion of the phrase “The only rationale”.	Please see comments in first row.	“Only” was deleted.

was stopped was: *"Clinicians at the UK advisory board confirmed that there would be a treatment effect for some years after stopping pembrolizumab but were unable to say how long this would last"* (CS, B.3.3.1.6)."			
Page 94: "the only justification given was that *"Clinicians at the UK advisory board confirmed that there would be a treatment effect for some years after stopping pembrolizumab but were unable to say how long this would last"*(CS, B.3.3.1.6).	The company proposes deletion of the phrase "The only rationale".	Please see comments in first row.	"only" was deleted.

Issue 13 Lack of contextualisation of discussion of severity modifier

Description of problem	**Description of proposed amendment**	**Justification for amendment**	**EAG comment**
Page 105, the EAG states:	Amend to read:	The current text indicates the level of SoC QALYs at which	Not a factual inaccuracy.

“Based on the abovementioned QALYs of 15.6 for the (age and gender matched) general population, the estimated QALYs for SoC should be >3.6 to result in an absolute QALY shortfall of less than 12, implying a 1.0 QALY weight.”	“In the EAG base case the estimated QALYs for SoC are [redacted], with only one company or EAG scenario (scenario 11: EAG pessimistic scenario) estimating >3.6 QALYs for standard of care.”	the decision around the severity weighting would change. The reason for highlighting this is unclear without the additional context of whether this threshold is reached in any of the analysis done by the company or EAG.	

Issue 14 Model validation

Description of problem	Description of proposed amendment	Justification for amendment	EAG comment
Page 108, the EAG states: “No further details of a systematic face validity assessment (including on the model structure) were provided in CS Section B.3.14.“	MSD suggest removing this statement.	MSD interpret face validity assessment to consider whether the decision problem, model structure, source and results correspond with reality. The key model inputs such as pre-landmark OS, baseline characteristics and treatment duration for pembrolizumab were based on the SACT, that is, the totality of UK real-world data that were available and	Not a factual inaccuracy.

		validated by clinicians at the UK Advisory Board. In addition, as stated in the CS, this indication is a particularly evidence-light area, therefore cross validating against data sources not included in this submission is much less feasible.	
Page 108, the EAG states: "No details of any technical verification were provided in CS Section B.3.14."	Amend to read: "No details of any technical verification were provided in CS Section B.3.14. The company later submitted their technical verification assessment in response to clarification."	Current wording suggests no technical verification was performed, which is misleading. MSD's proposed wording would now be consistent with the EAG's comment on page 109: "No technical verification was reported in the CS. In response to clarification questions, the Company provided their internal technical verification checklist."	Not a factual inaccuracy.

Issue 15 Description of relevance of KEYNOTE-204 to the decision problem

Description of problem	Description of proposed amendment	Justification for amendment	EAG comment
Page 78, the EAG states: "The Company is suggesting that if pembrolizumab is more effective than BV then it must, by implication, be at least that much more effective than SoC. Although this may sound plausible, such a statement should absolutely be supported by hard evidence, which the Company has not provided. A single clinical expert's opinion, while respected, cannot trump the lack of evidence."	MSD suggest amending this to "The KEYNOTE-204 trial provides evidence that pembrolizumab is more effective than BV in the R/R cHL population who are 3L+ and SCT naïve. NICE TA524 provides evidence that BV is likely to be more effective than SoC in this same population. The company is suggesting that an ITC using these two sources constitutes evidence that pembrolizumab is more effective than SoC in patients who have failed, on average, one more line of treatment. They further suggest that this analysis may be conservative because the indirectness in the patient population is likely to be a source of bias against pembrolizumab. This is because data supplied from KEYNTOE-204 and KEYNOTE-087 and clinical feedback have suggested that failing another line of chemotherapy is unlikely to meaningfully affect responses to pembrolizumab but their responses to standard chemotherapy treatments	The EAG's characterization is one of a total lack of hard evidence rather than an indirectness of population, which is misleading.	Not a factual inaccuracy.

	might be worse because the group would be more chemo-refractory on average. It is suggested that the NICE committee discusses the extent to which the indirectness of population affects the generalisability of the company's ITCs."		

Location of incorrect marking	**Description of incorrect marking**	**Amended marking**	**EAG comment**
Page 92	ICER is not redacted	The Company explored this in a scenario analysis where the standard mortality ratio was set to a factor of 1.2 and the ICER increased to £[redacted] per quality-adjusted life year (QALY) gained.	This was now marked as confidential.