in collaboration with:

ERASMUS UNIVERSITEIT ROTTERDAM
INSTITUTE OF HEALTH POLICY & MANAGEMENT

Maastricht University

Pembrolizumab for treating relapsed or refractory classical Hodgkin lymphoma

ERRATUM

This document contains errata in respect of the ERG report in response to the company's factual accuracy check.

The table below lists the page to be replaced in the original document and the nature of the change:

Page nr:	**Change:**
18	Replaced "confidential" by "commercial"
39	AiC marking has been added.
78	Corrected number of patients receiving alloSCT in KEYNOTE-087, added AiC marking
96	Corrected table 5.26 by replacing "Not addressed" by "Addressed in SA"

that any benefits will be obtained sooner than is likely to occur in clinical practice. Furthermore, the company's model assumed that no patients would progress after receiving alloSCT. These assumptions favour pembrolizumab.

The company informed alloSCT uptake conditional on response status at 12 weeks after treatment start through a UK clinician survey and then combined these survey results with the previously performed BMS survey results (from TA462). The appropriateness of combining both surveys is questionable. The appropriate approach for incorporating alloSCT in the model would have been to use time to alloSCT data directly from the main source of evidence. There remains major uncertainty about the alloSCT uptake estimates. Furthermore, the elicited alloSCT uptake (from the MSD survey) for patients with progressed disease was ignored. Both, the combining of both surveys and ignoring alloSCT uptake in progressed disease patients, were shown in scenario analysis to be major drivers of cost effectiveness.

A major limitation was the use of single-arm evidence to inform treatment effectiveness. There was uncertainty whether the MAIC or the naïve indirect comparison should be used. The company provided both and the ERG, like the company, used the naïve indirect comparison in the base-case and the MAIC in scenario analysis. Furthermore, the ERG viewed the immaturity of the OS data from KEYNOTE-087 as a major limitation as this necessitated the use of post-alloSCT OS and utility estimates from alternative data sources, one of which was based on 13 patients only. The methods used to extrapolate from this data source were also questionable. [redacted], and the ERG considers that these may be informative for the present analysis.

It is of note that the population used for the comparator was a mixed population of cohorts 1 and 2, that is, it included patients who did and did not receive autoSCT, derived from Cheah et al. The Cheah et al. population is more comparable with KEYNOTE-087 cohort 1 than with cohort 2 in terms of patient characteristics. The use of this mixed comparator population likely resulted in comparisons of pembrolizumab with SoC that may be favourable and non-favourable for pembrolizumab in cohorts 1 and 2 respectively, but this could not be formally explored in scenario analysis.

Of further note, the economic model, and the evidence from KEYNOTE-087, rely on the assumption that treatment with pembrolizumab is capped at 24 months, which is inconsistent with its SmPC. It is unclear whether in UK clinical practice pembrolizumab would also be provided for a maximum of 24 months. This assumption favoured pembrolizumab.

Model extrapolations lack face and external validity. For example, the company claims that End of Life criteria can be considered fulfilled, however, their model predicts life year gains of 53 months on standard of care.

1.7 Summary of exploratory and sensitivity analyses undertaken by the ERG

A number of issues were identified by the ERG. The ERG was able to adjust/correct some of these in its base-case. This resulted in ICERs (probabilistic) of pembrolizumab (with commercial access agreement (CAA)) versus SoC of £64,186 and £78,696 per QALY gained for cohorts 1 and 2 respectively.

Additional sensitivity analyses were performed to examine the potential impact of alternative assumptions on the cost effectiveness estimates. The scenarios with the largest impact were alternative assumptions for extrapolating post-alloSCT, an alternative survival model for extrapolating post-12-week PFS in cohort 2, the use of the MAIC instead of the naïve comparison and removing the cap of 24 months on TTD (Table 1.1).

Table 4.5: Quality assessment of the KEYNOTE-087 trial

	CS evaluation	ERG evaluation	ERG comment
Selection bias			
Representativeness of cohort	*	*	Representative of the cHL population but may not be representative of the UK population
Selection of non-exposed cohort	NA	NA	
Ascertainment of exposure	*	*	Assessment was made of number of patients who received at least one dose of treatment
Outcome of interest	*	*	Presence of the outcome of interest was assessed before exposure to the intervention.
Comparability of cohorts	NA	NA	
Outcome bias			
Outcome assessment	*	*	Outcomes were evaluated by an independent review committee (IRC).
Adequate duration of follow-up			Median follow up time was 15.9 months. This was adequate for ORR but not for PFS and OS.
Adequate follow-up of cohort	*	*	Explanations were provided regarding missing data or loss to follow up.

Source: CS, Table 12, page 68
CS = company submission; ERG = evidence review group; NA = non-applicable

ERG comments:

- The most important methodological aspect to note is that although the trial was well conducted, it represents a low level of evidence. It is a phase II, single arm, non-comparative trial which by its design has serious limitations. We cannot know whether the outcomes observed are a true reflection of the intervention. The role of natural history and baseline characteristics is not taken into account.
- The study had an adequate follow-up (median 15.9 months) for the main outcome evaluated (ORR defined as the proportion of patients who have complete remission (CR) or partial remission (PR)). However median progression free survival was immature and [redacted].

4.2.2.5 Main efficacy results of the KEYNOTE-087 trial

At the 21 March 2017 data cut off [redacted] of cohort 1 patients and [redacted] of cohort 2 patients remained on treatment. Table 4.6 gives the current status of the patients in the KEYNOTE-087 trial.

first in terms of AIC/BIC, (3) the ERG in TA462 considered the use of log-normal and Weibull models as more clinically plausible as they did not predict infinite survival, and (4) the company considered the Weibull more conservative than the lognormal. The lognormal was explored in the company's scenario analysis. Model predictions of the different models are shown in Table 5.9.

Table 5.1: Summary of the survival models (OS after alloSCT adjusted for all-cause mortality)

Item	Exponential	Weibull	Gompertz	Log-logistic	Log-normal	Generalised gamma	Lafferty 2017
Median (months)	53.13	64.62	266.78	58.41	61.86	87.39	--
Mean (months)	76.77	163.07	237.71	172.88	177.21	213.93	--
% at 1 year	85.73%	71.68%	63.33%	69.74%	70.01%	65.28%	64.17%
% at 2 years	73.39%	63.78%	55.90%	61.55%	61.93%	59.48%	53.47%
% at 5 years	53.77%	54.50%	53.58%	52.68%	53.33%	54.21%	53.47%
% at 10 years	21.09%	40.56%	52.90%	40.79%	41.77%	47.95%	--
% at 15 years	9.67%	34.13%	52.08%	35.78%	36.83%	45.43%	--
% at 20 years	4.43%	29.61%	50.80%	32.40%	33.45%	43.82%	--
% at 30 years	0.93%	23.46%	45.95%	27.88%	28.84%	39.63%	--
% at 40 years	0.20%	17.64%	34.77%	21.10%	21.83%	29.99%	--

Source: CS Table 69[1]

ERG comment: The ERG has concerns about (a) the appropriateness of using Lafferty et al.[21] for estimating post-alloSCT OS and (b) that the company over-estimates OS in post-alloSCT patients.

(a) The ERG questioned the appropriateness of using Lafferty et al[21] for post-alloSCT survival, given that in KEYNOTE-087, [redacted] patients had an alloSCT, of which only [redacted] were UK patients, compared with the 13 patients in Lafferty et al[21]. In response to the clarification letter,[10] the company explained that the KEYNOTE-087 study did not include the subsequent investigation of patients treated with pembrolizumab who were treated with a stem cell transplant. Furthermore, the company argued that OS data for the entire study population of KEYNOTE-087 were deemed to be too immature to provide robust extrapolations of survival and highlighted that Lafferty et al[21] was also used to inform TA462. Because Lafferty et al[21] is a very small study with questionable generalisability to the UK setting (see Section 4.2.3), its use means that there is substantial uncertainty around post-alloSCT survival, and alternative evidence was not explored.

(b) According to the company's Figure 3 in Appendix 17 of the CS,[27] (Figure 5.3) post-alloSCT survival is likely over-estimated. From this figure it appears that the company assumed no censoring after the last event until the end of the 5-year period. This results in an over-estimation of OS, as can be seen from the fitted curves that follow the plateau between 21 months and 5 years closely. It is unlikely that this plateau is a reflection of OS in patients post-alloSCT and the ERG considers it more likely that censoring occurred before the end of this 5-year period. The ERG acknowledges that there is uncertainty about the better approach, but notes that the company chose the approach that favoured pembrolizumab the most. The ERG therefore used the KM estimates from Figure 5.3 to reconstruct individual patient level data, allowing for censoring after the last event and before the end of the follow-up period, and used this in ERG scenario analysis, showing that the company's analysis significantly favoured pembrolizumab. The ERG's and the company's fitted curves are shown in Figure 5.4. As can be seen,

(c) Complete cross validation with TA462 was not performed by the company in both the CS and clarification response. The main differences between TA462 and the current assessment are the model structure, and how alloSCT is incorporated in the cost effectiveness model. TA462 used a three health states (progression-free, progressed, dead) semi-Markov model while the current model is composed of a short-term component (first 12 weeks), a decision tree element (at 12 weeks) and a long-term component (after 12 weeks). Additionally, progression was not allowed post-alloSCT in the current assessment while it was incorporated in TA462. Different assumptions were also made concerning the composition of SoC between the two assessments. All these discrepancies may have influenced the health benefits and costs obtained in the SoC arm. Table 5.25 compares the results of SoC between TA462 and the current assessment. The health benefits obtained from SoC were almost doubled and the costs of SoC were more than doubled in the current assessment compared to TA462. These discrepancies are most likely explained by the fact that patients in TA462 may receive alloSCT after 6 months while patients are considered for alloSCT after 12 weeks in the current assessment. These different assumptions have likely influenced health benefits and costs of SoC.

Table 5.2: Comparison of SoC results between TA462 and the current assessment

Assessment	Total QALY	Total costs
TA462[a]	1.870	£23,668
Current assessment[b]	3.684	£52,017
[a] Outcomes considered as the AC's most plausible analysis, retrieved from the committee papers for the second AC meeting, Table 4 of the ERG commentary on the company additional evidence [b] Retrieved from the corrected company's cost effectiveness model, post clarification response, Cohort 1		

5.3 Exploratory and sensitivity analyses undertaken by the ERG

Table 5.26 summarises all main issues highlighted by the ERG in Section 5.2, indicates the expected direction of bias introduced by these issues and whether these are examined in any analyses/incorporated in the ERG base-case.

Table Error! No text of specified style in document.**.3: Main ERG critique of company's submitted economic evaluation**

Issue	Bias introduced[a]	ERG analyses	Addressed in company analysis?
Model structure (section 5.2.2)			
• Incorporation of alloSCT at 12 weeks only	+/-	None	Not addressed
• No lag between decision and procedure	+	None	Not addressed
• No progressed disease health state post-alloSCT	+	SA	Addressed in SA
Population, interventions and comparators, perspective and time horizon (sections 5.2.3-5)			
• Comparator data based on mix of cohort 1 and 2	+ cohort 1, - cohort 2	None	Not addressed
• BSC only in scenario analysis	+/-	None	Not addressed
• Time horizon of 40 years	-	BC (FV)	Addressed in SA
Treatment effectiveness and extrapolation (section 5.2.6)			
• Use of alternative sources due to immature OS data from KEYNOTE-087	+/-	None	Requested, partially addressed

Issue	Bias introduced[a]	ERG analyses	Addressed in company analysis?
• Single-arm study used to inform treatment effectiveness	+/-	None	Not addressed