National Institute for Health and Care Excellence

Centre for Health Technology Evaluation

Pro-forma Response

ERG report

Bortezomib for induction therapy in multiple myeloma before high dose chemotherapy and autologous stem cell transplantation [ID610]

You are asked to check the ERG report from Southampton Health Technology Assessment Centre to ensure there are no factual inaccuracies contained within it.

If you do identify any factual inaccuracies you must inform NICE by **5pm**, **Friday 31 May**, using the below proforma comments table. All factual errors will be highlighted in a report and presented to the Appraisal Committee and will subsequently be published on the NICE website with the Evaluation report.

The proforma document should act as a method of detailing any inaccuracies found and how and why they should be corrected.

Issue 1 Proportion of patients who underwent SCT (p7)

Description of problem	Description of proposed amendment	Justification for amendment	SHTAC response
The ERG's assertion that "Data for the proportion of patients who underwent stem cell transplant (SCT) were not powered nor were statistical tests reported so results are uncertain." (p7) is incorrect.	We suggest this sentence be removed from the ERG report.	Data for the proportion of patients who underwent SCT is clearly reported in PETHEMA CSR p45 and MS p74, and the proportions are materially different between the TD group (61.4% 78/127) and the VTD group (80.8% 105/130). Also, whether a patient undergoes SCT or not is a tangible, binary and certain measure. Therefore, the ERG's conclusion that "results are uncertain" is unfounded.	The results are proportions of patients with the differences not tested statistically and the analysis was not powered. The text on p.7 has been amended to say that it is unclear whether there is a significant difference between groups.

Issue 2 Trial sponsorship / grants (p14)

Description of problem	Description of proposed amendment	Justification for amendment	SHTAC response
The use of the word "However" is incorrect and misleading in the following sentence on p14. "The MS states that the trials were independently conducted and not sponsored by the manufacturer (MS p.48). <u>However</u>, the Pethema trial publication[1] states that the trial was sponsored by the Spanish Pethema Foundation and that Janssen-Cilag and Pharmion supported the study costs through two grants to Pethema."	We suggest the word "however" be removed from the sentence.	While Janssen-Cilag (and Pharmion) indeed supported the trials by providing the study drugs free-of-charge and also the educational grants, the trials were investigator initiated studies over which Janssen-Cilag had no control (e.g. design of the trial, protocol writing, execution of the studies, analysis of the results). Therefore, the comments on MS and those on the PETHEMA study publication are entirely consistent. As such, the use of the word "however" is both incorrect and misleading.	Not a factual error. No action.

Issue 3 Definition of ORR (p18)

Description of problem	Description of proposed amendment	Justification for amendment	SHTAC response
The definition of ORR in p18 ("Overall response rate (ORR) defined as CR+nCR+VGPR+PR") is not correct in all trials.	We suggest this definition be changed to "Overall response rate (ORR) defined as the total proportion of patients who achieve PR or better"	Please see Issue 4 below.	See below.

Issue 4 Comparability of ORR results across trials (p18, p21, p31)

Description of problem	Description of proposed amendment	Justification for amendment	SHTAC response
The following statements are incorrect "The ORR results across the two trials therefore cannot be directly compared." (p18) "(D)ifferent definitions of ORR between the Pethema and Gimema trials means that results cannot be directly compared and should be interpreted with caution." (p21) "ORR is defined differently in the Pethema trial compared to the Gimema trial (and other three trials) making comparisons difficult." (p31)	We suggest the second paragraph on p18 ("ORR results for the Pethema and Gimema…") and the quoted sentence in p21 and p31 be removed from the ERG report.	The ORR results are in fact comparable between PETHEMA and GIMEMA trials because ORR is defined as "the total proportion of patients who achieve PR or better" (MS 6.3.5 p66). For example, in the case of PETHEMA (MS Table 24, ERG Table 5), the number of patients in the VTD group who achieve ORR can be calculated as 46 (CR) + 18 (nCR) + 46 (PR) = 110 (ORR) 33 for VGPR should not be added because it will result in double-counting. In GIMEMA (ERG Table 5), as nCR, VGPR and PR are distinct response categories, for the VTD group, 44 (CR) + 29 (nCR) + 73 (VGPR) + 74 (PR) = 220 (ORR) Since ORR means "PR or better" in both cases, ORR results are comparable across trials.	Thank you for the further clarification. Our comment stemmed from the statement in the MS (p.67): 'ORR is calculated from these measures combined, as the total proportion of patients who achieve PR or better; ORR=CR+nCR+VGPR+PR.' The paragraph on p.18 of the ERG report has been amended and we have removed the text stating that ORR results cannot be compared on p.21, 25 and 31.

Issue 5 Response rate as a predictor of long-term outcomes (p31)

Description of problem	Description of proposed amendment	Justification for amendment	SHTAC response
The following statement in the ERG report is incorrect. "Response rate is a surrogate outcome and <u>it is not clear how good a predictor of long term outcomes it is</u>; post-transplant response may be better than post-induction response." (p31)	We suggest the following sentence in p31 be removed. "There are uncertainties around the appropriateness of the primary outcome measure in these trials. Response rate is a surrogate outcome and it is not clear how good a predictor of long term outcomes it is"	As noted in MS Section 7.3.4, there is a strong relationship between depth of response from induction and overall survival (MS p131). van de Velde et al (2007) performed a meta-analysis of 21 prospective and retrospective published studies that specifically reported on the response rate post-induction/pre-HDT/SCT. The association between post-induction, pre-HDT/SCT response and overall survival was highly statistically significant (p=0.0015), Given the evidence, it is incorrect and misleading to make a statement as in the ERG report p31.	The SHTAC comment was informed by our clinical expert. Whether response predicts long-term outcomes is not in question, but rather the strength of the correlation. The van de Velde systematic review is of poor methodological quality (and of 21 studies, only 4 are prospective, single transplant studies). The association between post-induction/pre-HDT/SCT response and OS was not statistically significant for the 10 prospective studies. Therefore, it is still unclear how strong a predictor of long-term outcomes response is. 'Good' has been replaced with 'strong' in this sentence for clarity.

Issue 6 Label used in Table 14 (p40)

Description of problem	Description of proposed amendment	Justification for amendment	SHTAC response
The formula that explains CR (CR+nCR+VGPR) in Table 14 is incorrect.	We suggest the formula for CR in Table 14 be changed to "CR+nCR".	We acknowledge this error is due to a mistake in MS Table 50. As the CR and PR rates in MS Table 50 and in the economic model are computed correctly, we can confirm this label error has no other impacts (such as cost per	We agree this should be changed to CR + nCR.

		QALY gained).	

Issue 7 Use of post-transplant response rates in the economic model (p40)

Description of problem	Description of proposed amendment	Justification for amendment	SHTAC response
The following statement in p40 appears incorrect. “The ERG considers that incidence of comorbidities and other patient characteristics may be assumed to be balanced between treatment arms in a properly randomised trial, and that on this basis it would be appropriate to use post-SCT response in the economic model”.	We suggest this sentence be removed from the ERG report.	Post-induction response rates allow a non-confounded assessment of disease response following induction treatment. Subsequent assessments, such as post-transplant response rate, are potentially confounded by study design features that were introduced to ensure treatment success, to accommodate local practice, or to investigate other aspects of the overall treatment scheme. Although uniform procedures for stem cell collection, conditioning with melphalan and SCT were provided in the protocol, unlike the induction phase, these procedures could be replaced by institutional preferences, if required. Only general criteria were specified for advancement to stem cell collection or HDM/SCT. In the PETHEMA trial, the protocol required a single course of HDT/SCT, and the minimum number of CD34+ cells to be collected was left to the discretion of each centre, as were other procedures related to stem cell cryopreservation, storage, thawing, and infusion. In the event of mobilisation failure with G-CSF, use of a second mobilisation scheme including use of high doses of cyclophosphamide followed by normal or pegylated G-CSF or any other scheme used as part of the centre’s standard clinical practice could	This is not a factual inaccuracy, rather it is an opinion given by ERG. No action.

		be used. It is not known how these practices in Spain would differ from UK clinical practice or how they differed between the investigator hospitals and regions. Therefore, using the post-SCT response rates introduces an element of uncertainty in the model which can be negated by the use of post-induction response rates.	

Issue 8 Labels used in Table 18 (p44)

Description of problem	Description of proposed amendment	Justification for amendment	SHTAC response
The labels used in Table 18 of the ERG report (which was reproduced from MS Table 55, p125) is not accurate	We suggest the phrase "5 year survival time" be replaced with "Median overall survival"	While the labels used in MS Table 55 (and therefore ERG Table 18) were inaccurate, it will have no impact on the cost-effectiveness results because the economic model handled the OS data from MRC VII appropriately.	We agree this should be changed to "Median overall survival".

Issue 9 Use of the OS data from PETHEMA trial (p50)

Description of problem	Description of proposed amendment	Justification for amendment	SHTAC response
The comment on p50 ("The ERG considers that the OS data from the Pethema trial provide an appropriate contemporary validation dataset for the VTD vs. TD model") appears inconsistent with the ERG's assessment of robustness of OS data on p20 ("The CSR[8] for the Pethema trial shows that a high	We suggest a suitable caveat be added to p50 in line with the comment on p20.	The proposed amendment will remove an internal inconsistency within the ERG report. As noted in the ERG report p20, the OS data from PETHEMA study is highly censored and immature with the median follow-up of only 35.9 months. This contrasts with the median OS of 88.6 months for CR patients from the MRC VII study.	A caveat has been added to p.50 in line with the comment on p.20.

proportion of patients in both the VTD and TD arms were censored: 57.7% and 44.9% respectively in the PFS analysis, and 80.0% and 74.8% respectively in the OS analysis (Table 38 on p.72 and Table 44 on p.76 of the CSR[8]). Given this censoring, the ERG suggests that there is uncertainty about the robustness of the results.")			

Issue 10 Testing structural uncertainty (p52)

Description of problem	Description of proposed amendment	Justification for amendment	SHTAC response
The comment on p52 ("There is, however, no evidence that structural uncertainties have been addressed via sensitivity analysis") is incorrect.	We suggest this sentence be removed from p52.	The current NICE TA Methods Guide states "(t)he impact of structural uncertainty on estimates of cost effectiveness should be explored by separate analyses of a representative range of plausible scenarios. Examples of when this type of scenario analysis should be conducted are: when there is uncertainty about the most appropriate assumption to use for extrapolation of costs and outcomes beyond trial follow-up" (p47). MS in fact reports a series of sensitivity analyses to test structural uncertainties (MS Section 7.6). For example, different extrapolation methods / survival distributions were explored for the 1st line TPP. Also, alternative data source for long-term OS data (IFM90) was used to extrapolate the OS data in the economic model and to explore its impact on cost effectiveness.	This sentence has been removed from p.52.

Issue 11 Additional work undertaken by the ERG - a) (p55-56)

Description of problem	Description of proposed amendment	Justification for amendment	SHTAC response
The first additional analysis (p55-56) contains now incorrect and irrelevant comparisons between TD/CTD, VD and PAD.	We suggest the first additional analysis ("a) Comparing all treatment analyses") be removed from the ERG report.	As noted on p11 of the ERG report, the EMA will not approve PAD regimen. Therefore, the incremental analyses shown on p55-56 are now incorrect, irrelevant, and potentially misleading.	This is not a factual inaccuracy. A footnote has been added to Tables 21 and 22 to clarify that EMA has not approved the PAD regimen.