Response to the Appraisal Consultation Document (ACD)

Bortezomib for induction therapy prior to high dose chemotherapy and autologous stem cell transplantation for the treatment of multiple myeloma [ID610]

December 2013

Table of contents

1. Overview 2
2. Bortezomib in combination with thalidomide and dexamethasone (VTD) 4
   2.1 Janssen's comments on the critique by the ERG and the appraisal committee 4
   2.2. Amended economic analysis 10
3. Bortezomib in combination with dexamethasone (VD): Overall population 18
   3.1 Model structure and input data 18
   3.2 Results 23
4. Bortezomib in combination with dexamethasone (VD): when CTD is not suitable 25
   4.1 When is CTD not suitable as an induction therapy? 25
   4.2 Which regimen is used if CTD is not suitable? What is the most relevant comparator? 25
   4.3 Model structure and input data 25
   4.4 Results 26
5. Conclusion 28
References 29
Appendices 30

1. Overview

Janssen welcomes the opportunity to comment on the preliminary recommendation made by the appraisal committee and to respond to the request for additional analyses detailed in the appraisal consultation document (ACD).

We are disappointed that neither bortezomib in combination with thalidomide and dexamethasone (VTD) nor bortezomib in combination with dexamethasone (VD) is recommended for the induction treatment of adults with previously untreated multiple myeloma (MM) for whom high-dose chemotherapy with haematopoietic stem cell transplantation (SCT) is suitable.

VTD and VD remain the only licensed regimens for MM in the induction setting. The most commonly used treatment in England and Wales, cyclophosphamide in combination with thalidomide and dexamethasone (CTD), is unlicensed and suffers from a limited evidence base as this regimen is not widely used outside of the UK. Moreover, as induction therapy occurs shortly after diagnosis at the early stage of the patient pathway, the RCT data to support this indication are somewhat immature due to the prolonged survival in this patient population. This combination of immature trial data and comparator issues cause some problems from an HTA perspective and we hope NICE can work with us to try to find a pragmatic solution that will enable MM patients to access this important treatment option.

The license for VTD and VD was granted after the efficacy and safety of the bortezomib-based induction regimens was robustly demonstrated in multiple randomised clinical trials (RCTs) such as PETHEMA, GIMEMA and IFM-2005. The bortezomib-based regimens are now standard of care in many other European countries.

The decision problem for this single technology appraisal (STA) rests upon whether the addition of bortezomib to induction regimens demonstrates value for money to the NHS. As demonstrated in the original submission, in the addendums and in this document, we maintain that bortezomib is highly cost-effective when added to thalidomide plus dexamethasone (i.e. the VTD regimen), or to dexamethasone alone (i.e. the TD regimen).

Whilst the negative recommendation for VTD appears to be based on three key factors, we believe none of these factors are supported by evidence:

- Use of the Alvares long-term OS data: The Alvares data are from a non-randomised, *retrospective review* covering a highly select group of patients who actually received a SCT. (On this point, please note that the final scope for this STA specifies "adults with newly diagnosed multiple myeloma for whom induction therapy is considered appropriate prior to high dose chemotherapy and autologous stem cell transplantation" as the relevant population, not those who actually received a SCT.) In contrast, MRC VII is the most recent, UK-based prospective dataset in the relevant patient population for which long-term OS data by response category have been reported. Moreover, compared with the MRC VII data, the Alvares data are neither newer nor fit better to the PETHEMA Kaplan-Meier curve for patients with less than partial response (PR). (See Section 2.1.1 below for more details.)
- Use of the post-transplant response rates: For those patients who never received a SCT, post-induction response rates are the most relevant response rates. It is not logical to apply the post-transplant response rates to patients without a SCT. Moreover, contrary to the ERG's assertion, the "maximal response" from MRC VII is in fact defined as the post-induction response, not the post-transplant response. Therefore, it is not appropriate to apply the post-transplant response rates to the economic model submitted in July 2013. (See Section 2.1.2 below.)
- Incremental clinical efficacy of cyclophosphamide in addition to TD: The committee appears to believe that CTD is substantially more efficacious than TD when the most plausible ICER estimate was considered. This is not consistent with the opinion expressed by clinical experts; CTD and TD can be considered clinically equivalent and broadly similar. It is

important to acknowledge, therefore, that TD should be considered a relevant comparator for this appraisal. The final scope clearly states that "chemotherapy regimens containing thalidomide" are relevant comparators. (See Section 2.1.4 below.)

Furthermore, the ERG exploratory analysis presented at the appraisal committee meeting was inappropriate, and we are concerned that an inordinate amount of discussion during the meeting was taken up with this analysis, which was based upon a comparison of single arms from heterogeneous trials taken from two different economic models. The conclusions drawn from this analysis are spurious, particularly the assertion that VTD is dominated by CTD. This is wholly inappropriate and we are concerned about the inclusion of this statement in the ACD (paragraph 4.12).

It is also regrettable that, at the meeting which began a few hours behind schedule, clinical and patient experts were not given adequate time to explain the clinical need for VTD and VD as well as how these regimens would fit in to the patient pathway. The clinical discussions appeared to be rushed and, consequently, the ACD contains a number of important outstanding clinical questions. For example, the ACD seems to indicate that clinical need for VTD might be weaker than that for VD by stating: "there was no specific preference for bortezomib triple therapy also including thalidomide" (paragraph 4.12). Given the apparent misunderstanding around the clinical place in induction therapy of bortezomib, we would request that the clinicians and patient groups present at the first appraisal committee meeting be permitted to attend the second meeting in February 2014.

In the following sections, we discuss the cost effectiveness of:

- VTD compared with TD (Section 2),
- VD compared with CTD (Section 3), and,
- VD compared with an alternative comparator in circumstances when CTD is not suitable (Section 3).

In the ACD, the ERG and the appraisal committee made a number of suggestions that would enhance the robustness and transparency of the economic analysis. We welcome such suggestions and have adopted them to amend the economic model. This is discussed in details in Section 2.2.

In order to test the robustness of the ICER estimates, a number of different data sources, including the Kaplan-Meier curves from the PETHEMA and IFM-2005 studies, were used in the additional economic analysis undertaken for this response. The majority of the ICER estimates were below £30,000/QALY, and many of them were below £20,000/QALY. Along with the results already reported in our original submission and the addendums, the cost-effectiveness analyses in this document clearly indicate that both VTD and VD are cost-effective when compared to a relevant comparator.

2. Bortezomib in combination with thalidomide and dexamethasone (VTD)

This section focuses on bortezomib in combination with thalidomide and dexamethasone (VTD) compared with a thalidomide-based regimen, and is split into two sub-sections:

- Janssen's comments on the critique made by the ERG and by the appraisal committee, and;
- an amended economic analysis that reflects suggestions made in the ACD.

2.1 Janssen's comments on the critique by the ERG and the appraisal committee

2.1.1 Source of the long-term OS data

In the base case analysis, we used the median overall survival (OS) from the MRC Myeloma VII trial to extrapolate the OS in each response category. Both the ERG and the appraisal committee criticised the use of the MRC Myeloma VII, and suggested that the Alvares study be used instead.

We maintain that the MRC Myeloma VII is the most relevant and appropriate data source and preferable to the Alvares study for the following reasons:

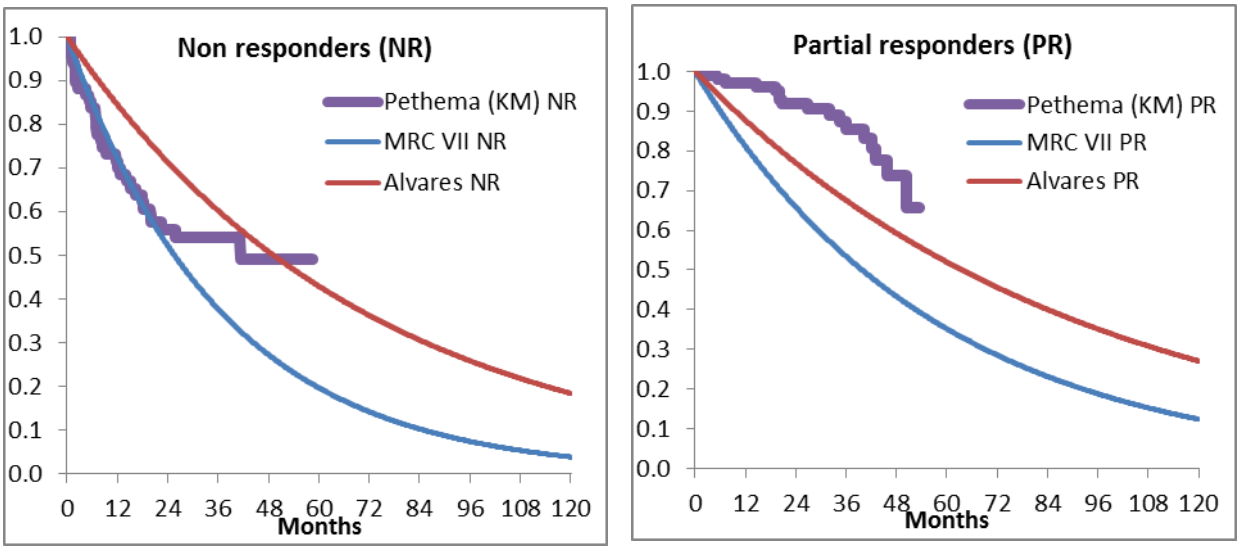
- MRC Myeloma VII was a prospective, randomised controlled trial based in the UK for the relevant patient population. While the Alvares study was also UK-based, it was a non-randomised, retrospective review in a single centre. Therefore, on the basis of the evidence hierarchy, the MRC Myeloma VII data are more robust and appropriate.
- The Alvares median OS figures quoted in the ERG report are for a highly select group of patients who actually received a SCT. For the MRC Myeloma VII trial, the median OS data are for the intensive-therapy group on an ITT basis. Therefore, patients included in the Alvares study may have had a better prognosis than those in the MRC Myeloma VII study, and this may have contributed to a longer median OS for patients with less than partial response (PR) in the Alvares study (49.2 months) than in the MRC Myeloma VII study (25.6 months).
- The ERG stated that "the MRC VII trial was old and its outcomes may not be reflective of the more advanced treatments available today" (ACD paragraph 3.33, page 25). However, since nearly half of patients in the Alvares study predated those in the MRC Myeloma VII trial (Table 1), the data from the latter are more current and relevant. It is important to note that the long-term OS data by response category must be derived from a historical study because the survival curves of different regimens in this population only begin to clearly diverge after several years post study initiation, given the relatively small number of patients enrolled in these pivotal studies, coupled with the relatively long survival in this population (please refer to the original February 2013 submission, page 131).

Table 1: Comparison of MRC Myeloma VII and Alvares studies

	MRC Myeloma VII 2003 NEJM	Alvares et al. 2005 BJH
Study design	RCT: standard chemo vs. HDT/SCT	Retrospective review
No. of patients	407 (201 in the HDT/SCT arm)	383
Patients – inclusion criteria	Previously untreated, age <65, suitable candidates for high-dose therapy	Patients with previously untreated MM were entered into the transplant programme
Year of enrolment	1993-2000	1985-2004 Out of 282 who received HDT; ○ 1985-89 n=66 (23%) ○ 1990-94 n=77 (27%) ○ 1995-99 n=82 (29%) ○ 2000-04 n=57 (20%)
Post-SCT response CR/PR/Less than PR	44% / 42% / 14%	50% / 10% / 40%
Median OS CR/PR/Less than PR	88.6 mos. / 39.8 mos. / 25.6 mos. 7.4 years / 3.3 years / 2.1 years	88.8 mos. / 63.6 mos. / 49.2 mos. 7.4 years / 5.3 years / 4.1 years

- The ERG commented that "the Alvares data provided the better fit to the PETHEMA overall survival data" (ACD paragraph 3.33, page 25). A close examination of the OS data from the PETHEMA trial revealed this statement is not accurate. In the PETHEMA study, more than 45% of patients with less than PR (non responders [NR] in the economic model) had died within 24 months and, consequently, the Kaplan-Meier curve is fairly steep (Figure 1, left). The exponential curve based on the Alvares result overestimates the survival of NR patients. In contrast, the MRC Myeloma VII curve fits extremely well to the PETHEMA Kaplan-Meier curve for NR. For PR patients, both the MRC Myeloma VII and the Alvares results underestimate OS compared with the PETHEMA Kaplan-Meier curves (Figure 1, right). It should be noted, however, that the OS data for the PR group are immature with the median OS not reached. Therefore, it is unclear which OS data would fit better to the Kaplan-Meier curve had more mature OS data been available from PETHEMA. The choice of the OS data source is explored further in Section 2.2.1.

Figure 1: OS curves: Kaplan-Meier curve from the PETHEMA study vs. MRC Myeloma VII and Alvares



2.1.2 Post-induction response and post-transplant response

The ERG commented that "data from the MRC VII trial related to maximal response to treatment rather than post-induction response rate" (ACD paragraph 3.29, page 22). This is, however, not an accurate description of how maximal response was defined in the MRC VII trial. According to Child et al.(1), the cycle of intensive therapy was "repeated every 21 days until a maximal response was attended" before "stem cells were harvested" (page 1876). Therefore, the definition of maximal response in the MRC VII study is in fact akin to the post-induction response in the economic model, rather than to the post-transplant response.

Moreover, it is illogical to apply the post-transplant response rates to those who never received a SCT. For such patients, the post-induction response rates are the only relevant response rates in the context of this economic evaluation.

However, we agree with the comment from the ERG: "it would have been more transparent to distinguish the separate effects on survival of post-induction response and stem cell transplantation" (ACD paragraph 3.29, page 23). Such a model design would allow the right type of response rates to be applied to the right patient group, i.e. post-induction response rates for the ITT population, and post-transplant response rates for those who actually received a SCT. This issue will be explored further in Section 2.2.1.

2.1.3 Most plausible ICER quoted in the ACD

As discussed above, it is not appropriate to use the OS data from Alvares or apply the post-transplant response rates to those who did not receive a SCT. Therefore, we do not agree with the assertion that "the ERG's exploratory analyses were appropriate and that £39,000 per QALY gained was an appropriate starting point for discussion on the most plausible ICER" for VTD compared with TD (ACD paragraph 4.13, page 38). Rather, £39,000/QALY is an extreme value based upon an unrealistic scenario. We maintain that the base case ICER of £20,468/QALY (from the July 2013 addendum) is the most realistic estimate.

2.1.4 Incremental clinical efficacy of cyclophosphamide in addition to TD

We acknowledge that CTD is currently the most commonly used regimen in the induction setting, and thus would be the ideal comparator. However, there is no comparative RCT evidence between VTD and CTD, unlike those between VTD and TD. It should be noted that both CTD and TD are within the scope of this STA as a relevant comparator: "chemotherapy regimens containing thalidomide".

Therefore, the PETHEMA study, which compared VTD against TD, provides robust and relevant RCT evidence upon which to base the economic evaluation. Once the plausible ICER is established based on the VTD-TD comparison, the next logical step is to assess how much, if any, the addition of 'C' (cyclophosphamide) to TD would enhance the response rates.

This issue was explicitly discussed in the ERG report, at the appraisal committee meeting, and in our original submission. The clinical experts agreed that TD and CTD are clinically broadly similar:

- "[T]he ERG clinical expert stated that this assumption [that TD and CTD are clinically equivalent] was reasonable in the absence of trial data." (Premeeting briefing paragraph 4.22, page 26)
- "The clinical specialists stated that thalidomide and dexamethasone may be a comparator and the 2 regimens [TD and CTD] could be considered broadly similar." (ACD paragraph 4.2, page 29)
- "[TD is] considered to be equivalent to CTD by UK clinicians" (Manufacturers original submission, page 13)

The clinical experts stated at the appraisal committee meeting: "the advantage of using a triple therapy such as cyclophosphamide, thalidomide and dexamethasone was that there was more flexibility to reduce doses in the case of toxicity" (ACD paragraph 4.2, page 29), rather than to enhance the response rates *per se* compared to the TD double therapy.

Therefore, we do not agree with the assertion that the ICER for VTD compared with CTD is "likely to be substantially higher than" £39,000/QALY (ACD, page 47), which implicitly assumes a substantial incremental efficacy benefit provided by 'C' in addition to TD alone.

In the addendum submitted in July 2013 (representing our base case, in line with the approved license), we stated that "given the uncertainty of the relative effectiveness of TD compared to CTD, a sensitivity analysis was undertaken in the form of a threshold analysis whereby post-induction response rates for TD were increased until the ICER reached £30,000 per QALY gained" (Addendum, page 23). The threshold analysis indicated that the CR rate for CTD would have to be nearly double of the PETHEMA result for TD (32.6% vs. 17.3%) for the ICER for VTD to reach the £30,000/QALY-mark, a clinically implausible result (Table 2).

Table 2: Threshold analysis for post-induction response rates from TD (from the July 2013 addendum)

	Post-induction response to TD			Incremental costs	Incremental QALYs	ICER
	CR	PR	NR			
Base case: *PETHEMA trial*						
	17.3%	44.1%	38.6%	+£20,682	+1.01	**£20,468**
Scenario 1: *Higher CR, lower PR, same NR rates vs. base case*						
	32.6%	28.8%	38.6%	+£19,991	+0.67	**£30,000**
Scenario 2: *Same NR rate as VTD in PETHEMA. Higher CR, lower PR rates vs. base case*						
	28.2%	56.4%	15.4%	+£21,270	+0.71	**£30,000**

Given the lack of any RCT data comparing VTD with CTD, the threshold analysis provides a useful way to simulate the impact of cyclophosphamide on the cost-effectiveness of VTD. Therefore, it is regrettable that the threshold analysis was not featured at all during the appraisal committee meeting.

Furthermore, it was equally regrettable that the slide set presented during the appraisal committee meeting heavily featured the ERG exploratory incremental analyses with severe limitations (Lead team presentation: Cost effectiveness, slides 22, 25, 27-29). The analyses were based on the comparison of individual arms of separate heterogeneous trials, did not account for potential differences in patient characteristics and on the outputs from *separate economic models*. Whilst the slides included a comment about the limitations, it was not appropriate to explore the cost-effectiveness of VTD compared with CTD using analyses with such severe limitations.

In the absence of comparative clinical data, it is sometimes necessary to match single arms of different trials to derive an estimate of comparative efficacy. However, when comparative RCT evidence is available against a relevant comparator (such as those from PETHEMA), we understand that it is not standard practice to match single arms from different economic models and compare them directly.

Instead, the discussion of cost-effectiveness of VTD compared with CTD should be based upon the clinical expert opinion surrounding the incremental clinical benefit of cyclophosphamide as well as on robust economic analyses using the PETHEMA RCT results.

We maintain that VTD is cost-effective, regardless of whether it is compared with TD or with CTD.

2.1.5 Transition probabilities to the second-line therapy

The rates of progression from the post-transplant or post-induction states (for those with or without a SCT, respectively) to the second-line treatment were computed using the time to progression (TTP) data from the PETHEMA study. Parametric curves were fitted to the patient-level post-induction TTP data from the PETHEMA study by induction treatment option (TD and VTD) and by post-induction response (CR, PR and NR). (Refer to Sections 7.3.2 and 7.3.7 of our original submission for further details.)

Commenting on this approach, the ERG stated that the exponential distribution fitted to the PETHEMA CR TTP for VTD patients resulted in a shorter median survival time (approximately 61 months) than the exponential distribution fitted to CR TTP data for patients receiving TD (median survival approximately 98 months). The ERG also stated that "this contrasted with overall findings for progression-free survival in the trial publication in which median progression-free survival was significantly higher" with VTD than with TD (ACD paragraph 3.31, page 24).

We accept the ERG's critique and agree that our original approach led to the counter-intuitive and clinically implausible results. In the PETHEMA study, hazard of progression was significantly lower in the VTD group compared to the TD group (HR=0.64, 95% CI 0.44-0.93, p=0.017) (Section 6.5.3 of our original submission).

The anomaly with the TTP data for CR patients (i.e. shorter [worse] TTP with VTD than with TD) was likely to be caused by splitting the overall PETHEMA TTP data into 6 subsets (2 treatment arms [VTD, TD] * 3 response categories [CR, PR and NR]) instead of analysing them by induction treatment arm. A small sample size in each subset resulted in the weakened statistical power; as illustrated in Table 3, none of the hazard ratios by response category is significantly different from 1 and the confidence intervals are wide. This is in contrast with the overall result which demonstrates a significant TTP benefit of VTD compared with TD. The TTP data by treatment arm is more robust and less uncertain than the TTP data by treatment arm and by response, and would have been more suitable to use in the economic model.

Table 3: Hazard ratios of time to progression from the PETHEMA study

Response category	Hazard ratio VTD vs. TD (95% CI; p-value)
All response categories	0.64 (0.44, 0.93; p=0.017)
CR	1.649 (0.621, 4.378; p=0.3156)
PR	1.024 (0.545, 1.921; p=0.9420)
NR	0.590 (0.307, 1.135; p=0.1141)
Note: Hazard ratio estimate is based on a Cox model.	

The issue with the transition probabilities to the second-line therapy highlighted by the ERG has a material impact on the cost-effectiveness of VTD compared with TD. It is clear from Table 4 that the use of TTP data by treatment arm and by response category resulted in the ICER estimates which were biased against VTD.

Table 4: ICER estimates (VTD vs. TD) by how the PETHEMA TTP data are analysed

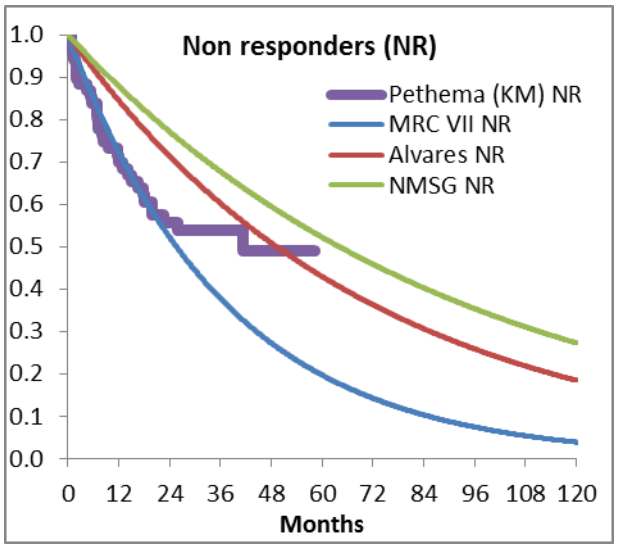
Scenarios	Deterministic ICER	Probabilistic ICER
Scenario 1: TTP analysed by treatment arm and by response category (Base case per the July addendum) Hazard ratio (95% CI; p-value): for CR 1.615 (0.609, 4.282; p=0.3356) for PR 1.007 (0.538, 1.886; p=0.9816) for NR 0.441 (0.230, 0.845; p=0.0136)	£20,468	£23,904
Scenario 2: TTP analysed by treatment arm Hazard ratio (95% CI; p-value): for all response categories 0.628 (0.434, 0.908; p=0.0133)	£17,038	£19,845
Note: Hazard ratio estimate is based on an exponential function. Except for the TTP data, there is no difference in the model structure or the data input between the two scenarios.		

2.1.6 Other comments

There is another limitation to the ACD comments.

- The NMSG 5/94 study does not provide fully relevant data for this STA because it does not report the median OS for the PR and NR groups separately. Moreover, for the NR population, the data from the NMSG 5/94 overestimate the OS from the PETHEMA study to an even greater extent than the Alvares data (Figure 2).

Figure 2: OS curves for the NR group: Kaplan-Meier curve from the PETHEMA study vs. MRC Myeloma VII, Alvares and NMSG 5/94.

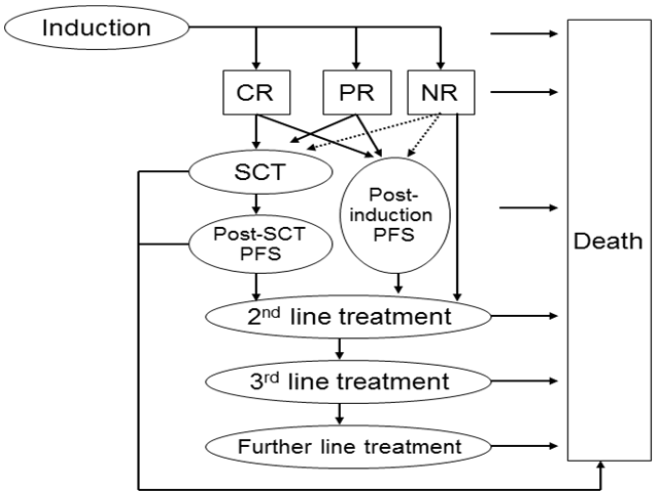


2.2. Amended economic analysis

2.2.1 Model design

The original submission in February 2013 and the addendum submitted in July 2013 were based on an economic model designed to utilise the post-induction response rates (Figure 3). The model was not designed to capture either the post-transplant response rates or the OS benefit arising from a SCT.

Figure 3: Schematic of the original state transition model (from page 117 of the original submission)



As discussed above, it would not be appropriate to apply the post-transplant response rates to those who have not received a SCT. Also, the ERG stated that "it would have been more transparent to distinguish the separate effects on survival of post-induction response and stem cell transplantation" (ACD paragraph 3.29, page 23), a statement we agree with.

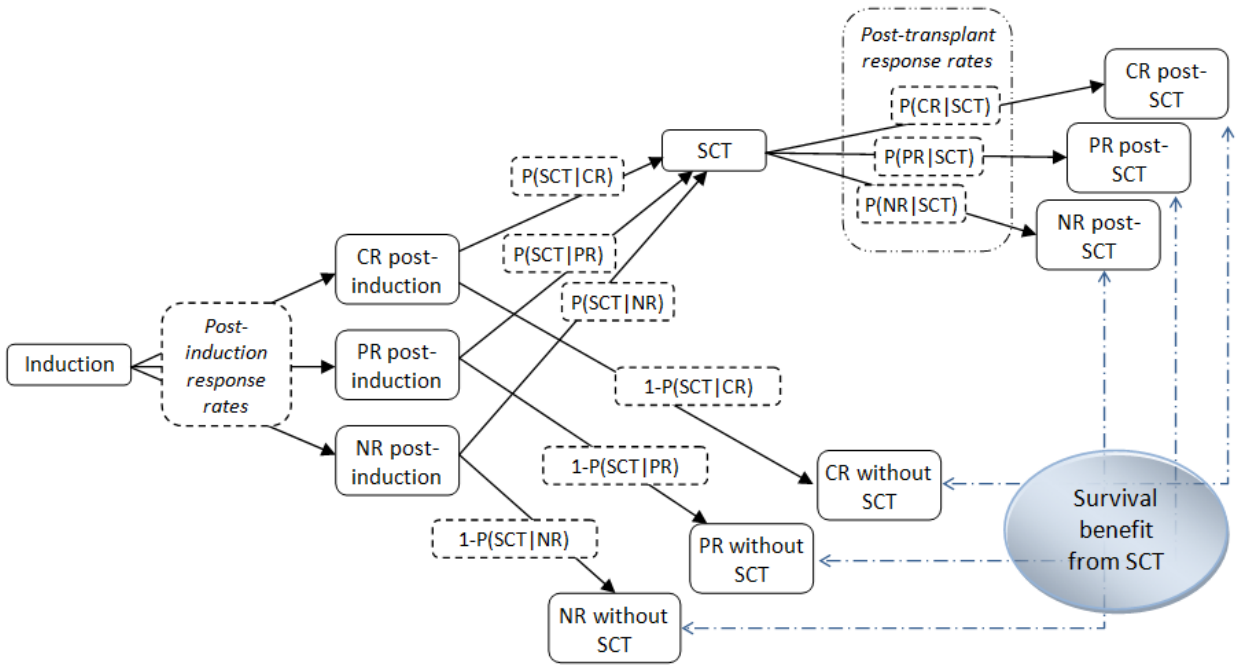
Based on the comments made in the ACD, we have made changes to the economic model so that:

- a survival benefit of SCT is explicitly captured,
- post-induction rates are applied on an ITT basis to all patients who enter the model,
- probabilities of receiving a SCT depend on the post-induction response,
- post-transplant rates are applied to only those who actually receive a SCT.

The amended model structure (up to the post-induction/SCT PFS stage) is illustrated in Figure 4 below.

- SCT rate: P(SCT|CR) is the probability of receiving a SCT given a CR is achieved from the induction therapy. Thus, 1-P(SCT|CR) is the probability of not receiving a SCT after achieving a CR.
- Post-transplant response rate: P(CR|SCT) is the probability of achieving a CR given that a SCT has been performed. Therefore, P(CR|SCT)+ P(PR|SCT)+ P(NR|SCT)=1.

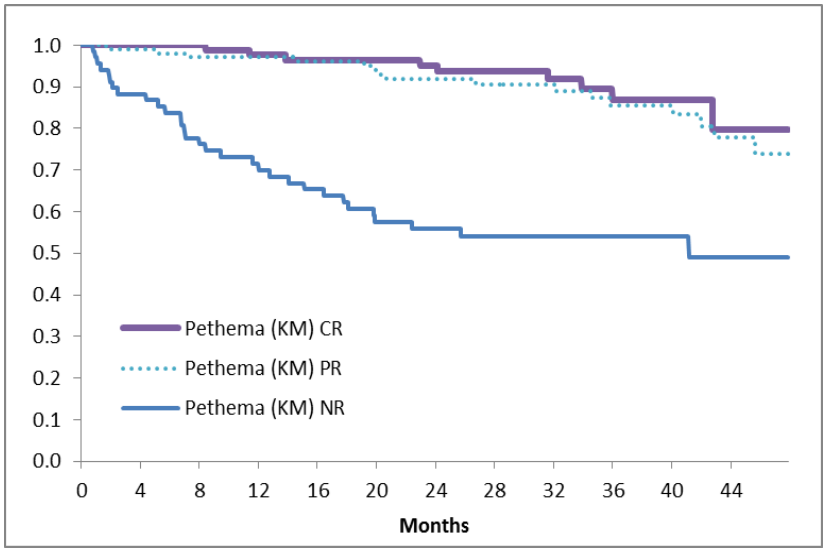
Figure 4: Schematic of the amended state transition model



2.2.2 Input data – long-term overall survival

The OS data from the PETHEMA study have been analysed further to plot the Kaplan-Meier curves by induction response (Figure 5). The hazard ratios between CR and NR, and between PR and NR are both significantly below 1 (CR vs. NR: HR=0.318 [95% CI 0.244-0.413, p<0.0001], PR vs. NR: HR=0.417 [95% CI 0.331-0.525, p<0.0001]). Also, the hazard ratio between CR and PR is significantly below 1 (HR=0.746 [95% CI 0.566-0.982, p<0.0368]).

Figure 5: Kaplan-Meier OS curves from the PETHEMA study by response category

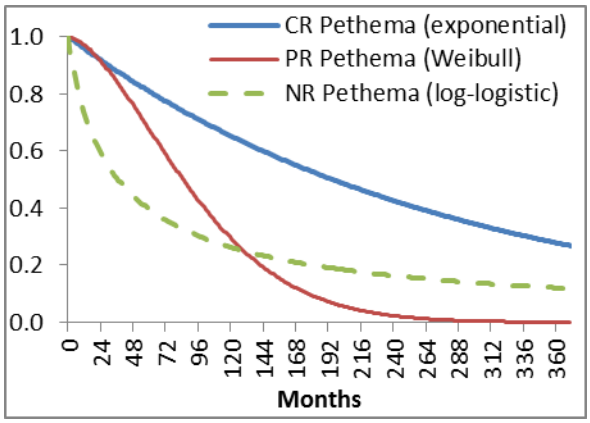


The parametric curves (exponential, Weibull and log-logistics) have been fitted to each Kaplan-Meier curve.

Whilst the median OS has been reached for NR (41 months), the OS data for CR and PR are immature with neither reaching the median, and with more than 70% of patients still alive after 4 years. Therefore, the choice of the OS data source and of the parametric curve must depend on a number of pragmatic considerations such as:

- fit to the PETHEMA Kaplan-Meier curves (visually and statistically [based on AIC]),
- face validity in the long term (e.g. It is not plausible that more than 25% of patients are alive at year 30 as in the case of the exponential function fitted to the Kaplan-Meier curve for CR [Figure 6])
- face validity across three response categories (e.g. It is not plausible that OS curves for PR and NR intersect with each other [Figure 6]).

Figure 6: An example of OS curves lacking face validity



The OS curves by response category are shown in Figure 7. The characteristics of each parametric function and suitability to the economic modelling are detailed in Table 5.

Figure 7: OS curves by response category: PETHEMA Kaplan-Meier curve, parametric curves fitted to the PETHEMA Kaplan-Meier curve and exponential curves based on the long-term OS data

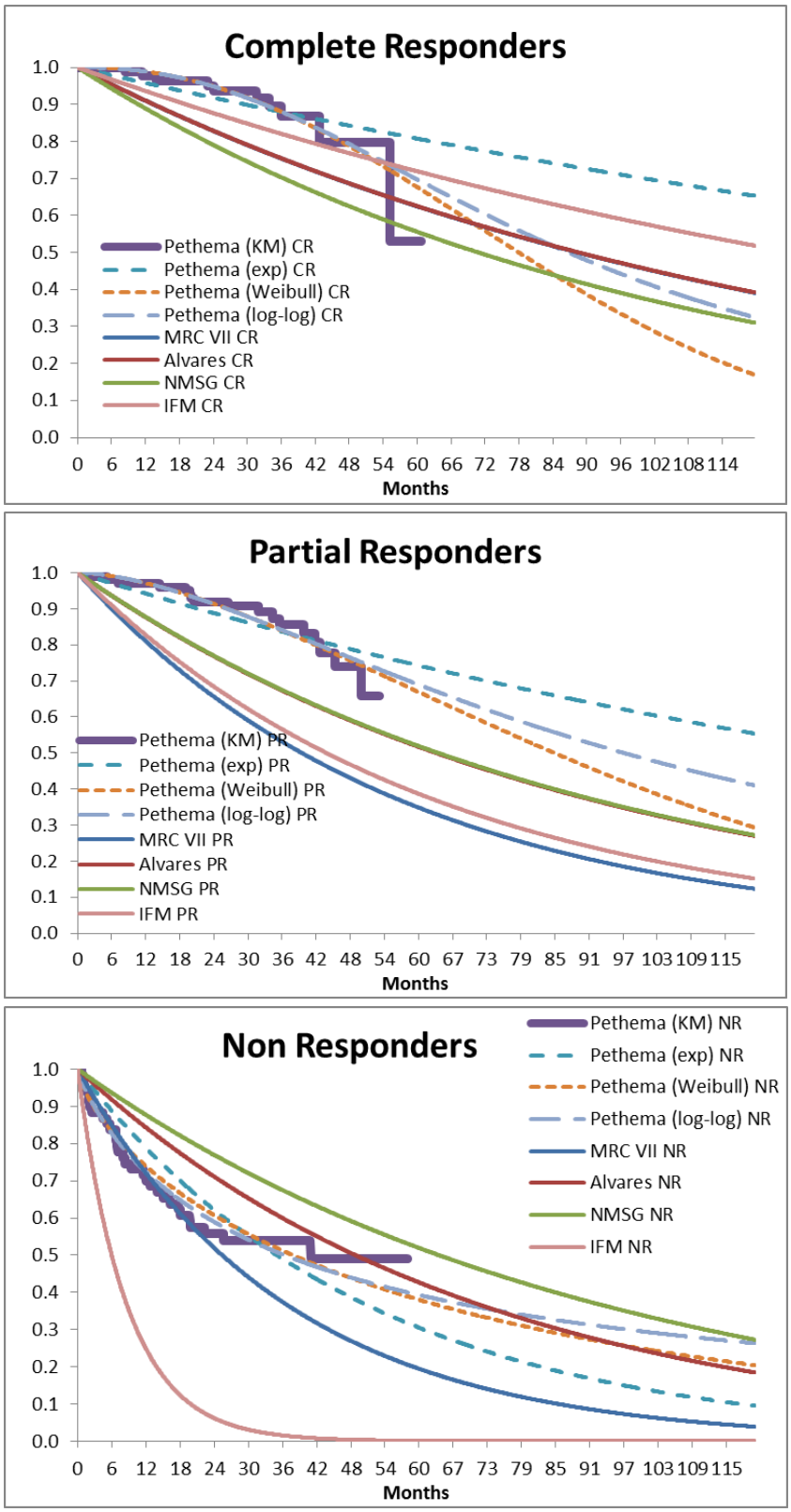


Table 5: Characteristics of each parametric function and suitability to the economic modelling

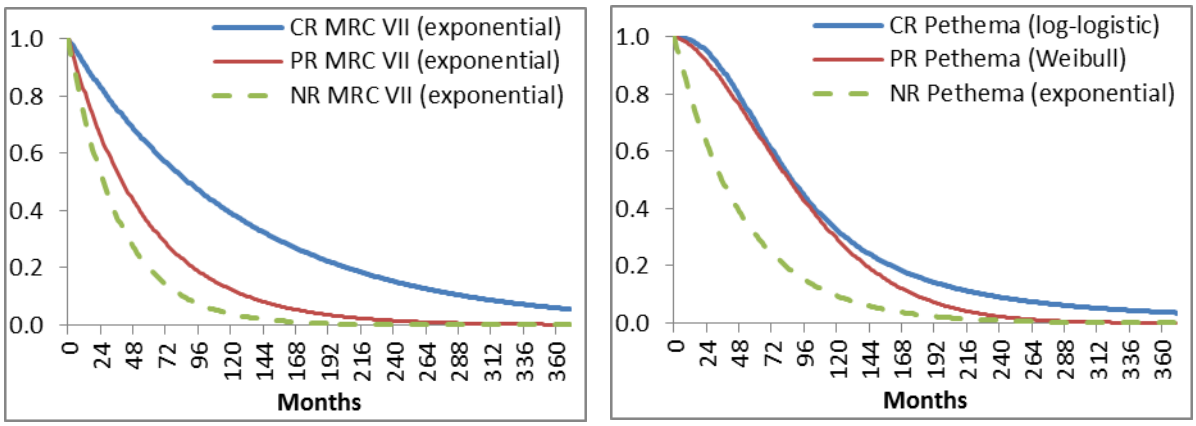
Response category	OS data source / parametric function	Comments
CR	Exponential	AIC=73.1. Highest (worst) AIC for CR. 30-year survival of 28% is unrealistic
	Weibull	AIC=69.1. AIC almost same as for log-logistic (69.3), but it is so steep after year 3 that it crosses over with any of 7 potential PR curves, which is unrealistic
	Log-logistic	AIC= 69.3. AIC almost same as for Weibull (69.1), but does not have the same issue as Weibull. Appears the **best choice** of three parametric curves
PR	Exponential	AIC=115.4. Highest (worst) AIC for PR
	Weibull	AIC=113.3. Lowest (best) AIC for PR, and does not have the same issue as log-logistic (see below). Appears the **best choice** of three parametric curves
	Log-logistic	AIC=114.0. It is so flat after year 6 that it crosses over the log-logistic curve for CR, which is unrealistic
NR	Exponential	AIC=190.5. While the AIC is the highest (worst) for NR, it does not have the face validity issue suffered by Weibull and log-logistic (see below). Appears the **best choice** of three parametric curves
	Weibull	AIC=187.9. It is so flat that it crosses over the Weibull curve for PR, which is unrealistic
	Log-logistic	AIC=185.6. It is so flat that it crosses over the Weibull curve for PR, which is unrealistic
Note: AIC, Akaike information criterion; CR, complete response; K-M, Kaplan-Meier curve; NR, no response (less than PR); PR, partial response		

In summary, the parametric curves fitted to the PETHEMA Kaplan-Meier curves, the best for each response category is:

- Log-logistic for CR,
- Weibull for PR, and,
- Exponential for NR

Figure 8 illustrates the OS curves from the MRC VII study (base case) and from the selected parametric functions. Note that neither chart indicates any obvious face validity issue such as those illustrated in Figure 6.

Figure 8: OS curves based on MRC VII (left) and the parametric curves fitted for PETHEMA Kaplan-Meier curves (right)



2.2.3 Input data – survival benefit of stem cell transplant

Using the pivotal result of the MRC Myeloma VII study (Child et al., 2003), a survival benefit of 11.8 months is assumed for those who have received a SCT over those who have not(1). This is line with a comment made by the ERG clinical expert: "stem cell transplantation offers a survival benefit of 12–18 months compared with no transplant" (ACD paragraph 3.29, page 23).

A survival benefit is now incorporated into the model by deducting 11.8 months from the median survival from the various OS data sources, i.e. (median OS without SCT) = (median OS with SCT) – (11.8 months). This is consistent with how the OS data were reported; the median OS from MRC VII, Alvares and NMSG 4/94, and the 5-year survival from IFM90 are for the SCT group.

2.2.4 Input data – drug administration costs

As noted in the ACD (paragraph 4.11), the original model used the day-patient costs to compute the administration cost associated with VTD, implicitly assuming bortezomib would be administered intravenously. However, since bortezomib can now be administered subcutaneously, the outpatient costs may be more appropriate for the modelling purposes.

In the original submission and in the July 2013 addendum, the following costs associated with bortezomib administration were used:

- SB12Z - Deliver simple parenteral chemotherapy at first attendance (Day case): £203
- SB15Z - Deliver subsequent elements: £284

In the amended model, the following outpatient costs are used instead:

- SB12Z - Deliver simple parenteral chemotherapy at first attendance (Outpatient): £197
- SB15Z - Deliver subsequent elements of a Chemotherapy cycle (Outpatient): £211

2.2.5 Input data – other parameters

Transition probabilities to the second-line therapy are based on the analysis of TTP from the PETHEMA study by treatment arm. Wherever possible, the model input parameters were kept equal to those for the July 2013 addendum to ensure consistency. See Appendix 1 for more details.

2.2.6 Results

The base case results from the deterministic and probabilistic analyses are shown in Table 6 and Table 7, respectively. The ICER estimates are below £23,000/QALY in both analyses.

Table 6: Base case cost-effectiveness results – VTD compared with TD (deterministic)

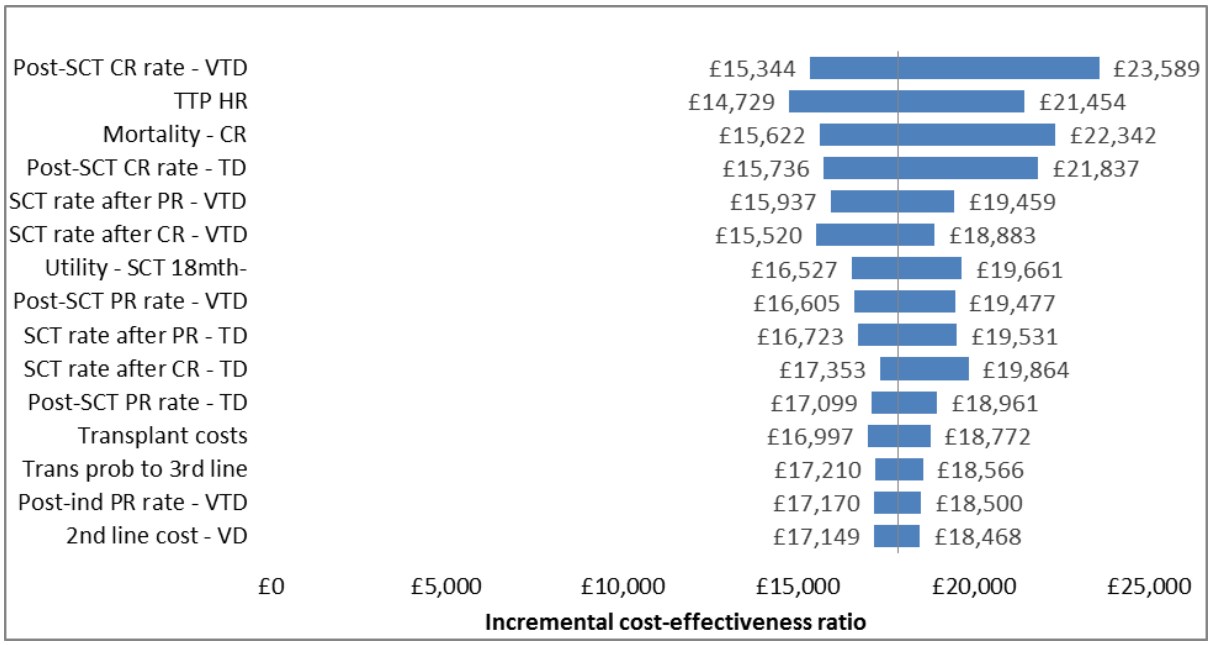
Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£) incremental (QALYs)
TD	£60,168	4.74	3.07	+£17,216	+1.35	+0.96	**£17,841**
VTD	£77,384	6.09	4.04				
ICER, incremental cost-effectiveness ratio; LYG, life years gained; QALY, quality-adjusted life year; TD, thalidomide and dexamethasone; VTD, Velcade® (bortezomib), thalidomide and dexamethasone							

Table 7: Base case cost-effectiveness results – VTD compared with TD (probabilistic)

Technologies	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER (£) incremental (QALYs)	Prob(VTD is cost effective @ £20k/QALY)	Prob(VTD is cost effective @ £30k/QALY)
TD	£60,380	3.44	+£17,109	+0.77	**£22,289**	47.1%	77.8%
VTD	£77,489	4.21					
ICER, incremental cost-effectiveness ratio; QALY, quality-adjusted life year; TD, thalidomide and dexamethasone; VTD, Velcade® (bortezomib), thalidomide and dexamethasone							

The tornado diagram (Figure 9) illustrates that, when each input parameter is varied within a plausible range, the ICER estimates remained below £24,000/QALY in all cases.

Figure 9: Tornado diagram - VTD compared with TD



The results of sensitivity analyses, the ICER estimate by OS data source, are shown in Table 8. The ICERs from both deterministic and probabilistic analyses are well below £30,000/QALY, with the exception of the use of OS data from NMSG 5/94. However, as discussed earlier, the data from NMSG 5/94 have severe limitations such as the identical median OS for the PR and NR groups, and the material overestimation of the OS for the NR group compared to the PETHEMA results.

The ICER based on the parametric curves fitted to the PETHEMA Kaplan-Meier curves (£19,359/QALY) is in line with those based on the historical data sources, indicating the robustness of the ICER estimates. (More detailed results are displayed in Appendix 2).

Table 8: Cost-effectiveness results by OS data source – VTD compared with TD

OS data sources	Deterministic ICER	Probabilistic ICER
Long-term OS data from MRC Myeloma VII for all response categories	£17,841	£22,289
Long-term OS data from IFM90 for all response categories	£15,883	£15,184
Long-term OS data from Alvares for all response categories	£22,696	£22,952
Long-term OS data from NMSG 5/94 for all response categories	£39,618	£39,881
Parametric curves fitted to PETHEMA K-M curve: CR – log-logistic, PR – Weibull, NR – exponential	£19,359	£19,668
Note: CR, complete response; K-M, Kaplan-Meier curve; NR, no response (less than PR); PR, partial response		

In addition, threshold analyses have been conducted to simulate the impact of cyclophosphamide on the cost-effectiveness of VTD. As shown in Table 9, addition of cyclophosphamide would need to materially enhance the post-transplant response rates for TD in order for the ICER to reach even the £20k/QALY-mark.

Table 9: Threshold analysis – How much would the post-transplant CR rate for TD have to rise before the ICER reaches £20k/QALY?

Description of scenarios	Post-SCT CR rate for TD would have to rise from 56.4% to:
Long-term OS data from MRC Myeloma VII for all response categories	67.4% (+11.0% point)
Long-term OS data from MRC VII for CR & NR, and Alvares for PR	85.2% (+28.8% point)
Parametric curves fitted to PETHEMA K-M curve: CR – log-logistic, PR – Weibull, NR – exponential	68.1% (+11.7% point)
NB – The post-transplant response rates for VTD are: CR 68.6%, PR 27.6%, NR 3.8% Note: CR, complete response; K-M, Kaplan-Meier curve; NR, no response (less than PR); PR, partial response	

Both the amended economic analyses and the base case from the July 2013 addendum (cost per QALY gained = £20,468) consistently demonstrate that VTD is a cost-effective treatment option compared with TD. Furthermore, threshold analyses illustrate that the addition of cyclophosphamide to TD would have to add a clinically implausible level of efficacy before VTD ceases to be cost-effective compared with CTD. Therefore, the cost-effectiveness of VTD is robust whether it is compared with TD or with CTD.

3. Bortezomib in combination with dexamethasone (VD): Overall population

In this section, VD is compared with CTD as requested in the NICE ACD:

"1.3 The Committee requests the following further information from the manufacturer:

An indirect comparison between bortezomib in combination with dexamethasone, compared with cyclophosphamide in combination with thalidomide and dexamethasone […].

In the absence of a network to facilitate a robust comparison, this should be a careful comparison using single arms from relevant clinical trials, taking into account differences in trial design and baseline characteristics."

3.1 Model structure and input data

3.1.1 Model structure

The amended model discussed in Section 2.2 was used to analyse the cost-effectiveness of VD compared with CTD.

3.1.2 Input data – response rates and SCT rates

As acknowledged in the ACD, there is no RCT evidence that compares VD and CTD directly. Therefore, the matching-adjusted indirect comparison (MAIC) was performed to adjust the response rates (post-induction and post-transplant) and the SCT rates for VD using the patient characteristics for the VD arm of the IFM study and the CTD arm of the MRC Myeloma IX study.

The MAIC was conducted using the method as suggested by Signorovitch et al(2;3). Individual patient data from the IFM-2005 study were matched to the aggregate data from the MRC IX study. The patient inclusion/exclusion criteria and baseline characteristics from the CTD arm in MRC IX formed the basis of a new population sample in which the VD arm was matched by re-weighting (i.e. adjusting) patients in the IFM-2005 study by their inverse odds of being enrolled in the MRC IX study. Further details of the MAIC are available in Appendix 3.

The MAIC created a new set of post-induction response rates, the SCT rates and the post-transplant response rates for the VD arm (Table 10).

Table 10: Response rates and SCT rates for VD before and after the MAIC

	Before MAIC (as reported in CSR)	After MAIC
Post-induction response rates (ITT)		
CR (CR+nCR+VGPR)	40.8%	45.3%
PR	36.7%	35.3%
NR (less than PR)	22.5%	19.5%
SCT rate		
	88.3%	89.1%
Post-transplant response rates (for those who received a SCT)		
CR (CR+nCR+VGPR)	73.6%	73.2%
PR	19.8%	20.4%
NR (less than PR)	6.6%	6.4%

Note: CR, complete response; CSR, clinical study report; ITT, intention-to-treat; MAIC, matching-adjusted indirect comparison ; nCR, near complete response; NR, non response; PR, partial response; SCT, stem cell transplant; VGPR, very good partial response

Consistent with our original submission, the CR category in the economic model consists of CR, nCR and VGPR from the clinical trials. We believe this is an appropriate way to categorise the various levels of response from the IFM-2005 and MRC IX studies.

Table 11: Comparison of the EBMT criteria (1998) and the IMWG criteria (2006)

Response	*EBMT criteria*	*IMWG criteria*
sCR	NA	CR as defined below plus normal FLC ratio and absence of clonal cells in bone marrow[a] by immunohistochemistry or immunofluorescence[b]
CR	No M-protein detected in serum or urine by immunofixation for a minimum of 6 weeks and fewer than 5% plasma cells in the bone marrow; No increase in size or number of lytic bone lesions; Disappearance of soft tissue plasmacytomas	Negative immunofixation on the serum and urine and disappearance of any soft tissue plasmacytomas and <5% plasma cells in bone marrow[a]
VGPR	NA	Serum and urine M-protein detectable by immunofixation but not on electrophoresis or ⩾90% reduction in serum M-protein plus urine M-protein level <100 mg/24 h
PR	>50% reduction in serum M-protein level and/or ⩾90% reduction in urine free light chain excretion or reduction to <200 mg/24 h for 6 weeks[c]	⩾50% reduction of serum M-protein and reduction in 24 h urinary M-protein by ⩾90% or to <200 mg/24 h If the serum and urine M-protein are unmeasurable,[d] a ⩾50% decrease in the difference between involved and uninvolved FLC levels is required in place of the M-protein criteria If serum and urine M-protein are not measurable, and serum free light assay is also not measureable, ⩾50% reduction in plasma cells is required in place of M-protein, provided baseline bone marrow plasma cell percentage was ⩾30% In addition to the above listed criteria, if present at baseline, a ⩾50% reduction in the size of soft tissue plasmacytomas is also required

VGPR is a relatively new concept formally introduced in the International Myeloma Working Group (IMWG) response criteria that "were developed in 2006 and have been commonly used by a larger number of physicians and have substituted for the EBMT criteria"(4) . In contrast, the EBMT criteria did not recognise VGPR (Table 11).

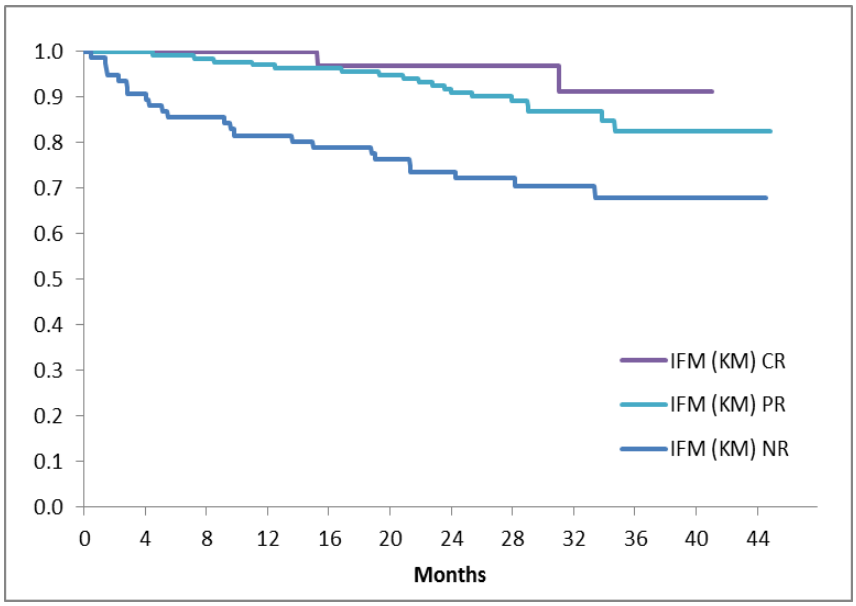
For the purpose of this economic evaluation, the pertinent question is whether long-term OS for VGPR (as defined in the IMWG criteria) is closer to that for CR, or to that for PR. However, long-term OS data specifically for VGPR do not exist yet, as VGPR was introduced relatively recently. Nevertheless, it has been shown that the achievement of response less than a VGPR both after induction therapy and after SCT is an adverse prognostic factor for PFS, particularly the achievement of response < VGPR after induction(5). Various analyses reported in recent publications have tended to group CR/VGPR together. In one such publication, it was cited that CR/VGPR remains a simple and powerful prognostic indicators in the context of HDT/ASCT for multiple myeloma(6). Moreover, from a practical point of view in the UK clinical setting, response of some patients may be categorised as VGPR instead of CR since immunofixation is not performed in some centres in the UK (a negative immunofixation result is required for confirmation of CR).

For this reason we have also chosen to group CR/VGPR together in line with the prevailing view that the attainment of at least a VGPR has a significant prognostic value.

3.1.3 Input data – long-term overall survival

The OS data from the IFM-2005 study have been analysed further to plot the Kaplan-Meier curves by induction response (Figure 10). The hazard ratios between CR and NR, and between PR and NR are both significantly below 1 (CR vs. NR: HR=0.204 [95% CI 0.077-0.232, p=0.0013], PR vs. NR: HR=0.445 [95% CI 0.232-0.854, p<0.0149]). However, the hazard ratio between CR and PR is not significantly below 1 (HR=0.456 [95% CI 0.166-1.258, p=0.1294]).

Figure 10: Kaplan-Meier OS curves from the IFM-2005 study



Similarly to the VTD-TD analysis discussed above, the choice of the OS data source and of the parametric curve must be made on pragmatic considerations. The OS curves by response category are shown in Figure 11. The characteristics and suitability of each parametric function are detailed in Table 12.

Figure 11: OS curves by response category: Kaplan-Meier curves from the IFM-2005 study, parametric curves fitted to the IFM-2005 Kaplan-Meier curve, and the long-term OS data from historical studies

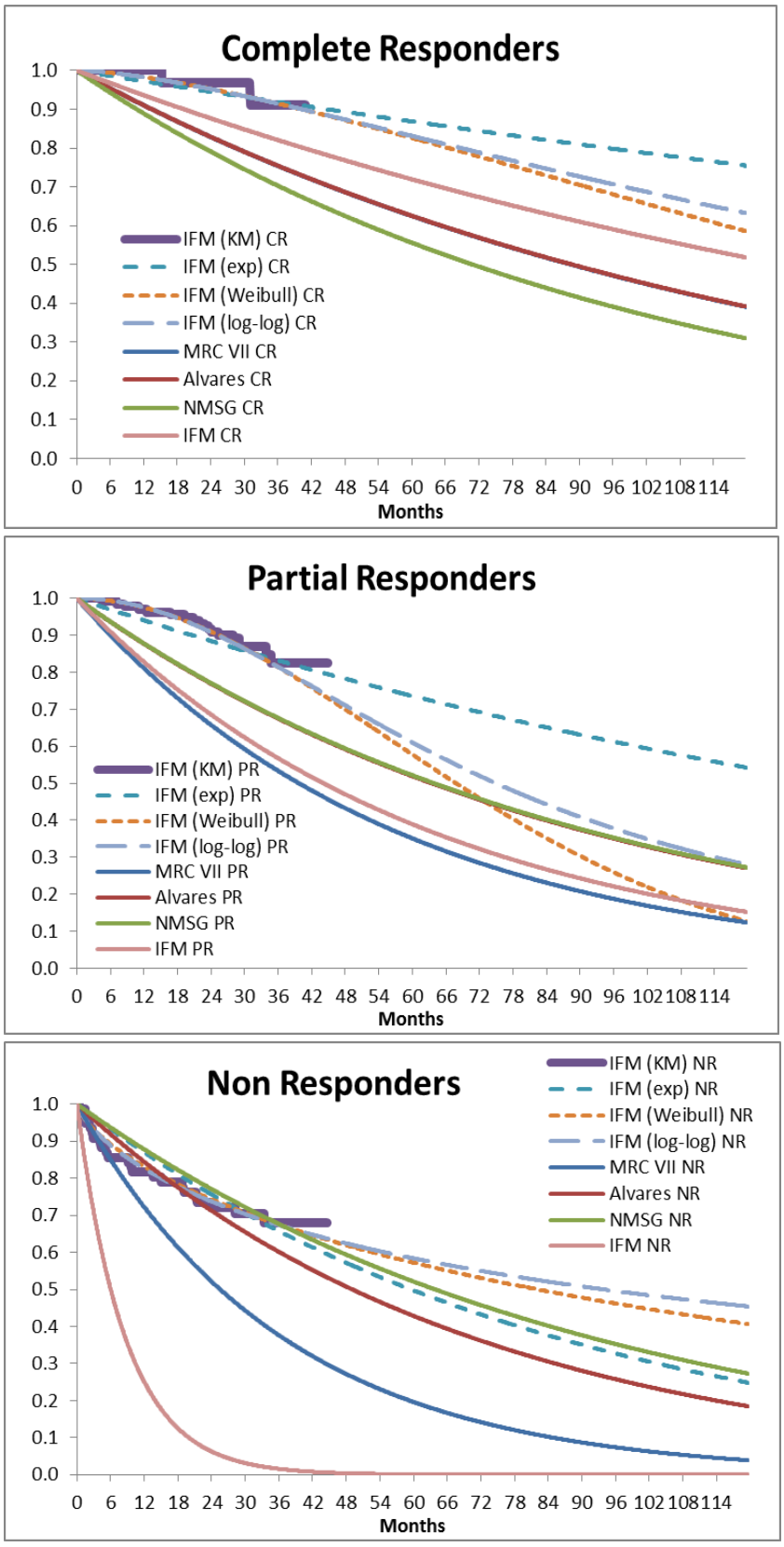


Table 12: Characteristics of each parametric function and suitability to the economic modelling

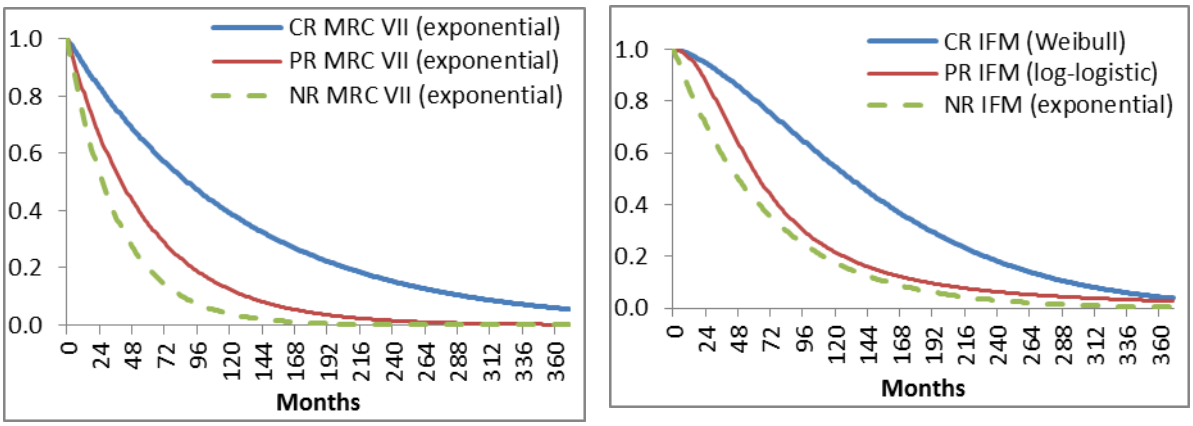
Response category	OS data source / parametric function	Comments
CR	Exponential	AIC=46.4. 30-year survival of 49% is unrealistic
	Weibull	AIC=47.7. Its AIC is almost same as for others and it does not have the same face validity issue as exponential and log-logistic. Appears the **best choice** of three parametric curves
	Log-logistic	AIC= 47.6. 30-year survival of 24% is unrealistic
PR	Exponential	AIC=103.5. Highest (worst) AIC for PR. 30-year survival of 15% may be unrealistic
	Weibull	AIC=99.9. It is so steep after year 3 that it crosses over with any of 3 potential parametric NR curves, which is unrealistic
	Log-logistic	AIC=99.8. It does not have the same face validity issue as exponential or Weibull. Appears the **best choice** of three parametric curves
NR	Exponential	AIC=167.8. While the AIC is the highest (worst) for NR, it does not have the face validity issue suffered by Weibull and log-logistic (see below). Appears the **best choice** of three parametric curves
	Weibull	AIC=165.7. It is so flat that it crosses over the log-logistic curve for PR, which is unrealistic
	Log-logistic	AIC=165.1. It is so flat that it crosses over the log-logistic curve for PR, which is unrealistic
Note: AIC, Akaike information criterion; CR, complete response; K-M, Kaplan-Meier curve; NR, no response (less than PR); PR, partial response		

In summary, among the parametric curves fitted to the IFM-2005 Kaplan-Meier curves, the best for each response category is:

- Weibull for CR,
- Log-logistic for PR, and,
- Exponential for NR

Figure 12 illustrates the OS curves for the base case (MRC VII) and parametric curves. Note that neither chart indicates any obvious face validity issues.

Figure 12: OS curves based on MRC VII (left) and the parametric curves fitted for IFM-2005 Kaplan-Meier curves (right)



3.2 Results

The base case results from the deterministic and the probabilistic analyses are shown in Table 13 and Table 14, respectively. The ICER estimates are below £23,000/QALY from both analyses.

Table 13: Base case cost-effectiveness results – VD compared with CTD (deterministic)

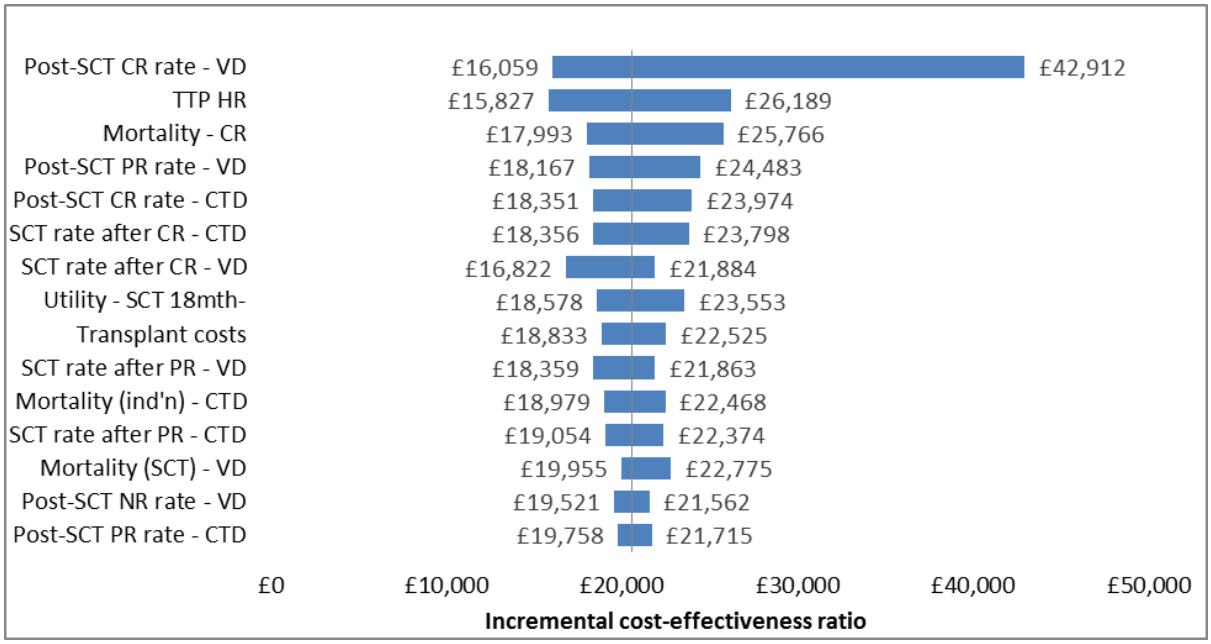
Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£) incremental (QALYs)
CTD	£58,037	5.61	3.70	+£12,225	+0.86	+0.59	**£20,588**
VD	£70,261	6.47	4.29				
ICER, incremental cost-effectiveness ratio; LYG, life years gained; QALY, quality-adjusted life year; CTD, cyclophosphamide, thalidomide and dexamethasone; VD, Velcade® (bortezomib) and dexamethasone							

Table 14: Base case cost-effectiveness results – VD compared with CTD (probabilistic)

Technologies	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER (£) incremental (QALYs)	Prob(VD is cost effective @ £20k/QALY)	Prob(VD is cost effective @ £30k/QALY)
CTD	£58,256	3.93	+£12,237	+0.55	**£22,305**	36.3%	79.0%
VD	£70,494	4.48					
ICER, incremental cost-effectiveness ratio; LYG, life years gained; QALY, quality-adjusted life year; TD, thalidomide and dexamethasone; VTD, Velcade® (bortezomib), thalidomide and dexamethasone							

The tornado diagram (Figure 13) illustrates that, when each input parameter is varied within a plausible range, the ICER estimates remained below £27,000/QALY in all cases with the exception of reducing the post-transplant CR rate for VD to 64.5%, well below 73.2% in the base case (Table 10).

Figure 13: Tornado diagram - VD compared with CTD



The results of sensitivity analyses, the ICER estimate by OS data source, are shown in Table 15. The ICERs from both deterministic and probabilistic analyses are well below £30,000/QALY, with the exception of when the OS data from NMSG 5/94 are used. However, as discussed earlier, the data from NMSG 5/94 have severe limitations such as the identical median OS for the PR and NR groups, which is clinically implausible.

Table 15: Cost-effectiveness results by OS data source – VD compared with CTD

Description of scenarios	Deterministic ICER	Probabilistic ICER
Long-term OS data from MRC Myeloma VII for all response categories	£20,588	£22,305
Long-term OS data from IFM90 for all response categories	£17,866	£16,438
Long-term OS data from Alvares for all response categories	£24,267	£23,816
Long-term OS data from NMSG 5/94 for all response categories	£33,435	£33,107
Parametric curves fitted to PETHEMA K-M curve: CR – Weibull, PR – log-logistic, NR – exponential	£18,864	£19,057
Note: CR, complete response; K-M, Kaplan-Meier curve; NR, no response (less than PR); PR, partial response		

As discussed above, in the context of this STA, VGPR should be categorised under CR in the economic model, i.e. CR in the model = CR+nCR+VGPR in the trials. Categorising VGPR under PR implies those who had a ≥90% reduction in serum M-protein (VGPR) would have a similar long-term OS to those who only had a >50% reduction (PR). However, in order to test the robustness of ICER estimates, the cost-effectiveness of VD compared with CTD has been assessed under this unlikely scenario.

As Table 16 shows, most ICER estimates remain below £30,000/QALY, even under this conservative assumption.

Table 16: Cost-effectiveness results by OS data source – VD compared with CTD (if CR in the model is defined as CR+nCR)

Description of scenarios	Deterministic ICER	Probabilistic ICER
Long-term OS data from MRC Myeloma VII for all response categories	£28,281	£32,640
Long-term OS data from IFM90 for all response categories	£25,070	£23,227
Long-term OS data from Alvares for all response categories	£28,121	£27,297
Long-term OS data from NMSG 5/94 for all response categories	£34,750	£34,167
Long-term OS data from MRC VII for CR, and Alvares for PR & NR	£28,135	£27,428
Parametric curves fitted to PETHEMA K-M curve: CR – Weibull, PR – log-logistic, NR – exponential	£24,589	£23,425
Note: CR, complete response; K-M, Kaplan-Meier curve; NR, no response (less than PR); PR, partial response		

As requested in the ACD, we have conducted a cost-effectiveness analysis of VD compared with CTD. Since it is heavily reliant on a comparison of single arms from separate heterogeneous studies (VD from IFM-2005 and CTD from MRC Myeloma IX), results are not as robust as those based on a well-designed RCT, such as the VTD-TD analysis presented above.

Whilst such limitations must be noted, the results shown here clearly indicate that VD is a cost-effective treatment option compared with CTD.

4. Bortezomib in combination with dexamethasone (VD): when CTD is not suitable

In this section, VD is compared with a regimen other than CTD as requested in the ACD:

"1.3 The Committee requests the following further information from the manufacturer:

An indirect comparison between bortezomib in combination with dexamethasone […] compared with an alternative comparator in circumstances when cyclophosphamide in combination with thalidomide and dexamethasone is not suitable."

In order to elicit clinical expert views on the clinical questions relevant to this request, advisory board meetings took place on 20th and 21st November 2013. The details of the meetings can be found in Appendix 6.

4.1 When is CTD not suitable as an induction therapy?

According to the clinical experts, CTD as an induction therapy is not suitable when a transplant-eligible MM patient:

- requires a rapid response due to renal failure or pending renal failure,
- has either prior or active thrombosis,
- is pregnant or is not willing to consent to pregnancy prevention or take part in a risk management programme,
- has existing peripheral neuropathy,
- is likely to have an allergic reaction or rash, or,
- may not adhere to dosing of thalidomide at home.

4.2 Which regimen is used if CTD is not suitable? What is the most relevant comparator?

The clinical experts commented that they would wish to use a bortezomib-based regimen (such as VD) when CTD is not suitable. Regarding the regimens without bortezomib, the following options were suggested:

- cyclophosphamide and dexamethasone (CD),
- vincristine, doxorubicin and dexamethasone (VAD), and,
- idarubicin and dexamethasone (Z-Dex).

Among these three options, there is only RCT evidence against VD for VAD (i.e., from the IFM-2005 trial). Therefore, the economic analysis discussed below focuses on VD compared with VAD.

However, it can be argued that VAD may be a reasonable proxy for CD as follows:

- VAD and CVAD can be assumed clinically equivalent (as agreed by the ERG clinical expert [ERG report page 23]), and,
- removing 'VA' (vincristine and doxorubicin) from CVAD would create CD, which is likely to have lower efficacy as well as lower toxicity compared with CVAD (vincristine is associated with peripheral neuropathy, and doxorubicin is cardiotoxic).

4.3 Model structure and input data

4.3.1 Model structure

The amended model discussed in Section 2.2 was used to analyse the cost-effectiveness of VD compared with VAD.

4.3.2 Input data

The input data are consistent with those in the March 2013 addendum, except for the parameters introduced in the amended model such as the OS benefit from a SCT.

4.4 Results

The base case results from the deterministic and the probabilistic analyses are shown in Table 17 and Table 18, respectively. The ICER estimates are below £21,000/QALY from both analyses.

Table 17: Base case cost-effectiveness results – VD compared with VAD (deterministic)

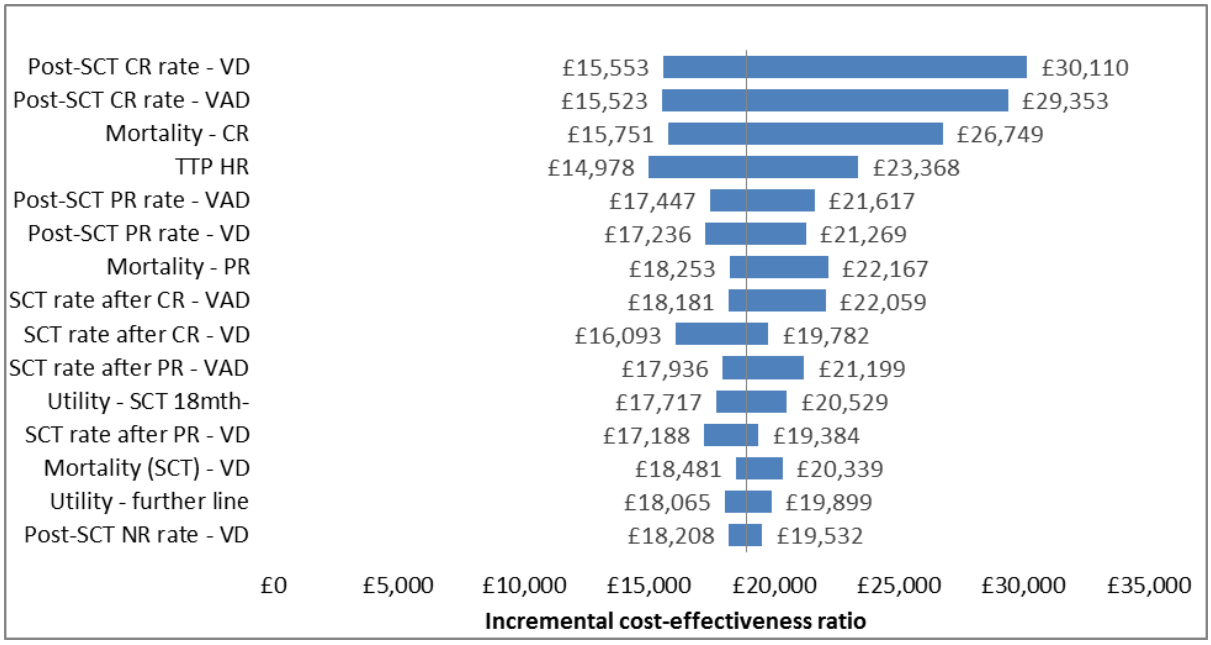
Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£) incremental (QALYs)
VAD	£56,455	5.35	3.52	+£14,000	+1.07	+0.74	**£18,914**
VD	£70,455	6.42	4.26				
ICER, incremental cost-effectiveness ratio; LYG, life years gained; QALY, quality-adjusted life year; VAD, vincristine, doxorubicin and dexamethasone; VD, Velcade® (bortezomib) and dexamethasone							

Table 18: Base case cost-effectiveness results – VD compared with VAD (probabilistic)

Technologies	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER (£) incremental (QALYs)	Prob(VD is cost effective @ £20k/QALY)	Prob(VD is cost effective @ £30k/QALY)
VAD	£56,730	3.76	+£13,846	+0.69	**£20,096**	49.2%	86.9%
VD	£70,576	4.45					
ICER, incremental cost-effectiveness ratio; QALY, quality-adjusted life year; VAD, vincristine, doxorubicin and dexamethasone; VD, Velcade® (bortezomib) and dexamethasone							

The tornado diagram (Figure 14) illustrates that when each input parameter is varied within a plausible range, the ICER estimates remained below approximately £30,000/QALY in all cases.

Figure 14: Tornado diagram - VD compared with VAD



The results of sensitivity analyses, the ICER estimate by OS data source, are shown in Table 19. Consistent with the results for VTD vs. TD/CTD and VD vs. CTD, the ICERs from both deterministic and probabilistic analyses are well below £30,000/QALY, except when the OS data from NMSG 5/94 are used. However, as discussed earlier, the data from NMSG 5/94 have severe limitations.

Table 19: Cost-effectiveness results – VD compared with VAD

Description of scenarios	Deterministic ICER	Probabilistic ICER
Long-term OS data from MRC Myeloma VII for all response categories	£18,914	£20,096
Long-term OS data from IFM90 for all response categories	£15,393	£14,392
Long-term OS data from Alvares for all response categories	£25,575	£25,494
Long-term OS data from NMSG 5/94 for all response categories	£42,811	£42,528
Long-term OS data from MRC VII for CR, and Alvares for PR & NR	£25,650	£25,189
Parametric curves fitted to PETHEMA K-M curve: CR – Weibull, PR – log-logistic, NR – exponential	£18,489	£18,761
Note: CR, complete response; K-M, Kaplan-Meier curve; NR, no response (less than PR); PR, partial response		

As aforementioned, VGPR should be categorised under CR in the economic model. However, in order to test the robustness of ICER estimates, the cost-effectiveness of VD compared with VAD has been assessed under an alternative, unlikely scenario where CR in the model is defined as CR+nCR in the trials. As Table 20 shows, most ICER estimates remain below £30,000/QALY even under this conservative assumption.

Table 20: Cost-effectiveness results – VD compared with VAD (if CR in the model is defined as CR+nCR)

Description of scenarios	Deterministic ICER	Probabilistic ICER
Long-term OS data from MRC Myeloma VII for all response categories	£26,146	£28,398
Long-term OS data from IFM90 for all response categories	£20,910	£19,123
Long-term OS data from Alvares for all response categories	£32,207	£31,224
Long-term OS data from NMSG 5/94 for all response categories	£48,425	£47,533
Long-term OS data from MRC VII for CR, and Alvares for PR & NR	£32,274	£31,136
Parametric curves fitted to PETHEMA K-M curve: CR – Weibull, PR – log-logistic, NR – exponential	£28,411	£23,634
Note: CR, complete response; K-M, Kaplan-Meier curve; NR, no response (less than PR); PR, partial response		

The results of the original cost-effectiveness analysis for VD compared with VAD have already been reported in the March 2013 addendum; the ICER estimate was £14,232/QALY, which is in line with the ICER estimates from the amended model shown above.

In summary, VD is a cost-effective treatment option compared with VAD when CTD is not suitable.

5. Conclusion

Bortezomib is the only licensed agent available during the induction phase of the multiple myeloma (MM) treatment pathway. As induction therapy occurs shortly after diagnosis, the trial data to support this indication are somewhat immature due to the prolonged survival in this early patient population. In addition, the most commonly used treatment in the UK (CTD) is unlicensed and suffers from a limited evidence base as this regimen is not widely used outside of the UK. This combination of immature trial data and comparator cause some problems from an HTA perspective and we hope NICE can work with us to try to find a pragmatic solution that will enable patients to access the only licensed treatment for this patient population.

We also hope that the current emphasis on the ERG's exploratory analysis of the cost-effectiveness of VTD and VD against CTD, comparing single arms derived from heterogeneous trials and different economic models, is lessened, and that the robust, comparative evidence of the PETHEMA and IFM-2005 trials is taken into account. This analysis is not in line with NICE's stringent evidence requirements, and should not carry any weight towards the appraisal committee's final decision.

We interpret the ACD as being principally concerned that the most plausible ICER was likely to be substantially higher than £39,000/QALY. Whilst we accept there is uncertainty around our submitted cost per QALY gained of £20,468 (VTD compared with TD), we believe our base case is highly robust, and represents the best estimate of the cost-effectiveness of bortezomib in the induction setting.

References

Reference List

(1) Child JA, Morgan GJ, Davies FE, Owen RG, Bell SE, Hawkins K, et al. High-dose chemotherapy with hematopoietic stem-cell rescue for multiple myeloma. N Engl J Med 2003 May 8;348(19):1875-83.

(2) Signorovitch J, Erder MH, Xie J, Sikirica V, Lu M, Hodgkins PS, et al. Comparative effectiveness research using matching-adjusted indirect comparison: an application to treatment with guanfacine extended release or atomoxetine in children with attention-deficit/hyperactivity disorder and comorbid oppositional defiant disorder. Pharmacoepidemiol Drug Saf 2012 May;21 Suppl 2:130-7.

(3) 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.

(4) Palumbo A, Sezer O, Kyle R, Miguel JS, Orlowski RZ, Moreau P, et al. International Myeloma Working Group guidelines for the management of multiple myeloma patients ineligible for standard high-dose chemotherapy with autologous stem cell transplantation. Leukemia 2009 Oct;23(10):1716-30.

(5) Moreau P, Attal M, Pegourie B, Planche L, Hulin C, Facon T, et al. Achievement of VGPR to induction therapy is an important prognostic factor for longer PFS in the IFM 2005-01 trial. Blood 2011;117(11):3041-4.

(6) Harousseau JL, Avet-Loiseau H, Attal M, Charbonnel C, Garban F, Hulin C, et al. Achievement of at least very good partial response is a simple and robust prognostic factor in patients with multiple myeloma treated with high-dose therapy: long-term analysis of the IFM 99-02 and 99-04 Trials. J Clin Oncol 2009 Dec 1;27(34):5720-6.

(7) Morgan GJ, Davies FE, Gregory WM, Bell SE, Szubert AJ, Navarro CN, et al. Cyclophosphamide, thalidomide, and dexamethasone as induction therapy for newly diagnosed multiple myeloma patients destined for autologous stem-cell transplantation: MRC Myeloma IX randomized trial results. Haematologica 2012 Mar 1;97(3):442-50.

Appendices

Appendix 1 Input parameters for the VTD vs. TD cost-effectiveness analysis

Post-induction response

Pethema CSR Table 25 (page 63)

	TD (N=127)		VTD (N=130)	
CR (CR+nCR)	22	17.3%	64	49.2%
PR	56	44.1%	46	35.4%
NR (less than PR)	49	38.6%	20	15.4%
TOTAL	127	100.0%	130	100.0%

SCT rate

Pethema CSR Table 52 (page 82)

	TD (N=127)		VTD (N=130)	
Transplant	78	61.4%	105	80.8%
No transplant	49	38.6%	25	19.2%
TOTAL	127	100.0%	130	100.0%

SCT rate by post-induction response

Post-hoc analysis (clarification stage)

	TD (N=127)			
	SCT		No SCT	
CR (CR+nCR)	21	95.5%	1	4.5%
PR	50	89.3%	6	10.7%
NR (less than PR)	7	14.3%	42	85.7%
TOTAL	78	61.4%	49	38.6%

	VTD (N=130)			
	SCT		No SCT	
CR (CR+nCR)	62	96.9%	2	3.1%
PR	38	82.6%	8	17.4%
NR (less than PR)	5	25.0%	15	75.0%
TOTAL	105	80.8%	25	19.2%

Post-transplant response (SCT population only)

Pethema CSR Table 31 (page 67)

	TD (N=78)		VTD (N=105)	
CR (CR+nCR)	44	56.4%	72	68.6%
PR	28	35.9%	29	27.6%
NR (less than PR)	6	7.7%	4	3.8%
TOTAL	78	100.0%	105	100.0%

Median OS by response category

(months)	CR (CR+nCR)	PR	NR (less than PR)
MRC VII	88.6	39.8	25.6
IFM90	126.6	44.2	6.0
Alvares	88.8	63.6	49.2
NMSG 5/94	71.0	64.4	64.4

Parameters for parametric curves fitted to the Kaplan-Meier curves from PETHEMA

	CR (CR+nCR)	PR	NR (less than PR)
Exponential function			
Scale	0.00012	0.00018	0.00097
Shape	1.0000	1.0000	1.0000
Weibull function			
Scale	0.0000000	0.0000025	0.0055024
Shape	2.1954	1.6279	0.7233
Log-logistic function			
Scale	-17.5653	-13.3030	-5.7598
Shape	2.2712	1.6926	0.8690

Appendix 2 Detailed results of the VTD vs. TD cost-effectiveness analysis

Deterministic analysis

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£) incremental (QALYs)
Long-term OS data from MRC Myeloma VII for all response categories							
TD	£60,168	4.74	3.07	+£17,216	+1.35	+0.96	**£17,841**
VTD	£77,384	6.09	4.04				
Long-term OS data from IFM90 for all response categories							
TD	£55,231	4.90	3.19	+£22,750	+2.07	+1.43	**£15,883**
VTD	£77,981	6.96	4.62				
Long-term OS data from Alvares for all response categories							
TD	£65,138	6.03	3.91	+£15,582	+0.91	+0.69	**£22,696**
VTD	£80,719	6.94	4.60				
Long-term OS data from NMSG 5/94 for all response categories							
TD	£65,005	6.11	3.96	+£14,270	+0.41	+0.36	**£39,618**
VTD	£79,275	6.52	4.32				
Parametric curves fitted to PETHEMA K-M curve: CR - log-logistic, PR - Weibull, NR - exponential							
TD	£66,313	5.61	3.65	+£16,279	+1.14	+0.84	**£19,359**
VTD	£82,592	6.74	4.49				
ICER, incremental cost-effectiveness ratio; LYG, life years gained; QALY, quality-adjusted life year; TD, thalidomide and dexamethasone; VTD, Velcade® (bortezomib), thalidomide and dexamethasone							

Probabilistic analysis

Technologies	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER (£) incremental (QALYs)	Prob(VTD is cost effective @ £20k/QALY)	Prob(VTD is cost effective @ £30k/QALY)
Long-term OS data from MRC Myeloma VII for all response categories							
TD	£60,380	3.44	+£17,109	+0.77	**£22,289**	47.1%	77.8%
VTD	£77,489	4.21					
Long-term OS data from IFM90 for all response categories							
TD	£55,186	3.39	+£23,297	+1.53	**£15,184**	82.3%	99.0%
VTD	£78,483	4.93					
Long-term OS data from Alvares for all response categories							
TD	£64,967	3.96	+£15,633	+0.68	**£22,952**	28.1%	82.5%
VTD	£80,601	4.64					
Long-term OS data from NMSG 5/94 for all response categories							
TD	£64,788	4.00	+£14,348	+0.36	**£39,881**	1.7%	22.6%
VTD	£79,136	4.36					
Parametric curves fitted to PETHEMA K-M curve: CR - log-logistic, PR - Weibull, NR - exponential							
TD	£65,608	3.68	+£16,284	+0.83	**£19,668**	50.8%	86.6%
VTD	£81,892	4.51					
ICER, incremental cost-effectiveness ratio; QALY, quality-adjusted life year; TD, thalidomide and dexamethasone; VTD, Velcade® (bortezomib), thalidomide and dexamethasone							

Appendix 3 The matching-adjusted indirect comparison (MAIC) analysis

Summary

The matching-adjusted indirect comparison (MAIC) was conducted in SAS version 9.2 using the method as suggested by Signorovitch et al.(2;3) Individual patient data (IPD) from the IFM-2005 study were matched to the aggregate data from the MRC IX study. The patient inclusion/exclusion criteria and baseline characteristics from the CTD arm in MRC IX formed the basis of a new population sample in which VD arm was matched by re-weighting (i.e. adjusting) patients in the IFM-2005 study by their inverse odds of being enrolled in the MRC IX study.

Background

Signorovitch et al. proposes a method that leverages IPD from trials of one treatment matched to baseline aggregate statistics reported from trials of another treatment. This involves re-weighting patients in the trial and treatment arm(s) for which IPD is available so that their average baseline characteristics match those reported from the trials of the comparator treatment(s). Matching is accomplished by re-weighting patients in the trial with IPD by their inverse odds of having enrolled in that trial versus having enrolled in the trial that only contains the aggregate data. After matching, the weighted mean baseline characteristics match those reported for the trial without IPD and the treatment outcomes can thus be compared across balanced trial populations.

This method accounts, amongst others, for cross-trial differences in patients' baseline characteristics which can bias more traditional indirect comparisons and satisfies the request by the Appraisal Committee: "In the absence of a network to facilitate a robust comparison, … a careful comparison using single arms from relevant clinical trials, taking into account differences in trial design and baseline characteristics" (ACD paragraph 1.3).

Methods

The MAIC was conducted in SAS version 9.2 using the method as suggested by Signorovitch et al.(2;3)

Step 1: Selection of studies

The Myeloma IX study is the only study available evaluating the comparator of choice, namely CTD. Only aggregate data were available from the publication by Morgan et al.(7) IPD were available from the IFM-2005 study.

Step 2: Alignment of inclusion/exclusion criteria

Characteristics of the MRC IX study were compared against those of the IFM-2005 study. Specifically, inclusion and exclusion criteria were compared and with the exception of the patient's age (see next paragraph), were found to be similar between the MRC IX and IFM-2005 studies.

The MRC IX study had no upper age limit for inclusion in the trial, whereas IFM-2005 study specified that patients >65 years were to be excluded. The limitation of having only aggregate data for the MRC IX study and no patients older than 65 in the PETHEMA/GEM and IFM-2007 meant it was not possible to fully balance the age distribution between trials.

Therefore, no patients were excluded from the IPD data based on inclusion/exclusion criteria.

Step 3: Selection of matching variable(s)

The baseline characteristics for the CTD arm, as presented in the publication by Morgan et al., were matched with the baseline characteristics in the VD arm in IFM-2005 study. Additional data for type of myeloma (e.g. IG-A, IG-G) was also available. When the median of a characteristic is reported by Morgan et al. instead of the mean, such as is the case for age, the results were interpreted as a binary characteristic, e.g. age >59 years, with a frequency of 50%.

Ideally, baseline characteristics which are likely to be predictive of patient outcomes should be considered as potential matching variables. However, Morgan et al. only reports a few of the factors

considered prognostic in myeloma : ISS stage, certain cytogenetic abnormalities and serum β2-microglobulin concentration.

Five categories of baseline characteristics were specified:

1. ISS stage 1, 2, 3
2. Serum β2-microglobulin concentration; patients with cytogenetic abnormality t4;14
3. Age
4. Male gender
5. Light chain myeloma, IG-A, IG-D, IG-G

Step 4: Matching of patient-level data to aggregate data

The matched population is obtained according to the algorithm as described by Signorovitch at al. (2)

Step 5: Deriving outcome measures

The distributions of patients between post-induction and post-transplant response rate categories in the VD arm for IFM-2005 were determined for the new matched sample populations. The response rate categories were CR, PR and NR where 1) CR=CR+nCR+VgPR, PR=PR and NR = all other and 2) CR=CR+nCR, PR=VgPR+PR and NR=all other categories. The SCT vs. no SCT rates were also determined for the IFM-2005 study.

Limitations

It was not possible to fully match the age distribution of the MRC IX trial as the IFM-2005 study excluded patients >65 years of age. As limited baseline characteristics were published in the MRC IX trial, unreported baseline characteristics which affect prognosis could not be matched between the studies. Such unobserved differences may result in residual confounding and bias.

Appendix 4: Input parameters for the VD vs. CTD cost-effectiveness analysis

Efficacy data for VD and CTD

CR=CR+nCR+VGPR

	CTD induction (MRC Myeloma IX study)				VD induction (IFM study, adjusted)			
Post-induction response rate	**N**	**N %**			**N**	**N %**		
CR (=CR+nCR+VGPR)	240	43.2%			54	45.3%		
PR	218	39.3%			42	35.3%		
NR (less than PR)	97	17.5%			23	19.5%		
TOTAL	555				120			
SCT rate given post-ind. response	**SCT**	**SCT %**	**No SCT**	**No SCT %**	**SCT**	**SCT %**	**No SCT**	**No SCT %**
CR (=CR+nCR+VGPR)	204	85.0%	36	15.0%	53	98.0%	1	2.0%
PR	160	73.4%	58	26.6%	40	95.0%	2	5.0%
NR (less than PR)	6	6.2%	91	93.8%	14	58.0%	10	42.0%
TOTAL	370	66.7%	185	33.3%	107	89.1%	13	10.9%
Post-ind. response given SCT status	**SCT**	**Resp rate**	**No SCT**	**Resp rate**	**SCT**	**Resp rate**	**No SCT**	**Resp rate**
CR (=CR+nCR+VGPR)	204	55.1%	36	19.5%	53	49.7%	1	8.5%
PR	160	43.2%	58	31.4%	40	37.6%	2	16.2%
NR (less than PR)	6	1.6%	91	49.2%	14	12.7%	10	75.3%
TOTAL	370	100.0%	185	100.0%	107	100.0%	13	100.0%
Post-transplant response rate	**N**	**N %**			**N**	**N %**		
CR (=CR+nCR+VGPR)	273	73.8%			88	73.2%		
PR	66	17.8%			24	20.4%		
NR (less than PR)	31	8.4%			8	6.4%		
TOTAL	370				120			

CR=CR+nCR

	CTD induction (MRC Myeloma IX study)				VD induction (IFM study)			
Post-induction response rate	**N**	**N %**			**N**	**N %**		
CR (=CR+nCR)	72	13.0%			26	21.7%		
PR (=PR+VGPR)	386	69.5%			71	58.9%		
NR (less than PR)	97	17.5%			23	19.5%		
TOTAL	555				120			
SCT rate given post-ind. response	**SCT**	**SCT %**	**No SCT**	**No SCT %**	**SCT**	**SCT %**	**No SCT**	**No SCT %**
CR (=CR+nCR)	64	88.9%	8	11.1%	25	98.0%	1	2.0%
PR (=PR+VGPR)	300	77.7%	86	22.3%	67	95.0%	4	5.0%
NR (less than PR)	6	6.2%	91	93.8%	14	61.5%	9	38.5%
TOTAL	370	66.7%	185	33.3%	107	89.1%	13	10.9%
Post-ind. response given SCT status	**SCT**	**Resp rate**	**No SCT**	**Resp rate**	**SCT**	**Resp rate**	**No SCT**	**Resp rate**
CR (=CR+nCR)	64	17.3%	8	4.3%	25	23.8%	1	4.1%
PR (=PR+VGPR)	300	81.1%	86	46.5%	67	62.8%	4	27.0%
NR (less than PR)	6	1.6%	91	49.2%	14	13.4%	9	68.9%
TOTAL	370	100.0%	185	100.0%	107	100.0%	13	100.0%
Post-transplant response rate	**N**	**N %**			**N**	**N %**		
CR (=CR+nCR)	185	50.0%			48	39.9%		
PR (=PR+VGPR)	154	41.6%			64	53.6%		
NR (less than PR)	31	8.4%			8	6.4%		
TOTAL	370				120			

Median OS by response category

(months)	CR	PR	NR (less than PR)
MRC VII	88.6	39.8	25.6
IFM90	126.6	44.2	6.0
Alvares	88.8	63.6	49.2
NMSG 5/94	71.0	64.4	64.4

Parameters for parametric curves fitted to the Kaplan-Meier curves from IFM-2005

	CR (CR+nCR+VGPR)	PR	NR (less than PR)
Exponential function			
Scale	0.00008	0.00017	0.00038
Shape	1.0000	1.0000	1.0000
Weibull function			
Scale	0.00000	0.00000	0.00316
Shape	1.49009	1.91791	0.68875
Log-logistic function			
Scale	-13.02252	-15.60218	-6.02551
Shape	1.52138	2.01735	0.75706

	CR (CR+nCR)	PR (PR+VGPR)	NR (less than PR)
Exponential function			
Scale	0.00006	0.00015	0.00038
Shape	1.00000	1.00000	1.00000
Weibull function			
Scale	0.00000	0.00000	0.00316
Shape	2.31589	1.74642	0.68875
Log-logistic function			
Scale	-19.02385	-14.39924	-6.02551
Shape	2.36128	1.82050	0.75706

Appendix 5 Detailed results of VD vs. CTD cost-effectiveness analysis

Deterministic analysis (CR in the model = CR+nCR+VGPR in the trials)

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£) incremental (QALYs)
Long-term OS data from MRC Myeloma VII for all response categories							
CTD	£58,037	5.61	3.70	+£12,225	+0.86	+0.59	**£20,588**
VD	£70,261	6.47	4.29				
Long-term OS data from IFM90 for all response categories							
CTD	£57,093	6.31	4.16	+£14,304	+1.17	+0.80	**£17,866**
VD	£71,396	7.49	4.96				
Long-term OS data from Alvares for all response categories							
CTD	£61,603	6.49	4.27	+£11,369	+0.66	+0.47	**£24,267**
VD	£72,971	7.15	4.74				
Long-term OS data from NMSG 5/94 for all response categories							
CTD	£60,458	6.17	4.06	+£10,768	+0.44	+0.32	**£33,435**
VD	£71,226	6.61	4.38				
Parametric curves fitted to PETHEMA K-M curve: CR - Weibull, PR - log-logistic, NR - exponential							
CTD	£67,229	7.91	5.21	+£12,530	+0.95	+0.66	**£18,864**
VD	£79,760	8.85	5.88				

CTD, cyclophosphamide, dexamethasone and dexamethasone; ICER, incremental cost-effectiveness ratio; LYG, life years gained; QALY, quality-adjusted life year; VD, Velcade® (bortezomib) and dexamethasone

Probabilistic analysis (CR in the model = CR+nCR+VGPR in the trials)

Technologies	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER (£) incremental (QALYs)	Prob(VD is cost effective @ £20k/QALY)	Prob(VD is cost effective @ £30k/QALY)
Long-term OS data from MRC Myeloma VII for all response categories							
CTD	£58,256	3.93	+£12,237	+0.55	**£22,305**	36.3%	79.0%
VD	£70,494	4.48					
Long-term OS data from IFM90 for all response categories							
CTD	£57,488	4.45	+£14,743	+0.90	**£16,438**	68.7%	94.2%
VD	£72,231	5.35					
Long-term OS data from Alvares for all response categories							
CTD	£61,472	4.30	+£11,373	+0.48	**£23,816**	24.7%	77.0%
VD	£72,845	4.78					
Long-term OS data from NMSG 5/94 for all response categories							
CTD	£60,378	4.10	+£10,748	+0.33	**£33,017**	4.6%	37.9%
VD	£71,126	4.43					
Parametric curves fitted to PETHEMA K-M curve: CR - Weibull, PR - log-logistic, NR - exponential							
CTD	£65,913	5.18	+£12,321	+0.65	**£19,057**	51.1%	75.7%
VD	£78,234	5.83					

CTD, cyclophosphamide, dexamethasone and dexamethasone; ICER, incremental cost-effectiveness ratio; QALY, quality-adjusted life year; VD, Velcade® (bortezomib) and dexamethasone

Deterministic analysis (CR in the model = CR+nCR in the trials)

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£) incremental (QALYs)
Long-term OS data from MRC Myeloma VII for all response categories							
CTD	£55,467	4.92	3.24	+£10,943	+0.54	+0.39	**£28,281**
VD	£66,411	5.46	3.62				
Long-term OS data from IFM90 for all response categories							
CTD	£53,905	5.34	3.51	+£12,809	+0.73	+0.51	**£25,070**
VD	£66,714	6.07	4.02				
Long-term OS data from Alvares for all response categories							
CTD	£60,580	6.17	4.05	+£10,727	+0.52	+0.38	**£28,121**
VD	£71,307	6.69	4.43				
Long-term OS data from NMSG 5/94 for all response categories							
CTD	£60,130	6.08	3.99	+£10,443	+0.39	+0.30	**£34,750**
VD	£70,572	6.47	4.29				
Parametric curves fitted to PETHEMA K-M curve: CR - log-logistic, PR - Weibull, NR - Alvares exponential							
CTD	£63,925	6.34	4.18	+£11,325	+0.63	+0.46	**£24,589**
VD	£75,250	6.97	4.65				
CTD, cyclophosphamide, dexamethasone and dexamethasone; ICER, incremental cost-effectiveness ratio; LYG, life years gained; QALY, quality-adjusted life year; VD, Velcade® (bortezomib) and dexamethasone							

Probabilistic analysis (CR in the model = CR+nCR in the trials)

Technologies	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER (£) incremental (QALYs)	Prob(VD is cost effective @ £20k/QALY)	Prob(VD is cost effective @ £30k/QALY)
Long-term OS data from MRC Myeloma VII for all response categories							
CTD	£55,698	3.46	+£10,950	+0.34	**£32,640**	10.3%	46.1%
VD	£66,647	3.79					
Long-term OS data from IFM90 for all response categories							
CTD	£54,125	3.72	+£13,156	+0.57	**£23,227**	29.6%	72.2%
VD	£67,282	4.29					
Long-term OS data from Alvares for all response categories							
CTD	£60,454	4.08	+£10,753	+0.39	**£27,297**	13.1%	60.0%
VD	£71,206	4.48					
Long-term OS data from NMSG 5/94 for all response categories							
CTD	£60,091	4.03	+£10,390	+0.30	**£34,167**	4.0%	34.6%
VD	£70,482	4.34					
Parametric curves fitted to PETHEMA K-M curve: CR - log-logistic, PR - Weibull, NR - Alvares exponential							
CTD	£63,305	4.44	+£11,427	+0.49	**£23,425**	30.8%	67.8%
VD	£74,732	4.93					
CTD, cyclophosphamide, dexamethasone and dexamethasone; ICER, incremental cost-effectiveness ratio; QALY, quality-adjusted life year; VD, Velcade® (bortezomib) and dexamethasone							

Appendix 6 Minutes of the advisory board meetings (November 2013)

Janssen held two advisory boards on the 20th and 21st of November. The minutes for both advisory boards are summarised below.

Attendees

20 November 2013

[redacted]

21 November 2013

[redacted]

20 November 2013 & 21 November 2013

[redacted]

Introductions

RM welcomed the KOLs, explained the purpose of the advisory boards and thanked participants for their time.

Overview of the NICE ACD

TI explained the key points from the NICE ACD.

- Bortezomib in combination with thalidomide and dexamethasone (VTD) is not recommended within its marketing authorisation, that is, for the induction treatment of adults with previously untreated multiple myeloma for whom high-dose chemotherapy with haematopoietic stem cell transplantation is suitable.
- The Committee is minded not to recommend bortezomib in combination with dexamethasone (VD) within its marketing authorisation, that is, for the induction treatment of adults with previously untreated multiple myeloma for whom high-dose chemotherapy with haematopoietic stem cell transplantation is suitable.
- NICE has requested further information about VD vs. CTD, and VD vs. alternative comparator when CTD is not suitable.

Discussion

<u>What does “CTD is not suitable” mean from a clinical perspective?</u>

- CTD as an induction therapy is not suitable when a transplant-eligible MM patient:
- requires a rapid response due to renal failure or pending renal failure,
- has either prior or active thrombosis,
- is pregnant or is not willing to consent to pregnancy prevention or take part in a risk management programme,
- has existing peripheral neuropathy,
- is likely to have an allergic reaction or rash, or,
- may not adhere to dosing of thalidomide at home.

<u>When CTD is not suitable for a particular patient, what regimen(s) do you currently use?</u>

- Clinicians would give Velcade if they could not use CTD. They would apply for funding.
- Idarubicin and dexamethasone (Z-Dex)
- Cyclophosphamide and dexamethasone (Cyclo-Dex, or CD)
- Vincristine, doxorubicin and dexamethasone (VAD)
- Oral therapies
- Note - no comparative data in the subgroups in papers, so there will be uncertainty.
- Do not want to give bendamustine.

Appendix 7: Input parameters for the VD vs. VAD cost-effectiveness analysis

Efficacy data for VD and CTD

Post-induction response — Post-hoc analysis at the clarification stage

	VAD (N=121)		VD (N=120)	
For econ model				
CR (=CR+nCR+VGPR)	19	15.7%	49	40.8%
PR	53	43.8%	44	36.7%
NR (less than PR)	49	40.5%	27	22.5%
TOTAL	121	100.0%	120	100.0%
CR (=CR + nCR)	9	7.4%	23	19.2%
PR (=PR+VRPR)	63	52.1%	70	58.3%
NR (less than PR)	49	40.5%	27	22.5%
TOTAL	121	100.0%	120	100.0%

SCT rate — Post-hoc analysis at the clarification stage

	VAD (N=121)		VD (N=120)	
Transplant	100	82.6%	106	88.3%
No transplant	21	17.4%	14	11.7%
TOTAL	121	100.0%	120	100.0%

SCT rate by post-induction response — Post-hoc analysis at the clarification stage

	VAD (N=121)			
	SCT		No SCT	
For econ model				
CR (=CR+nCR+VGPR)	18	94.7%	1	5.3%
PR	50	94.3%	3	5.7%
NR (less than PR)	32	65.3%	17	34.7%
TOTAL	100	82.6%	21	17.4%
CR (=CR + nCR)	9	100.0%	0	0.0%
PR (=PR+VRPR)	59	93.7%	4	6.3%
NR (less than PR)	32	65.3%	17	34.7%
TOTAL	100	82.6%	21	17.4%

	VD (N=120)			
	SCT		No SCT	
For econ model				
CR (=CR+nCR+VGPR)	48	98.0%	1	2.0%
PR	43	97.7%	1	2.3%
NR (less than PR)	15	55.6%	12	44.4%
TOTAL	106	88.3%	14	11.7%
CR (=CR + nCR)	22	95.7%	1	4.3%
PR (=PR+VRPR)	69	98.6%	1	1.4%
NR (less than PR)	15	55.6%	12	44.4%
TOTAL	106	88.3%	14	11.7%

Post-transplant response (SCT population only) — Post-hoc analysis at the clarification stage

	VAD (N=100)		VD (N=106)	
For econ model				
CR (=CR+nCR+VGPR)	52	52.0%	78	73.6%
PR	37	37.0%	21	19.8%
NR (less than PR)	11	11.0%	7	6.6%
TOTAL	100	100.0%	106	100.0%
CR (=CR + nCR)	31	31.0%	43	40.6%
PR (=PR+VRPR)	58	58.0%	56	52.8%
NR (less than PR)	11	11.0%	7	6.6%
TOTAL	100	100.0%	106	100.0%

Median OS by response category

(months)	CR (CR+nCR)	PR	NR (less than PR)
MRC VII	88.6	39.8	25.6
IFM90	126.6	44.2	6.0
Alvares	88.8	63.6	49.2
NMSG 5/94	71.0	64.4	64.4

Parameters for parametric curves fitted to the Kaplan-Meier curves from IFM-2005

	CR (CR+nCR+VGPR)	PR	NR (less than PR)
Exponential function			
Scale	0.00008	0.00017	0.00038
Shape	1.0000	1.0000	1.0000
Weibull function			
Scale	0.00000	0.00000	0.00316
Shape	1.49009	1.91791	0.68875
Log-logistic function			
Scale	-13.02252	-15.60218	-6.02551
Shape	1.52138	2.01735	0.75706

	CR (CR+nCR)	PR (PR+VGPR)	NR (less than PR)
Exponential function			
Scale	0.00006	0.00015	0.00038
Shape	1.00000	1.00000	1.00000
Weibull function			
Scale	0.00000	0.00000	0.00316
Shape	2.31589	1.74642	0.68875
Log-logistic function			
Scale	-19.02385	-14.39924	-6.02551
Shape	2.36128	1.82050	0.75706

Appendix 8 Detailed results of VD vs. VAD cost-effectiveness analysis

Deterministic analysis (CR in the model = CR+nCR+VGPR in the trials)

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£) incremental (QALYs)
Long-term OS data from MRC Myeloma VII for all response categories							
VAD	£56,455	5.35	3.52	+£14,000	+1.07	+0.74	**£18,914**
VD	£70,455	6.42	4.26				
Long-term OS data from IFM90 for all response categories							
VAD	£54,938	5.86	3.85	+£16,238	+1.55	+1.05	**£15,393**
VD	£71,176	7.40	4.90				
Long-term OS data from Alvares for all response categories							
VAD	£60,832	6.44	4.23	+£12,353	+0.68	+0.48	**£25,575**
VD	£73,184	7.11	4.71				
Long-term OS data from NMSG 5/94 for all response categories							
VAD	£60,164	6.26	4.11	+£11,331	+0.35	+0.26	**£42,811**
VD	£71,495	6.60	4.37				
Parametric curves fitted to PETHEMA K-M curve: CR - Weibull, PR - log-logistic, NR - exponential							
VAD	£66,433	7.78	5.12	+£13,480	+1.04	+0.73	**£18,489**
VD	£79,913	8.82	5.85				
ICER, incremental cost-effectiveness ratio; LYG, life years gained; QALY, quality-adjusted life year; VAD, vincristine, doxorubicin and dexamethasone; VD, Velcade® (bortezomib) and dexamethasone							

Probabilistic analysis (CR in the model = CR+nCR+VGPR in the trials)

Technologies	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER (£) incremental (QALYs)	Prob(VD is cost effective @ £20k/QALY)	Prob(VD is cost effective @ £30k/QALY)
Long-term OS data from MRC Myeloma VII for all response categories							
VAD	£56,730	3.76	+£13,846	+0.69	**£20,096**	49.2%	86.9%
VD	£70,576	4.45					
Long-term OS data from IFM90 for all response categories							
VAD	£55,161	4.09	+£16,530	+1.15	**£14,392**	76.5%	93.1%
VD	£71,691	5.24					
Long-term OS data from Alvares for all response categories							
VAD	£60,582	4.27	+£12,359	+0.48	**£25,494**	19.1%	67.1%
VD	£72,941	4.76					
Long-term OS data from NMSG 5/94 for all response categories							
VAD	£59,936	4.15	+£11,308	+0.27	**£42,528**	1.8%	19.9%
VD	£71,245	4.42					
Parametric curves fitted to PETHEMA K-M curve: CR - Weibull, PR - log-logistic, NR - exponential							
VAD	£65,214	5.10	+£13,041	+0.70	**£18,761**	52.1%	67.1%
VD	£78,255	5.79					
ICER, incremental cost-effectiveness ratio; QALY, quality-adjusted life year; VAD, vincristine, doxorubicin and dexamethasone; VD, Velcade® (bortezomib) and dexamethasone							

Deterministic analysis (CR in the model = CR+nCR in the trials)

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£) incremental (QALYs)
Long-term OS data from MRC Myeloma VII for all response categories							
VAD	£54,290	4.75	3.12	+£12,619	+0.69	+0.48	**£26,146**
VD	£66,909	5.44	3.60				
Long-term OS data from IFM90 for all response categories							
VAD	£52,233	5.01	3.29	+£14,448	+1.00	+0.69	**£20,910**
VD	£66,680	6.01	3.98				
Long-term OS data from Alvares for all response categories							
VAD	£59,962	6.16	4.05	+£11,772	+0.50	+0.37	**£32,207**
VD	£71,734	6.66	4.41				
Long-term OS data from NMSG 5/94 for all response categories							
VAD	£59,886	6.17	4.05	+£11,149	+0.29	+0.23	**£48,425**
VD	£71,035	6.47	4.28				
Parametric curves fitted to PETHEMA K-M curve: CR - Weibull, PR - log-logistic, NR - Alvares exponential							
VAD	£63,675	6.37	4.20	+£11,909	+0.57	+0.42	**£28,411**
VD	£75,583	6.94	4.62				
ICER, incremental cost-effectiveness ratio; LYG, life years gained; QALY, quality-adjusted life year; VAD, vincristine, doxorubicin and dexamethasone; VD, Velcade® (bortezomib) and dexamethasone							

Probabilistic analysis (CR in the model = CR+nCR in the trials)

Technologies	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER (£) incremental (QALYs)	Prob(VD is cost effective @ £20k/QALY)	Prob(VD is cost effective @ £30k/QALY)
Long-term OS data from MRC Myeloma VII for all response categories							
VAD	£54,574	3.36	+£12,468	+0.44	**£28,398**	14.4%	56.4%
VD	£67,042	3.80					
Long-term OS data from IFM90 for all response categories							
VAD	£52,396	3.47	+£14,707	+0.77	**£19,123**	49.0%	83.3%
VD	£67,103	4.24					
Long-term OS data from Alvares for all response categories							
VAD	£59,820	4.09	+£11,737	+0.38	**£31,224**	5.9%	43.4%
VD	£71,557	4.46					
Long-term OS data from NMSG 5/94 for all response categories							
VAD	£59,752	4.09	+£11,074	+0.23	**£47,533**	0.5%	10.3%
VD	£70,826	4.32					
Parametric curves fitted to PETHEMA K-M curve: CR - Weibull, PR - log-logistic, NR - Alvares exponential							
VAD	£62,978	4.40	+£12,002	+0.51	**£23,634**	33.5%	55.4%
VD	£74,980	4.91					
ICER, incremental cost-effectiveness ratio; QALY, quality-adjusted life year; VAD, vincristine, doxorubicin and dexamethasone; VD, Velcade® (bortezomib) and dexamethasone							