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

ERG overview of manufacturer's response to the ACD

Produced by Southampton Health Technology Assessments Centre (SHTAC)

Authors Keith Cooper
Debbie Hartwell
Jackie Bryant

Correspondence to Debbie Hartwell
Senior Research Fellow
Southampton Health Technology Assessments Centre (SHTAC)
[redacted]

Date completed 23rd December 2013

Introduction

In the Appraisal Consultation Document (ACD) of 5/11/13, NICE requested further information from the manufacturer regarding the bortezomib and dexamethasone (VD) treatment regimen for induction treatment of adults with previously untreated multiple myeloma suitable for high-dose chemotherapy with stem-cell transplantation (SCT). NICE requested from the manufacturer:

- An indirect comparison of VD compared with cyclophosphamide, thalidomide and dexamethasone (CTD), and compared with an alternative comparator in circumstances when CTD is not suitable;
- Sensitivity analyses using assumptions suggested by the Evidence Review Group (ERG);
- Probabilistic incremental cost-effectiveness ratios (ICER) for the revised comparisons.

Full details of the requested information are available in the ACD.

The manufacturer provided their response in the form of a report (supplemented with eight appendices), four economic models and a CiC checklist. These were received by the ERG on 5/12/13. SHTAC was requested to critique the manufacturer's response and has focused on the primary issues in Sections 2, 3 and 4 of the manufacturer's response.

Manufacturer Section 2: Bortezomib in combination with thalidomide and dexamethasone

The manufacturer submitted an amended economic analysis for the comparison between VTD and TD. This analysis contains changes to the model structure with regard to mortality (2.2.1) and the transition probabilities to second-line therapy (2.1.5). In addition, the manufacturer has further analysed alternative parametric curves and data for long term overall survival (OS) by response category (2.2.2). The manufacturer states that the reason for these changes is in response to comments made by the ERG. These changes are briefly summarised below.

i) Survival benefit of stem cell transplant (2.2.1)

The original submission in February 2013 and the addendum submitted in July 2013 was not designed to capture either the post-transplant response rate or the OS benefit arising from an SCT. In response to ERG comments on this approach, that 'it would have been more transparent to distinguish the separate effects on survival of post-induction response and SCT' (ERG report, p.38), the manufacturer has made changes to the economic model so that: the survival benefit of SCT is explicitly captured; post-induction rates are applied on an intention-to-treat (ITT) basis to all patients who enter the model; probabilities of receiving an SCT depend on the post-induction response; and post-transplant rates are applied to only those who receive an SCT.

The amended structure is illustrated in Figure 4 of the manufacturer's response to the ACD.

ii) Transition probabilities to second-line therapy

The original submission in February 2013 and the addendum submitted in July 2013 used time to progression (TTP), from either SCT or the end of induction to the start of second-line therapies, based upon treatment response and treatment arm (i.e. 6 subsets). The ERG noted an anomaly caused by this approach whereby the modelled TTP contrasted with overall findings for progression-free survival (PFS) given in the trial publication (ERG report, p.42). The manufacturer acknowledged that this approach led to 'counter-intuitive and clinically implausible results' and have changed the estimation of TTP so that is based upon only treatment response.

The amended hazard ratios of time to progression from the Pethema trial are shown in Table 3 of the manufacturer's response to the ACD.

iii) Parametric curves for OS survival

The manufacturer discusses the use of the OS survival data from the MRC Myeloma VII trial. The manufacturer defends the use of the MRC Myeloma VII trial and states preference for using these data. The manufacturer compares the OS curves from the Pethema trial to that from MRC Myeloma VII and other sources for complete responders, partial responders and non-responders (Figure 7 of the manufacturer's response to the ACD). The manufacturer considers that there are no obvious face validity issues between the MRC Myeloma VII OS curves and the parametric curves fitted for Pethema (Figure 8). The manufacturer therefore continues to use the MRC Myeloma VII OS curves in the new amended analysis.

Results

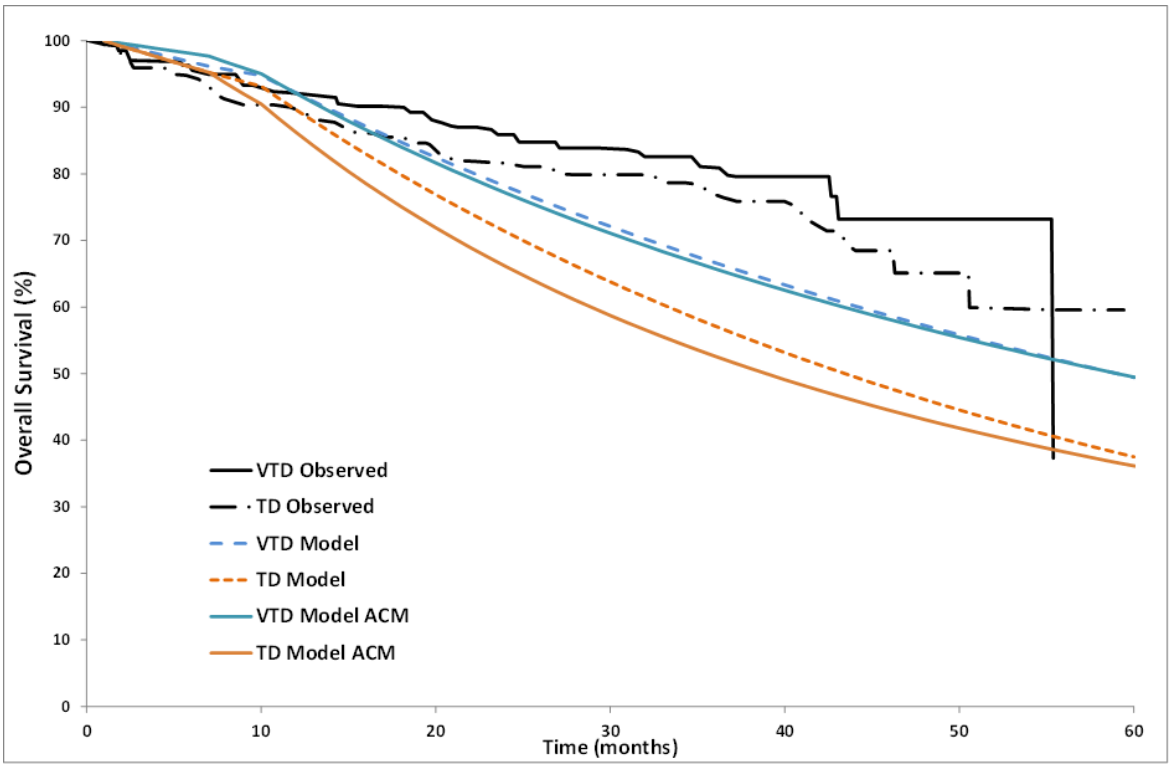
The base case results from the deterministic and probabilistic analyses from the amended economic model are shown in Tables 6 and 7 of the manufacturer's response to the ACD. The changes to the model results in a reduced ICER of £17,841 per QALY for VTD compared to TD (deterministic analysis).

ERG commentary for VTD vs. TD

The ERG notes that the model structure has changed substantially from the original model submitted in February 2013. Due to time and resource constraints, it was not possible for the ERG to check the internal validity of the new approach with the same rigour as that used for the original model. The ERG considers that the new approach appears more intuitive; however the ERG notes that the manufacturer has not checked the external validity of the new approach by validating OS against the Pethema trial. The ERG explored this in more detail below.

In its original report, the ERG compared the OS curves of the economic model to the observed Pethema OS data (p.51). The ERG considered that the manufacturer had not adequately validated the model, that the model consistently underestimated OS and that the model was systematically biased in favour of VTD. The ERG has repeated this analysis for the new economic model. Figure 1 shows the OS from the Pethema trial compared to the original model analyses and the new economic model analyses (Model ACM). As shown in Figure 1, the new model analyses continue to underestimate OS and are biased in favour of VTD. This bias appears even more pronounced in the new analyses than in the original analyses.

Figure 1: Comparison of overall survival predicted by original model and updated model and overall survival observed in Pethema trial, by treatment arm



Manufacturer Section 3: Bortezomib in combination with dexamethasone: overall population

The manufacturer submitted an amended analysis for VD compared with CTD, as requested in the NICE ACD, and this section forms the focus of the ERG's critique. The manufacturer has used the matching-adjusted indirect comparison (MAIC) method 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 carried out by the manufacturer in the absence of direct RCT evidence of VD vs CTD and was conducted using the method as suggested by Signorovitch *et al*.[1,2] This method is purported to account for cross-trial differences in patients' baseline characteristics but has some limitations. Results of an MAIC should be interpreted as those of an observational study since causation cannot be inferred from the results due to the loss of randomisation.[3] An MAIC may also be limited by sample size and this method is recommended for studies which require matching on only one or two main variables.[3] Individual patient data from the IFM, 2005 study were matched to the aggregate data from the MRC Myeloma IX study, and the response rates were re-calculated. The MAIC set of post-induction response rates are shown in Table 10 of the manufacturer's response to the ACD. The base case results from the deterministic and probabilistic analyses are shown in Table 13 and 14 of the manufacturer's response to the ACD. These show an ICER of £20,588 per QALY for VD vs. CTD (deterministic analysis.)

Commentary on adjustment of response rates using MAIC

The ERG notes that the analysis for VD compared with CTD is dependent on the results of an MAIC but it is not possible for the ERG to conduct a formal quality assessment of the MAIC because limited information regarding the methodology and assumptions used is presented; details of the resulting VD patient characteristics are also not given. Therefore, it is not clear how well the new sample population of VD patients matches the aggregate MRC Myeloma IX population, particularly with regard to age and stage of disease which differed between the original (IFM and MRC Myeloma IX) trials and which are factors typically expected to impact prognosis. The post-induction response rates for VD before and after MAIC in the manufacturer's Table 10 show that the effect of adjustments increases CR from 40.8% to 45.3% and decreases NR from 22.5% to 19.5%. Results may be at risk of confounding and bias and therefore should be treated with caution.

ERG commentary for VD vs. CTD.

The ERG notes that SCT rates for the analysis are taken from the respective trials for the treatment arms of VD vs. CTD. The SCT rate for CTD is 67% (MRC Myeloma IX) and 89% for VD (IFM 2005). As noted above, the new economic model has introduced a separate benefit for SCT, thus in their analysis patients receiving VD have an additional survival benefit over those receiving CTD. However, the proportion of patients receiving SCT is dependent upon the response rate for these patients. In the adjusted analysis, patients with CTD have a similar response to VD (Table 1). Therefore assuming a much lower SCT for CTD is inconsistent with the response data and provides a substantial bias to the VD cohort.

Table 1 Efficacy data for VD and CTD (from Appendix 4 of manufacturer's response to ACD)

	CTD, N	%	VD, 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**	**100%**	**120**	**100%**

The ERG has re-run the manufacturer's model for the analysis of VD vs. CTD assuming the same SCT rate for both treatment arms. The ICER for VD vs. CTD increases to £73,025 per QALY gained when both treatment arms have an SCT rate of 89.1% (as used for VD). When both arms have an SCT rate of 67% (as used for CTD), VD is dominated by CTD, i.e. is more expensive and less effective.

The ERG has re-run the manufacturer's model using the values for VD before MAIC (Table 10 of manufacturer's response to ACD). For this scenario, the ICER increases only slightly from £20,588 to £20,667 per QALY.

Manufacturer Section 4: Bortezomib in combination with dexamethasone when CTD is not suitable

The manufacturer carried out an amended analysis of VD compared with an alternative regimen for circumstances where CTD is not suitable, as requested by NICE. The views of clinical experts were sought to elicit the patient circumstances where CTD is not suitable, and these are presented. Clinical experts also advised on the most relevant comparator in the absence of CTD. A bortezomib-based regimen was cited as the preference, with three non-bortezomib alternative regimens suggested:

- cyclophosphamide + thalidomide (CD);
- vincristine + doxorubicin + dexamethasone (VAD);
- idarubicin + dexamethasone (Z-Dex).

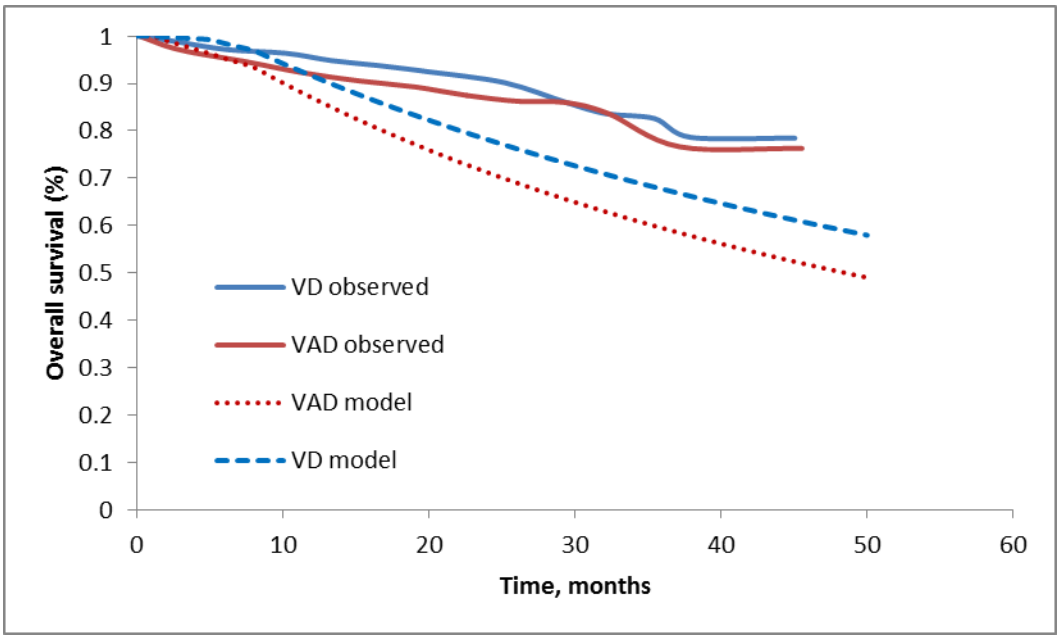
The manufacturer stated that as RCT evidence was only available for the VAD option compared to VD (i.e. the IFM trial), this was the focus of their subsequent economic analysis. The ERG would point out that only the CD option contains thalidomide as per the NICE scope.

For this analysis, the manufacturer has compared VD with VAD using the IFM 2005 trial data. The analysis has used the new model structure as described above for the analyses for VTD vs. TD. The base case results from the deterministic and probabilistic analyses are shown in Table 17 and 18 of the manufacturer's response to the ACD. The ICER for VD vs. VAD is £18,914 per QALY (deterministic analysis).

ERG commentary for VD vs. VAD

The ERG notes that the manufacturer has provided no validation for the model results for the new analysis for VD vs. VAD by comparing the predicted OS against the trial data in IFM 2005. The ERG has digitised the OS Kaplan-Meier curves presented in the trial publication and Figure 16B of the original MS for VD and VAD arms of the IFM trial and plotted these against OS predicted by the model for these treatment arms (Figure 2). As for the analyses for VTD vs. TD, the model results for OS show a poor fit to the trial results; the model underestimates survival compared to the trial data. In addition there is a greater difference in survival benefit between the treatment arms in the model than in the trial, indicating the model results are biased in favour of VD.

Figure 2: Comparison of overall survival predicted by updated model and overall survival observed in IFM trial, by treatment arm



Conclusion

The manufacturer has addressed the request by NICE by providing additional analyses for VTD vs. TD, VD vs. CTD and VD vs. VAD. The economic model for these analyses has been modified to include several structural changes. However, the ERG notes that the estimation of long term survival and progression free survival is based upon surrogate outcomes for post-induction response, and that there is not a good fit between OS predicted in the model and the trial data. The ERG considers that the results presented in the response to the ACD are systematically biased in favour of the intervention.

`

Reference List

1. 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. Pharmacoepidemiology and Drug Safety 2012; 21 Suppl 2: 130-37.

2. Signorovitch J, Wu EQ, Yu AP, Gerrits CM, Kantor E, Bao Y et al. Comparative effectiveness without head-to-head trials. *Pharmacoeconomics* 2010; 28(10): 935-45.

3. Malagone E and Sherman S. Matching-Adjusted Indirect Comparison Analysis Using Common SAS® 9.2. PROCEDURES. SAS Global Forum 2011; Poster 228-2011.