BORTEZOMIB AND THALIDOMIDE FOR THE FIRST LINE TREATMENT OF MULTIPLE MYELOMA

APPEAL AGAINST THE FINAL APPRAISAL DETERMINATION ISSUED BY THE NATIONAL INSTITUTE FOR HEALTH AND CLINICAL EXCELLENCE ON 19th AUGUST 2010

EXECUTIVE SUMMARY

Janssen-Cilag Ltd provides its notice of appeal in relation to the Final Appraisal Determination for bortezomib and thalidomide for the first line treatment of multiple myeloma. The appeal is brought under grounds one and two of NICE's appeal procedures.

1. Ground 1: Procedural Fairness
   1.1. NICE's failure to disclose to consultees the economic model upon which its guidance is based, lacks transparency and is unfair
   1.2. The Appraisal Committee's reasons for limiting use of bortezomib to patients who have contraindications to thalidomide, rather than those for whom thalidomide is clinically inappropriate are unexplained.
   1.3. In deciding to place less weight on the thalidomide studies which included a "maintenance" phase, the Appraisal Committee has relied on evidence from the Assessment Group which has not been disclosed to consultees
   1.4. While the Appraisal Committee decided that the thalidomide studies which included a "maintenance" phase should receive less weight than those with no "maintenance" phase, the weighting carried out is unexplained and appears inconsistent with NICE's Guide to the Methods of Technology Appraisal
2. Ground 2: Perversity
   2.1. The exclusion of critical evidence from thalidomide trials has resulted in a fundamentally flawed evidence synthesis that is not a sound basis for decision-making
   2.2. The Appraisal Committee have demonstrated a lack of consistency in considering clinical experts' opinion to inform its decision
   2.3. The Appraisal Committee's conclusion at paragraph 4.3.8 of the FAD that an assessment of cost-effectiveness which assumes use of 31.5 vials of bortezomib

"should be considered the most optimistic estimate for clinical practice", is inconsistent with the available evidence.

2.4. Failure to consider vial sharing of bortezomib is inconsistent with the available evidence and with the approach followed in other appraisals

INTRODUCTION

Bortezomib (Velcade) is the subject of a marketing authorisation granted by the European Commission under the centralised procedure and held by JANSSEN-CILAG INTERNATIONAL N.V.. It is supplied in the UK by Janssen-Cilag Ltd.

Bortezomib is a novel antineoplastic agent that inhibits the proteasome enzymes which influence cell proliferation leading to apoptosis (programmed cell death). It currently has several indications for the treatment of multiple myeloma, a haematological malignancy affecting B lymphocytes (plasma cells) which are responsible for the production of immunoglobulins (antibodies). The average age at presentation of multiple myeloma is 68-70 years and, median survival is approximately 4-5 years from diagnosis. The disease is approximately twice as common in black as white populations. More detailed information regarding multiple myeloma is provided in Janssen's original submission for the purposes of this appraisal.

PROCEDURAL HISTORY OF THE APPRAISAL

Following a single technology appraisal in 2006-7, guidance was issued by NICE in October 2007, recommending bortezomib monotherapy, for the treatment of progressive multiple myeloma in people who are at first relapse having received one prior therapy and who have undergone, or are unsuitable for, bone marrow transplantation (Technology Appraisal Guidance No 129).

In 2008, the license for bortezomib was extended to front-line usage in patients who had not been previously treated and who were unsuitable for a stem-cell transplant. The Multiple Technology Appraisal (MTA) of bortezomib and thalidomide for front line treatment of multiple myeloma was referred to NICE in March 2009 and the final scope was issued in May 2009. Following discussions at the first scoping workshop, a group of

Haematologists representing myeloma physicians met with the Department of Health (DH) to propose broadening the scope of the appraisal. Their objective was to seek approval to appraise bortezomib and thalidomide in combination with an alkylating agent and a corticosteroid, rather than being restricted specifically to the evaluation of combination therapy with melphalan and prednisolone (MP), in line with the specific combinations referred to in the SmPCs of both technologies. Clinicians requested this amendment to the original scope to enable NICE to appraise the unlicensed CTDa (cyclophosphomide, thalidomide and dexamethasone attenuated) combination which is commonly used in UK clinical practice and which has been studied in the recent MRC sponsored Myeloma IX study (Owen, IMW 2009).

Janssen provided its submission to NICE in relation to bortezomib in October 2009. An Assessment Report dated 1 February 2010 was prepared by the Southampton Health Technology Assessments Centre (SHTAC) and provided to consultees to the appraisal; Janssen submitted its comments in relation to the Assessment Report on 16 March 2010.

The Appraisal Committee met for the first time to consider this appraisal on 13 April 2010 and an Appraisal Consultation Document (ACD) was issued to consultees in May 2010. The preliminary recommendations in the ACD provided in relation to bortezomib:

"Bortezomib in combination with an alkylating agent and a corticosteroid is recommended as an option for the first line treatment of multiple myeloma in people for whom:

High-dose chemotherapy with stem cell transplantation is considered inappropriate and

The person is unable to tolerate or has contraindications to thalidomide".

Janssen submitted comments to NICE in response to the ACD, as did other stakeholders. The Appraisal Committee met for a second time to consider bortezomib on 8 July 2010. Following this meeting a Final Appraisal Determination (FAD) was prepared and issued to consultees on 19 August 2010. The draft guidance for bortezomib contained within section 1 of the FAD remained unchanged from the ACD.

GROUNDS OF APPEAL

1. Ground 1: Procedural Unfairness

1.1. NICE's failure to disclose to consultees the economic model upon which its guidance is based, lacks transparency and is unfair

The draft recommendations of the Appraisal Committee set out in the FAD are primarily based on an assessment of the clinical and cost-effectiveness of bortezomib and thalidomide, carried out using an evidence synthesis and economic model developed by the Assessment Group in this appraisal, SHTAC. During the consultation process, and despite repeated requests from Janssen to reconsider the decision, the economic model has not been released to consultees and, accordingly, the principal basis for the recommendations of the Committee is unclear and consequently is not able to be investigated or tested by consultees. As a consequence, Janssen is unable to fully understand the reasons for the draft guidance set out in the FAD and has been prejudiced in its inability to participate fully in the consultation process.

NICE's Guide to the Multiple Technology Appraisal Process (the MTA Guide) confirms at paragraph 3.2.38 that the economic model relied upon by the Appraisal Committee will be disclosed to consultees:

> *"...To make sure that the appraisal process is transparent, NICE considers it essential that evidence on which the Appraisal Committee's decisions are based is publicly available. All the evidence seen by the Committee should be available to all consultees and commentators. This includes an executable version of the economic model. Under exceptional circumstances, unpublished evidence is accepted under agreement of confidentiality. It is important to consider carefully the information that is marked as confidential and therefore not releasable, because it may be difficult to identify how evidence has been used and interpreted. NICE will ask for restrictions on release of evidence to be reconsidered if there appears to be no obvious reason for the restrictions, or when such restrictions would make it difficult or impossible for NICE to show the evidence on which the guidance is based."*

The MTA Guide expressly considers the status of any economic model developed by the Assessment Group at paragraph 3.4.7

> *"If the Assessment Group has produced an economic model in support of the assessment report, NICE offers to send it (in its executable form) to consultees and commentators during consultation on the assessment report. This offer is made if the economic model does not contain confidential information. If it does contain confidential material NICE will ask the Group to lift any restrictions if possible (in discussion with the data owners) or remove the confidential material if this can be done without severely limiting the model's function. Consultees and commentators must make requests for a copy of the model in writing. NICE supplies the model on the basis that the consultee or commentator agrees, in writing, to [specified conditions]."*

Against this background, at the time the Assessment Report in this appraisal was provided to consultees in February 2010, NICE indicated in a covering letter:

> '*The Assessment Group have prepared an economic model, which is unavailable to consultees and commentators because it contains information designated as confidential and cannot be redacted without producing severe limitations on the functionality of the model*'.

In its response to the Assessment Report dated 16 March 2010, Janssen therefore asked NICE to "*reconsider releasing the economic model produced by Southampton Health Technology Assessments Centre to all consultees*", indicating "*if Janssen-Cilag's commercial in confidence data is considered an obstruction, then Janssen-Cilag would like to discuss how this might be addressed to allow a complete consultation process to take place*".

NICE responded to Janssen's request by email dated 18 June 2010, stating: "*please be aware that as well as Janssen-Cilag's commercial and confidence information, the Assessment Group's economic model also contains information designated academic in confidence and which cannot be redacted without producing severe limitations on the functionality of the model. Therefore, in accordance with NICE MTA process, the Institute was unable to release the economic model to consultees and commentators for this appraisal*".

Janssen responded to this information, by a further email dated 22 June 2010 stating "*please could you provide me with further information on this issue. In particular, I would like to know which stakeholder is preventing the release of the model and the exact nature of the evidence which has been deemed to be academic in confidence. Finally, I would like to register that the fact that (1) as the model is not available and (2) the assessment report generally presents only ICERs, rather than disaggregated costs and effectiveness results, we do not feel that we have been able to fully understand the Assessment Group modelling in any level of detail during this consultation process.*"

No response was received to this request and, in particular, NICE did not provide Janssen with the requested information regarding the nature of the evidence which seemingly precluded disclosure of the Assessment Group's model.

The fact that Janssen has not been provided with the Assessment Group's economic model has substantially prejudiced its ability to understand the conclusions reached by the Assessment Group, relied upon by the Appraisal Committee, and to participate in the consultation process. The differences between the assessment of cost effectiveness carried out by Janssen and by the Assessment Group are summarised at Section 4.2.5 - 4.2.9 of the FAD. These are very substantial: the Assessment Group concluded that melphalan prednisolone with thalidomide (MPT) dominated a regime comprising bortezomib, melphalan and prednisolone (VMP) (i.e. it is more effective and cheaper) or that the ICER for VMP vs. MPT was around £320,000 depending on the assumptions made; in contrast Janssen assessed this comparison as resulting in an ICER of £14,400 and £21,600 (in scenario 4 and 5 respectively, presented in Janssen's response to the ACD (paragraph 4.2.30 of the FAD)). Various explanations for this discrepancy between the ICERs calculated by the Assessment Group and by Janssen are proposed in the FAD (paragraphs 4.2.25 - 4.2.29), including differences in the numbers of vials of bortezomib used, the inclusion or exclusion of costs after first line treatment, the modelling of adverse event and the estimates of QALY benefits. However, Janssen has not been able adequately to investigate these matters or to test the reliability of the Assessment Group's model as a result of the refusal by NICE to disclose the model.

During the consultation on the ACD, Janssen's only option was to attempt to replicate the Assessment Group's model using its own model in order to try and better understand the reasons for these discrepancies and to enable them to more effectively participate in the consultation process. The Assessment Group states in paragraph 4.2.34 of the FAD that:*'there was close agreement between the two models when using the same assumptions and data for both models'*. Although the Assessment Group claimed that the models generated similar results, it cannot validate whether this is indeed the case as it has been unable to independenly scrutinise or validate the Assessment Group model. Furthermore, Janssen remain unclear why, if the models do show close agreement that their basecase and the Assessment Group basecase results vary so dramatically. Without access to the model it has been impossible to understand this. The lack of transparency is compounded further by an assessment report which fails to follow the NICE methods guide in the way it reports the economic model results. Our assertion is that the Assessment Report is of poor quality in that critically important cost-effectiveness results are systematically presented only as summary cost-effectiveness ratios without first providing a standard disaggregated breakdown of the costs and outcomes that are used in their derivation. Sections 5.9.2. and 5.9.3 of the Guide to methods of technology appraisal are explicit on this point, stating that:

> *"The expected value of each component of cost and expected total costs should be presented; expected QALYs for each option compared in the analysis should also be detailed in terms of their main contributing components. ICERs should be calculated as appropriate"*
>
> and
>
> *"The main individual components comprising both costs and QALYs for the intervention and control treatment pathways should be tabulated. For QALYs this includes presenting the life-year component separately....."*

These basic requirements were not adhered to in the assessment report with the result that consultees were not able to scrutinise the costs and outcomes that the model generates prior to their impenetrable combination into a cost-effectiveness ratio. This situation simply compounds the problems caused by a failure to release the economic model and has further

prejudiced Janssen from being able to understand the economic model results and input into the consultation process adequately. Taken together these two failings of transparency and deviation from NICE's methods guide make it impossible to have meaningfully participated in this element of the consultation process, which was so material to the final decision.

Finally, the fact that the model and its input information was withheld on an "academic in confidence" basis has meant that as well as not being able to understand how the model generates results, we are unable to verify and validate many of the key input assumptions and variables relied upon by the Assessment Group in their model, including and in relation to overall survival, response rates to treatment, adverse events, deaths during treatment, patient withdrawals and the duration and intensity of first line treatment. These matters are all fundamental to the cost effectiveness assessment in this case and the fact that evidence which has formed the basis for the Appraisal Committee's recommendations has not been disclosed represents a very serious flaw in the process. The effect of this lack of transparency is that it is impossible either to understand or to critique the conclusions of the Assessment Group, relied upon by the Appraisal Committee in their consideration of thalidomide and bortezomib. This is clearly unfair.

In summary it is unfair that Janssen was not able to fully participate in the consultation process due to the non-release of the economic model, the lack of reporting of disaggretated results and the withholding of 'academic in confidence' data given the implications of the decision being made.

Whilst we do not propose to set out the legal arguments in full in this appeal letter, the requirement to disclose the key information relied upon by the Appraisal Committee for the purposes of its guidance, including the economic model which forms the basis for the cost effectiveness assessment, has been confirmed by the Court of Appeal in R ota Eisai Ltd v NICE [2008] EWCA CIV 438, and subsequent authorities.

The issue of disclosure of an economic model containing information categorized "academic in confidence" was specifically considered by the Administrative Court in R ota Servier Laboratories v NICE [2009] EWHC 281. In that case, the Court stated: "*even after a*

confidentiality undertaking has been justifiably given, NICE remains under a positive duty, at appropriate stages in the process, to take all reasonable steps to obtain permission to disclose the information. In deciding what are reasonable steps it must keep firmly in mind the high importance of fairness and transparency, and the importance of the respective information to understanding the appraisal. Having regard to the decision of the Court of Appeal in Eisai v NICE, it must particularly strive to seek permission to disclose the economic model and/or the data contained therein".

In the context of the current appraisal, it is Janssen's position that NICE has not met those standards. It is impossible for consultees to judge what efforts have been made by NICE to obtain permission to disclose the "academic in confidence" (AIC) information relied upon by SHTAC, however the fact that NICE has refused to inform us even of the nature of the information or the detailed reasons for non-disclosure raise concerns. (We assume the AIC information relates to the data from the MRC Myeloma IX Study (paragraph 4.1.11 of the FAD) although it is unclear whether this is the only reason for NICE's refusal to disclose the model). Furthermore, NICE has provided no information as to the efforts, if any, which have been made to persuade the data owners to release the information, including whether they are willing to disclose it on the basis that access is limited to designated individuals, upon provision of confidentiality undertakings, as suggested in the Servier case. It is difficult to see why the holder of AIC information would refuse to disclose information on such a basis.

NICE's refusal to provide information regarding any efforts that have been made to persuade the data owners to agree to their release suggests that NICE have not, in fact, taken every reasonable step to secure agreement to disclosure. The result is an appraisal that is incapable of being tested or investigated and wholly lacks transparency. This is clearly unfair.

1.2. The Appraisal Committee's reasons for limiting use of bortezomib to patients who have contraindications to thalidomide, rather than those for whom thalidomide is clinically inappropriate are unexplained

Paragraph 1.2 of the FAD recommends use of bortezomib regimens in patients who are "unable to tolerate or [have] contraindications to thalidomide".

During the consultation process there was a agreement amongst consultees that this wording could be interpreted by the NHS to inappropriately limit the use of bortezomib to those patients for whom thalidomide is contraindicated as defined by the summary of product characteristics (SmPC) approved by the regulatory authority rather than those patients for whom thalidomide is clinically inappropriate (e.g. because they have coagulation disorders, recent ischaemic or thromboembolic events, neuropathy or impaired renal function). The SmPC for thalidomide contraindicates its use in patients with hypersensitivity to thalidomide or to any of the excipients, pregnant women, women of childbearing potential unless all the conditions of the Thalidomide Celgene Pregnancy Prevention Programme are met and patients unable to follow or comply with the required contraceptive measures.

In this instance, a group of patients would fall into a significant third group, who their clinicians would not want to use thalidomide in for clear clinical reasons, but for whom a specific intolerance or SmPC contraindication cannot be proved. A "postcode lottery" with local variable interpretations of what exactly the committee intended would inevitably follow. These "third group" of patients would be denied bortezomib and would ironically only have an option of receiving MP alone, which has been clearly demonstrated in this approasial to be less cost-effective than bortezomib. In other words, unclear wording would have the unintended consequence of forcing patients to receive a less effective and less cost-effective treatment.

Clinical expert groups along with patients' associations and Janssen recommended that this confusion could be avoided altogether through a minor change to the wording and by substituting 'contraindications' with 'clinically inappropriate'. Out of eleven stakeholders who submitted comments in response to the ACD, eight commented on section 1.2, all of whom supported this change of wording. These eight stakeholders are: The Royal College of Pathologists, British Society of Haematology, UK Myeloma Forum, Myeloma UK, Macmillan, Leukaemia Care, Royal College of Physicians, NHS Quality Improvement Scotland and Janssen. The other three stakeholders did not comment on this section at all.

There is a precedent in a recent appraisal where a committee has adopted a similar approach to the one proposed here. In the Topotecan NICE guidance (TA 184), the

committee recommended it as an option where retreatment with a first line agent was '*not considered appropriate*'.

The Appraisal Committee's response to the concerns expressed by consultees is provided at paragraph 4.3.11 of the FAD: "*The Committee heard different opinions for and against restricting the wording of the guidance around the contraindications to thalidomide and agreed that the contraindications specified in the SPC for thalidomide covered the safety risks adequately*". However, this is a misleading. As stated above, stakeholders who commented on this issue were unanimous in their support for this change and there were no discussions contradicting this consensus during Part 1 of the second appraisal committee meeting. The reasons why the committee chose to ignore these recommendations therefore remain unclear.

1.3. In deciding to place less weight on the thalidomide studies which included a "maintenance" phase, the Appraisal Committee relied on evidence from the Assessment Group which has not been disclosed to consultees

In considering the ACD for this appraisal the Appraisal Committee rejected those studies which included a thalidomide monotherapy "maintenance" phase for the purpose of its assessment of the clinical effectiveness of thalidomide. The reason for this approach was that the Committee concluded that thalidomide monotherapy is neither licensed treatment in this indication nor included in the Scope for this appraisal. Importantly the clinical outcomes were less favourable in studies where a "maintenance" phase was included as compared with those where there was no "maintenance" phase.

Janssen believed that this approach by the Appraisal Committee was incorrect and excluded important data from consideration by the Committee. We therefore submitted a graph to the Appraisal Committee (figure 1, under point 2.1 below, presented in section 2.3.2 of the cover letter responding to the ACD) which demonstrated that the principal differences in clinical outcome between the studies with a "maintenance" phase and those with no "maintenance" phase occurred during the early parts of the studies before the commencement of any maintenance treatment. Therefore in 5 of the 6 excluded studies, the highest hazard ratio in terms of overall survival occurs earliest, and this is before the

start of the maintenance phase. The heterogeneity in survival outcomes between the included and excluded studies therefore occurred during the MPT treatment phase, that is, before the commencement of the monotherapy phase and they should accordingly be considered an receive equal weight as studies with no "maintenance" phase.

However paragraph 4.1.20 of the FAD states that '*the assessment group commented on the additional evidence and stated that it was not possible to make conclusions about the relative effects of maintenance vs. 1st line treatment from the evidence submitted*'. No evidence is provided in the FAD or in any of the associated consultation documents made available to consultees exactly what the comments are from the Assessment Group were that support this statement and this appears inconsistent with the available data. This is unfair. We would like to be provided with the Assessment Group's reasons for stating that it is not possible to make a conclusion from the analyses presented by Janssen and given an opportunity to respond to them.

1.4. While the Appraisal Committee decided that the thalidomide studies which included a "maintenance" phase should receive less weight than those with no "maintenance" phase, such weighting is unexplained and appears inconsistent with NICE's Guide to the Methods of Technology Appraisal

In addition, the assertion that *"where possible (that is, when available for first-line treatment without maintenance), outcome data (for example, complete response) had been included in the Assessment Group's systematic review of clinical effectiveness"* (paragraph 4.3.4 of the FAD) is incorrect. In our comments on the ACD Janssen presented methods to allow the estimation of survival effects from these study phases, however these data have not been included in the Assessment Group's review. The Assessment Group and the Appraisal Committee provides no reasons why the results of such studies, at least up to the commencement of the maintenance phase, should not receive equal weight to other studies which do not include a maintenance phase.

NICE's Method Guide (section 5.3.9) states that '*Synthesis of outcome data through meta-analysis is appropriate provided there are sufficient relevant and valid data that use measures of outcome that are comparable'.* Meta-analysis provides a methodologically

robust approach to synthesizing data assigning weight to different studies based on the associated statistical uncertainty, and is clearly preferable to implicit assumptions based on an arbitrary and an un-transparent cognitive weighting process. It should be noted that given the treatment outcomes were in fact worse during the induction than the maintenance phases of the studies, whilst the weight given to these studies in an evidence synthesis will be reduced when only including data from these phases, the impact on the efficacy estimates for thalidomide may still be substantial.

In this context we would also like to highlight that even if a less formal weighting of study data were attempted, the period of induction treatment (combination treatment as per the SmPC) in these studies was relatively long with reference to the overall dosing period. For example Palumbo *et al* incorporated a 6-month induction period, whilst the median period of thalidomide dosing was 8 months, suggesting that the contribution of the induction outcomes vs the maintenance outcomes to the overall study estimates (and hence the appropriate weight) will in fact be high.

The relevant dataset on CTDa from the MMIX trial has been excluded from the systematic review of clinical effectiveness without any reference to these data in the FAD or any explanation. This point was also made by Clinical Trial Research Unit, who had provided data from this study in an academic in confidence submission, in their comments on the ACD. Importantly their submission included data for patients not randomised to maintenance and therefore highly relevant to the decision problem, which given the committees conclusion of equivalence between CTDa and MPT should have contributed to the overall estimate of benefit of thalidomide in combination with an alkylating agent. This point was made clearly in our comments on the ACD and such analysis is specified as part of the NICE methods (section 5.3.12) "*A group of related technologies, whether or not they are formally identified as part of a recognised 'class', might have similar but not necessarily identical effects. When the Institute is appraising a number of related technologies within a single appraisal, analyses based both on a class effect and individual effects should normally be undertaken, unless specified otherwise in the final scope for the appraisal.*"

In conclusion the Appraisal Committee does not appear to have followed NICE's procedures in giving the studies of MPT with no maintenance treatment (which enrolled only 550 of the

total 2018 patients enrolled in identified thalidomide studies) such a weight as to conclude that thalidomide or bortezomib in combination with an alkylating agent improve outcomes to a similar degree.

2. Ground 2: Perversity

2.1. The exclusion of critical evidence from thalidomide trials has resulted in a fundamentally flawed evidence synthesis that is not a sound basis for decision-making

In assessing overall survival associated with thalidomide treatment the Assessment Group excluded trials which included a thalidomide monotherapy "maintenance" phase following initial treatment with MPT. We have already stated and explained in our response to the ACD that the exclusion of these studies from the systematic evidence review was inappropriate (section 2.3.2 of the cover letter).

The FAD contains the committee's deliberations on this point stating *"The Committee noted that maintenance with thalidomide monotherapy after first-line treatment with a combination regimen did not fall within the appraisal scope. It also noted that, where possible (that is, when available for first-line treatment without maintenance), outcome data (for example, complete response) had been included in the Assessment Group's systematic review of clinical effectiveness"* and with respect to the survival data *"The Committee concluded (see section 4.3.10) that to assign studies (published and ongoing) in which the results were confounded by treatment outside the appraisal scope equivalent weight to the two key studies without maintenance treatment was not justified. Nevertheless it was prepared to bear in mind these data without over emphasising them".* Based on this the committee concludes that *"it was likely that bortezomib in combination with an alkylating agent and corticosteroid improved outcomes to a similar degree to thalidomide in combination with an alkylating agent and corticosteroid".*

We contend that the committee's conclusion is perverse in the light of the evidence presented for three reasons.

(a) Firstly, the committee's view that the maintenance studies fall outside the scope of the appraisal is incorrect as the MPT induction phase of these studies uses the MPT combination in a manner which is consistent with the license, in the appropriate patient population.

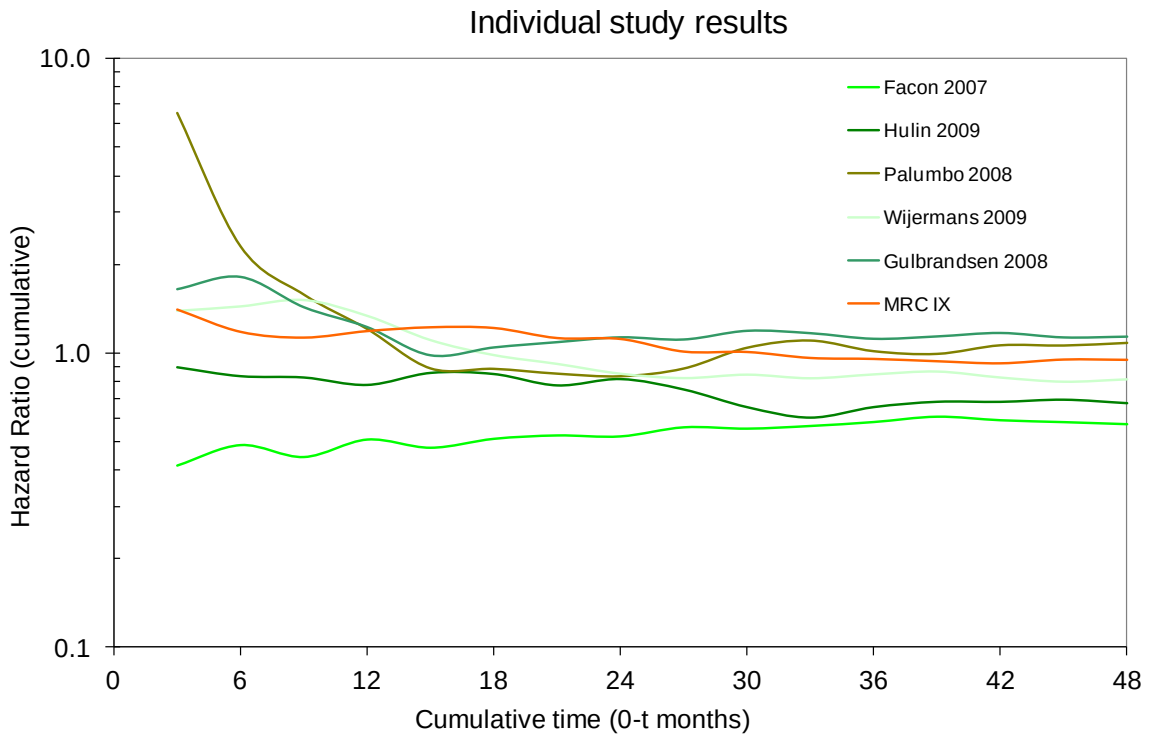
These data are relevant to the decision problem and clearly should be considered in the context of this appraisal.

(b) Secondly, the committee claims to have taken the data into account but not with "equivalent weight" to the two non-maintenance studies (paragraph 4.3.4), however to conclude that bortezomib and thalidomide based combinations "*improved outcomes to a similar degree*" the committee cannot have attached any materially important weight to the maintenance studies. Such an approach is wholly perverse in view of the information they contain to inform the comparison of MPT to MP (given that all of these studies included an 'induction' phase are clearly within the scope of the appraisal).

(c) Finally, having concluded that CTDa was clinically equivalent to MPT the data from the Myeloma IX trial is relevant to the overall efficacy conclusions of thalidomide in combination with an alkylating agent. In this study patients were randomised to either maintenance or no maintenance and thus the study clearly provides information relevant to the decision problem (more specifically the data from the 'induction' period and further follow up data from the arm that were randomised to the "no maintenance" regimen). We provide further details of these issues below.

Janssen submitted a graph to the Appraisal Committee (figure 1 below, presented in section 2.3.2 of the cover letter responding to the ACD) which shows that in 5 of the 6 excluded studies, the highest hazard ratio in terms of overall survival occurs earliest, and this is before the start of the maintenance phase. The heterogeneity in survival outcomes between the included and excluded studies therefore occurred during the MPT treatment phase, that is, before the commencement of the monotherapy phase. While the FAD states at paragraph 4.1.20 that '*the assessment group commented on the additional evidence and stated that it was not possible to make conclusions about the relative effects of maintenance vs. 1st line treatment from the evidence submitted',* the evidence and/or reasons of the Assessment Group have not been disclosed (see point 1.3 above) however Janssen does not accept that such an assessment is correct.

Figure 1. Cumulative hazard ratio for overall survival in individual studies



2.2. The Appraisal Committee have demonstrated a lack of consistency in considering clinical experts' opinion to inform its decision

Two combination regimes for thalidomide were considered in the course of this appraisal: melphalan, prednisolone and thalidomide (MPT) (the licensed combination) and cyclophosphamide, thalidomide and attenuated dexamethasone (CTDa) (which is not licensed). After considering the data, the Assessment Group concluded that CTDa was dominated by MPT (i.e. MPT is cheaper and more effective than CTDa) and in comparison with CTDa, a regimen of bortezomib, melphalan and prednisolone (VMP) produced an ICER of £28,907 per QALY gained (paragraph 4.2.18 of the FAD). While Janssen believes that this ICER is substantially too high it demonstrates that even with the extreme assumptions used by the assessment group that bortezomib is more cost-effective than CTDa within the range NICE is generally willing to recommend.

However, the Appraisal Committee made the decision to view the two thalidomide regimens as interchangable and to therefore recommend use of unlicensed CTDa alongside

MPT. This is justified on the basis of clinical opinion at the first committee meeting. The ICER for MPT vs. MP was applied to CTDa as it was considered to be more robust than the one for CTDa vs. MP due to the limited availability of data for the evidence synthesis. In coming to its conclusion, the Committee has disregarded the available large RCT dataset for CTDa which, for the reasons explained under 2.1 above, are actually directly relevant. While reference is made in the FAD to "past studies" mentioned by the clinical experts, which demonstrated similar effects of the two regimens before the addition of thalidomide", such studies are not identified. It is therefore impossible to determine whether the studies in question investigated comparable regimens to those under consideration in this appraisal and there is no indication that the studies were considered in detail or at all by the Committee.

While we broadly accept the committee's decision to use clinical opinion as the basis to justify the interchangeability of these regimens, we are concerned that the committee has then simply ignored clinical experts' opinion with regards to the appropriate wording in section 1.2 without any explanation or justification. Both should carry similar weight in the hierarchy of evidence, but in this instance, it appears that the committee have picked and chosen which pieces of clinical advice they have acted upon to support their decision instead of using clinical opinion to inform their decision in a systematic manner.

2.3. The Appraisal Committee's conclusion at paragraph 4.3.8 of the FAD that an assessment of cost-effectiveness which assumes use of 31.5 vials of bortezomib "should be considered the most optimistic estimate for clinical practice", is inconsistent with the available evidence.

The data for the effectiveness of bortezomib is taken from the VISTA trial in which patients received an average of 31.5 vials. This number was derived from the drug accounting records that are maintained as part of the Good Clinical Practice process for the conduct of registration studies. However in assessing the cost-effectiveness of bortezomib, the Assessment Group assumed that 52 vials were used, reflecting the maximum possible number of vials and treatment cycles permitted within the trial protocol. The Assessment Group identified that the Janssen model had lower costs and incorrectly stated that this was

because the company had assumed vial sharing. Janssen made NICE aware of this extremely significant error immediately on receipt of the assessment report.

With such a fundamental error, we respectfully requested that a new assessment report be issued to consultees which relied on a correct costing assumption. Unfortunately, and for reasons that were never justified, this was not forthcoming meaning that all stakeholders received an assessment report with basecase analyses for bortezomib that were fundamentally and inarguably flawed. Some effort was made to redress this by providing a "scenario analysis" with the correct number of vials, but this would inevitably carry less weight that the basecase analyses and meant that sensitivity analyses still relied on the original flawed basecase assumptions. It is important to understand that the figure of 31.5 vials reflected actual usage in the VISTA trial and is a result of doses being missed, doses being reduced and treatment being discontinued rather than vial sharing. The inclusion of fully loaded per protocol dosing of bortezomib in the cost-effectiveness analysis clearly had the substantial effect of artificially inflating the ICERs by exaggerating the costs of bortezomib to an enormous degree and completely at odd with what happened in the trial.

As stated above SHTAC did accept that fewer vials had been used in the VISTA trial and noted that other consultees had indicated that, in practice, patients receive fewer than the maximum number of cycles of treatment. While the Appraisal Committee accepted at paragraph 4.3.8 of the FAD that the Assessment Group's original analysis had assumed too many vials of bortezomib, the Committee referred to "*consultation comments from the Assessment Group*" and "*further discussion with both the manufacturer and the Assessment Group at the second meeting*" to form its view that "*the costs of delayed doses might still reflect clinical practice and need to be considered*". The Appraisal Committee therefore concluded that "*the manufacturer's preference for modelling 31.5 vials should be considered the most optimistic estimate for clinical practice*".

The Appraisal Committee's conclusion that use of 31.5 vials is "*the most optimistic estimate*" is simply incorrect. This actually should have been correctly reported to be the basecase assumption and one reflected in a re-issued, fit for purpose assessment report. It is our contention that the continued presentation throughout this consultation of a flawed basecase using deeply misleading costs of bortezomib gives the impression that this is

indeed in some way the most optimistic set of assumptions. As indicated in our response to the ACD (page 3 of the tabulated comments) as well as in the 2nd appraisal committee meeting, the use of 31.5 vials is not optimistic and indeed in clinical practice, the true ICER will be considerably lower than that presented in the submission. That is because the economic models assume single patient use of bortezomib vials, whilst in reality, we know that it is common practice to vial share bortezomib in the UK. It is our contention that the impact of vial sharing on the cost-effectiveness ratios should have been explored in the Assessment Group model and should have been considered by the Appraisal Committee.

2.4 Failure to consider vial sharing of bortezomib is inconsistent with the available evidence and with the approach followed in other appraisals

Janssen notes that there is a precedent for considering vial sharing within an economic evaluation of a cancer medicine by NICE. In the recent appraisal of trastuzumab for gastric cancer, the committee has accepted an assumption that vial sharing would systematically occur in 80% of centres. We further note that without acceptance of widespread vial sharing, this technology would have failed to have demonstrated cost-effectiveness. Janssen contends that it is inequitable for NICE to consider vial sharing in gastric cancer but not multiple myeloma.

Our own analyses suggest that vial sharing would have a significant impact on the cost-effectiveness results negating a large portion of the incremental cost of bortezomib in comparison to thalidomide.

REQUESTED ACTIONS

In the context of the above concerns, Janssen respectfully request the Appeal Panel to refer this appraisal back to the Appraisal Committee for further consideration with the following directions:

- Release of the economic model to fully participate in the consultation process
- Data identified in the Assessment Group's economic model and analyses should be explored in detail with the data owners and consultees provided with data to allow

consideration of this appraisal under confidentiality undertakings or other appropriate arrangements.

- The Appraisal Committee should be asked to re-consider the wording in section 1.2 . in line with the consensus reached by stakeholders
- Inclusion of the five MPT trials in the evidence synthesis of the survival endpoints and in the economic analysis. At minimum the MPT phase of these trials should be included in the revised economic analysis. ICERs provided by Janssen could be used by the Appraisal Committee
- Inclusion of the relevant large dataset of patients not receiving maintenance thalidomide in the MRC Myeloma IX.

THE DETERMINATION OF THIS APPEAL

Janssen requests an oral hearing for the determination of this appeal.

Janssen

September 2010

References

Facon,T., et al. Melphalan and prednisone plus thalidomide versus melphalan and prednisone alone or reduced-intensity autologous stem cell transplantation in elderly patients with multiple myeloma (IFM 99-06): a randomised trial. *Lancet*, 2007; 370, 1209-1218

Hulin,C., et al. (2009) Efficacy of Melphalan and Prednisone Plus Thalidomide in Patients Older Than 75 Years With Newly Diagnosed Multiple Myeloma: IFM 01/01 Trial. *J.Clin Oncol. 2009*

Kapoor et al 2009. Melphalan and Prednisone (MP) Versus Melphalan, Prednisone and Thalidomide (MPT) as Initial Therapy for Previously Untreated Elderly and/or Transplant Ineligible Patients with Multiple Myeloma: A Meta-Analysis of Randomized Controlled Trials ASH 2009 (Abstract 615)

Kumar et al. Thalidomide Versus Bortezomib-Based Regimens for Relapsed Myeloma: Meta-Analysis and Indirect Meta-Analysis. ASH 2008 (abstract 2362). *Although the title is in relapsed myeloma, the abstract is on front line treatment.*

Owen RG, Morgan GJ, Jackson H et al. MRC Myeloma IX: preliminary results from the non-intensive study. Abstract 547 presented at XII International Myeloma Workshop, 2009.

Palumbo,A., et al. Oral melphalan, prednisone, and thalidomide in elderly patients with multiple myeloma: updated results of a randomized controlled trial. *Blood 2008*, 112, 3107-3114.

Wijermans P. et al. Phase III study of the value of thalidomide added to melphalan plus prednisolone in elderly patients with newly diagnosed multiple myeloma.: the HOVON 49 study. J Clin Oncol. 2010. Published ahead of print on June 1, 2010 as 10.1200/JCO.2009.26.1610

Waage A. et al. Melphalan and prednisolone plus thalidomide or placebo in elderly patients with multiple myeloma. Blood 2010. Prepublished online May 6, 2010; DOI10.1182/blood-2009-08-237974

Waage A, Palumbo AP, Fayers P, et al. MP versus MPT for previously untreated elderly patients with multiple myeloma: a meta-analysis of 1,682 individual patient data from six randomized clinical trials [abstract]. J Clin Oncol. 2010;28(suppl):15s Abstract 8130.