NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

SINGLE TECHNOLOGY APPRAISAL

Lenalidomide for the treatment of multiple myeloma in people who have received at least one prior therapy with bortezomib (partial review of TA171) [ID667]

The following documents are made available to the consultees and commentators:

1. **Appraisal Consultation Document (ACD)** as issued to consultees and commentators in July 2014

2. **Consultee and commentator comments on the Appraisal Consultation Document** from:
   - Myeloma UK
   - UK Myeloma Forum
   - Janssen

3. **Patient Access Scheme submission** from the company, Celgene (submitted for the April 2016 Committee meeting)

4. **Evidence Review Group addendum** from the ERG, Peninsular Technology Assessment Group
   - Addendum
   - Additional analyses

5. **Request for additional evidence** as sent to Celgene (in June 2016)

6. **Celgene's response to the request for additional information**

7. **Evidence Review Group addendum critiquing the additional information from Celgene (**prepared in September 2016)

8. **Celgene's commentary on the Evidence Review Group critique of additional information**

9. **Evidence Review Group addendum responding to Celgene's commentary** (prepared in October 2016)

Any information supplied to NICE which has been marked as confidential, has been redacted. All personal information has also been redacted.

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Appraisal consultation document

Lenalidomide for treating multiple myeloma after 1 prior treatment with bortezomib (part-review of TA171)

The Department of Health has asked the National Institute for Health and Care Excellence (NICE) to produce guidance on using lenalidomide in the NHS in England. The Appraisal Committee has considered the evidence submitted by the manufacturer and the views of non-manufacturer consultees and commentators, and clinical specialists and patient experts.

This document has been prepared for consultation with the consultees. It summarises the evidence and views that have been considered, and sets out the draft recommendations made by the Committee. NICE invites comments from the consultees and commentators for this appraisal (see section 10) and the public. This document should be read along with the evidence base (the evaluation report).

The Appraisal Committee is interested in receiving comments on the following:

- Has all of the relevant evidence been taken into account?
- Are the summaries of clinical and cost effectiveness reasonable interpretations of the evidence?
- Are the provisional recommendations sound and a suitable basis for guidance to the NHS?
- Are there any aspects of the recommendations that need particular consideration to ensure we avoid unlawful discrimination against any group of people on the grounds of race, gender, disability, religion or belief, sexual orientation, age, gender reassignment, pregnancy and maternity?

Note that this document is not NICE's final guidance on this technology. The recommendations in section 1 may change after consultation.

After consultation:

The Appraisal Committee will meet again to consider the evidence, this appraisal consultation document and comments from the consultees.

At that meeting, the Committee will also consider comments made by people who are not consultees.

After considering these comments, the Committee will prepare the final appraisal determination (FAD).

Subject to any appeal by consultees, the FAD may be used as the basis for NICE's guidance on using lenalidomide in the NHS in England.

For further details, see the Guides to the technology appraisal process.

The key dates for this appraisal are:

Closing date for comments: 21 August 2014

Third Appraisal Committee meeting: 17 September 2014

Details of membership of the Appraisal Committee are given in section 9, and a list of the sources of evidence used in the preparation of this document is given in section 10.

Note that this document is not NICE's final guidance on this technology. The recommendations in section 1 may change after consultation.

1 Appraisal Committee's preliminary recommendations

1.1 Lenalidomide in combination with dexamethasone is not recommended for treating multiple myeloma in people:

- whose condition has relapsed for the first time **and**
- who have had 1 prior treatment with bortezomib **and**
- for whom thalidomide is contraindicated or cannot be tolerated **and**
- for whom stem cell transplantation is not appropriate.

1.2 People currently receiving treatment initiated within the NHS with lenalidomide that is not recommended for them by NICE in this guidance should be able to continue treatment until they and their NHS clinician consider it appropriate to stop.

2 The technology

2.1 Lenalidomide (Revlimid) is a derivative of thalidomide and has immunomodulatory, anti-neoplastic, anti-angiogenic and pro-erythropoietic activity. It is administered orally. Lenalidomide in combination with dexamethasone has a marketing authorisation for treating multiple myeloma in adults who have received at least 1 prior therapy.

2.2 The summary of product characteristics includes the following adverse effects for lenalidomide: neutropenia, anaemia and thrombocytopenia. Because lenalidomide is structurally related to thalidomide, a known human teratogen that causes severe birth

defects, a risk minimisation plan has been developed and agreed with the Medicines and Healthcare products Regulatory Agency to avoid fetal exposure to lenalidomide. For full details of adverse reactions and contraindications, see the summary of product characteristics.

2.3 Lenalidomide is available as a 21-capsule pack. The cost per pack varies according to capsule size: £3570 (5 mg), £3780 (10 mg), £3969 (15 mg) and £4368 (25 mg; excluding VAT; British National Formulary edition 65). The recommended starting dose is 25 mg orally once daily on days 1–21 of repeated 28-day cycles. Costs may vary in different settings because of negotiated procurement discounts.

3 The manufacturer's submission

The Appraisal Committee (section 9) considered evidence submitted by the manufacturer of lenalidomide and a review of this submission by the Evidence Review Group (ERG; section 10). The purpose of the manufacturer's submission was to demonstrate the effectiveness and cost effectiveness of lenalidomide compared with bortezomib, bortezomib plus dexamethasone, bendamustine or chemotherapy in adults who have been treated with bortezomib first line, and so for adults in whom treatment with thalidomide or stem cell transplant was not appropriate (see NICE technology appraisal guidance 228, Bortezomib and thalidomide for first-line treatment of multiple myeloma) and who would consider second-line treatment.

3.1 The manufacturer identified, using a systematic literature review, 2 identically designed randomised controlled trials that compared lenalidomide plus dexamethasone with placebo plus dexamethasone (MM-009 and MM-010). The manufacturer did not identify any randomised controlled trials that directly compared

lenalidomide with the comparators defined in the scope (bortezomib, bendamustine or chemotherapy, which included regimens based on melphalan, vincristine, cyclophosphamide and doxorubicin). The manufacturer presented data from the following studies, identified through a systematic literature review, to estimate the efficacy of bortezomib and bendamustine:

- **Bortezomib:** The manufacturer identified 6 observational studies. Taverna et al. (2012), a small retrospective survey of people who had received bortezomib re-treatment, was selected as the most relevant (see sections 3.13 and 3.14).
- **Bendamustine:** The manufacturer identified 1 retrospective observational study that included people with relapsed or refractory multiple myeloma after prior treatment (Damaj et al. 2012; see sections 3.15 and 3.16).

Lenalidomide plus dexamethasone compared with placebo plus dexamethasone (MM-009 and MM-010)

Clinical effectiveness

3.2 MM-009 and MM-010 compared treatment with lenalidomide plus dexamethasone with placebo plus dexamethasone in people with multiple myeloma who had received at least 1 prior therapy. The prior therapy did not necessarily include bortezomib (and was therefore different to the population set out in the decision problem). The dose schedule on the trial was a once-daily starting dose of 25 mg oral lenalidomide or placebo on days 1–21 of each 28-day cycle, and a daily dose of 40 mg oral dexamethasone on days 1–4, 9–12 and 17–20 for the first 4 cycles. After the fourth cycle, 40 mg of dexamethasone was administered on days 1–4 only. Treatment was continued until the disease progressed, unless treatment was stopped because of adverse reactions. Those who had placebo plus dexamethasone were offered lenalidomide when

they stopped treatment or when the study was unblinded (that is, they could cross over). The study designs of MM-009 (n=353) and MM-010 (n=351) were identical other than study location (MM-009 was carried out in the USA and Canada, whereas MM-010 was carried out in Australia, Europe [including the UK] and Israel). The primary outcome was time to disease progression. Secondary outcomes included overall survival, progression-free survival (defined as time from randomisation to disease progression or death), response rates, adverse reactions and time to decrease in performance status. Randomisation was stratified according to the serum concentration of beta-2 microglobulin, whether or not the patient had previously had stem cell transplantation, and to the number of previous anti-myeloma therapies (1 compared with 2 or more). About 35% had received only 1 prior therapy (first relapse) and about 65% had received at least 2 prior therapies. Only 2 people in the combined MM-009 and MM-010 trials reflected the population in this appraisal, that is, 1 prior treatment only, with bortezomib. Exclusion criteria included people previously treated with lenalidomide or whose disease was refractory to dexamethasone. The investigators assessed response using the European Group for Blood and Marrow Transplantation criteria. In its submission, the manufacturer identified several 'unspecified' (post hoc) subgroup analyses using the pooled populations of both trials, including (but not limited to) whether prior treatment with thalidomide or bortezomib had been received. The median age of the population in both studies was between 62 years and 64 years.

3.3 The manufacturer presented primary and secondary outcomes of the MM-009 and MM-010 trials analysed individually and as a pooled dataset using the intention-to-treat population, and intention-to-treat analyses. The absolute values observed in the lenalidomide plus dexamethasone group informed the modelling (see section 3.20), rather than the relative effectiveness of

lenalidomide plus dexamethasone compared with placebo plus dexamethasone. At the end of the study ('unblinding'), randomisation to lenalidomide plus dexamethasone was associated with a statistically significantly reduced median time to progression compared with placebo plus dexamethasone (48.1 weeks compared with 20.1 weeks, p<0.001 in MM-009; and 48.7 weeks compared with 20.1 weeks, p<0.001 in MM-010). In addition, lenalidomide plus dexamethasone was associated with a statistically significantly increased median progression-free survival compared with placebo plus dexamethasone (41.1 weeks compared with 20.1 weeks, p<0.001 in MM-009; and 'not yet reached' compared with 20.1 weeks, p<0.001 in MM-010). In an extended follow-up (median 48 months), lenalidomide plus dexamethasone was associated with statistically significantly increased median overall survival compared with placebo plus dexamethasone (29.6 months compared with 20.2 months, p<0.001 in MM-009; and 'not estimable' – because too few people in the treatment arm had died for it to be possible to estimate the median – compared with 20.6 months, p<0.05 in MM-010).

3.4 The manufacturer presented pooled analyses of MM-009 and MM-010. Lenalidomide plus dexamethasone was associated with a statistically significant longer median time to progression (13.4 months compared with 4.6 months, p<0.001) and median progression-free survival (11.1 months compared with 4.6 months, p<0.001) compared with placebo plus dexamethasone at unblinding (median 17.5 months; n=704). The manufacturer presented pooled data for overall survival after a median follow-up of 48 months. Overall survival was statistically significantly longer with lenalidomide plus dexamethasone than placebo plus dexamethasone (median of 38.0 months compared with 31.6 months respectively; p=0.045).

Health-related quality of life

3.5 The MM-009 and MM-010 trials did not collect information on health-related quality of life. Instead, the manufacturer presented Eastern Cooperative Oncology Group-performance status (ECOG-PS) data from the trials and it used this to assess progression of disease and impact on daily living abilities. The median time to 'first worsening' was statistically significantly longer with lenalidomide plus dexamethasone (36.3 weeks) than with placebo plus dexamethasone (12.1 weeks in MM-009; p=0.012). The median time to first worsening with lenalidomide plus dexamethasone was shorter than with placebo plus dexamethasone in MM-010 (lenalidomide plus dexamethasone: 10.1 weeks; placebo plus dexamethasone: 12.3 weeks). However, this difference was not statistically significant (p=0.271).

Adverse events

3.6 The manufacturer presented data on adverse events from MM-009 and MM-010 separately and pooled. The most common adverse effects associated with lenalidomide plus dexamethasone were haematological. The manufacturer suggested that clinicians could manage these by reducing the dose. Anaemia, neutropenia, thrombocytopenia, constipation, pneumonia, decreased weight, hypokalaemia, hypocalcaemia, tremor, rash and deep vein thrombosis were reported considerably more frequently in the lenalidomide plus dexamethasone group than in the placebo plus dexamethasone group. There was an increased risk of developing thromboembolic adverse events (deep vein thrombosis, pulmonary embolism) with lenalidomide plus dexamethasone compared with placebo plus dexamethasone (9.1% compared with 4.3%, and 4.0% compared with 0.9% respectively).

3.7 The manufacturer stated that bortezomib and thalidomide can be associated with peripheral neuropathy, whereas pooled analyses of

MM-009 and MM-010 showed that lenalidomide plus dexamethasone neither increased nor decreased the risk of peripheral neuropathy compared with placebo plus dexamethasone.

3.8 Lenalidomide is structurally related to thalidomide, a known human teratogen that causes severe birth defects. As such, a plan was developed and agreed with the Medicines and Healthcare products Regulatory Agency to avoid fetal exposure to lenalidomide.

Subgroups

3.9 The manufacturer presented 2 subgroup analyses from MM-009 and MM-010. The first analysis was a post hoc comparison of lenalidomide plus dexamethasone with placebo plus dexamethasone after 1 prior therapy. The second analysis was a pre-specified comparison of lenalidomide plus dexamethasone after 1 prior therapy with lenalidomide plus dexamethasone after 2 or more prior therapies (Stadtmauer et al. 2009).

3.10 The first subgroup analysis showed that, in the MM-009 trial, after 1 prior therapy only, lenalidomide plus dexamethasone was associated with a statistically significantly longer median progression-free survival (16.6 months, 95% confidence interval [CI] 11.0 to 36.8 months) compared with placebo plus dexamethasone (4.6 months, 95% CI 4.0 to 5.7 months; hazard ratio [HR] 0.3, 95% CI 0.19 to 0.47; $p<0.0001$). Similarly, in MM-010, after 1 prior therapy, lenalidomide plus dexamethasone was associated with a statistically significantly longer median progression-free survival (13.3 months, 95% CI 5.1 to 26.9 months) compared with placebo plus dexamethasone (4.5 months, 95% CI 2.8 to 5.6 months, HR 0.39, 95% CI 0.24 to 0.62; $p<0.0001$). There were no statistically significant differences in median overall survival in MM-009 or MM-010 for lenalidomide plus

dexamethasone compared with placebo plus dexamethasone ($p>0.05$).

3.11 Stadtmauer et al. (2009) looked at the effectiveness of lenalidomide at different lines of therapy. Compared with lenalidomide plus dexamethasone after 2 or more prior therapies (n=220), lenalidomide plus dexamethasone after 1 prior therapy (n=133) had longer:

- median time to progression: 17.1 months compared with 10.6 months (HR 0.68, 95% CI 0.48 to 0.97; p=0.026)
- median progression-free survival: 14.1 months compared with 9.5 months (HR 0.71, 95% CI 0.2 to 0.99; p=0.047)
- median overall survival from study enrolment: 42.0 months compared with 35.8 months; p=0.041.

3.12 The manufacturer identified the following factors that increased the risk of death among people in MM-009 and MM-010 ($p<0.05$):

- treatment with placebo versus lenalidomide
- high versus low percent of plasma cells in the bone marrow
- high versus low serum concentrations of beta-microglobulin
- shorter versus longer duration of myeloma
- more versus fewer previous therapies for myeloma
- earlier versus later dexamethasone therapy
- higher versus lower international staging system.

The manufacturer also identified the following predictors of progression and treatment failure: beta-2 microglobulin; time since diagnosis of multiple myeloma; number of prior therapies; baseline presence or absence of bone lesions; and ECOG performance score.

Bortezomib

Clinical effectiveness and adverse events

3.13 The manufacturer presented data on the efficacy of bortezomib from 6 single-arm observational studies. The manufacturer presented a retrospective review of patient medical records by Taverna et al. (2012), which included 42 patients across 26 centres in Switzerland. People whose records were surveyed for the study had multiple myeloma that had responded to initial bortezomib therapy, but whose disease had subsequently progressed or relapsed, and were then re-treated with bortezomib, although not necessarily in second line. The study inclusion criteria specified that the initial treatment with bortezomib therapy had achieved complete response, near complete response or partial response, and that people had completed a re-treatment regimen with bortezomib after their disease had relapsed or progressed. Of those in the study, 31% had previously received a stem cell transplant, 12 people had received a different therapy between bortezomib treatments, and a third of people did not receive dexamethasone when re-treated with bortezomib. People had received a median of 2 prior therapies (range 1 to 11). The study end points were listed as complete response, near complete response and partial response, but were not clearly defined. How frequently patients had been followed was not defined, and results were not presented as a Kaplan–Meier plot.

3.14 The Taverna et al. (2012) study showed:

- Median time to progression after bortezomib re-treatment was 10.5 months (range 0.4 to 39.5 or more months).
- Median overall survival from first diagnosis was 9.3 years, after prior bortezomib 3.5 years and after bortezomib re-treatment 1.7 years.

Bendamustine

Clinical effectiveness

3.15 The manufacturer identified a retrospective study to show the clinical effectiveness of bendamustine in relapsed or refractory multiple myeloma (n=110, Damaj et al. 2012). The study consisted of a review of medical records of people who had been treated with bendamustine in a compassionate-use programme in France, and therefore had no control group. People whose records were reviewed had been previously treated with all of the following: alkylators, corticosteroids, immunomodulatory drugs and bortezomib. The study evaluated the response rate to bendamustine, the duration of response, progression-free survival and overall survival. Overall survival was calculated from the first dose of bendamustine and progression-free survival included death from any cause or progression as events.

3.16 The Damaj et al. (2012) study showed that after treatment with bendamustine, at a median follow-up of 10 months, 49 people who had received bendamustine had died after progression (or from other causes related to myeloma). Median progression-free survival was 9.3 months and median overall survival was 12.4 months.

Evidence Review Group critique of clinical effectiveness

3.17 The ERG reviewed the manufacturer's literature review and concluded that it was broadly suitable. The ERG reviewed the designs of MM-009 and MM-010 and concluded that they were of high quality. In addition, the ERG noted several issues:

- Both trials started in 2004. Since then, management of multiple myeloma has changed (specifically, the introduction of bortezomib) and therefore the trials do not reflect current clinical practice.

- The mean age of people was lower (63 years) than that of people in clinical practice with multiple myeloma in the UK (around 70 years).
- The proportion of people who received 2 or 3 prior therapies was higher than the proportion who had 1 prior therapy only.

3.18 The ERG commented that the manufacturer did not present time to treatment failure, an end point of MM-009 and MM-010, in the clinical section of its submission, although the manufacturer included it in its model.

Cost-effectiveness evidence

3.19 The cost-effectiveness evidence presented by the manufacturer consisted of a systematic literature review and a Markov model developed by the manufacturer. The manufacturer's systematic review did not identify relevant cost-effectiveness studies. The manufacturer's model compared lenalidomide with bortezomib re-treatment for people with multiple myeloma for whom stem cell transplant was not suitable, for whom thalidomide was contraindicated, and who had received 1 prior therapy only with bortezomib. The model had 3 health states: pre-progression on treatment, pre-progression off treatment, and post-progression, plus death. The manufacturer chose a lifetime time horizon (25 years), a cycle length of 28 days, and discount rate of 3.5% for costs and quality-adjusted life years (QALYs). The model took an NHS and personal social services perspective. Although a patient access scheme is available for bortezomib after 1 prior therapy, the manufacturer did not include this in the model's base case.

3.20 The transition probabilities between states in the model were based on overall survival, progression-free survival and time to treatment failure curves. To extrapolate beyond the period observed in studies, the manufacturer presented 6 parametric distributions

(exponential, Weibull, log-logistic, log-normal, Gompertz and gamma) for overall survival, progression-free survival, and time to treatment failure for lenalidomide plus dexamethasone. The manufacturer used a log-logistic distribution in the model base case for progression-free survival and time to treatment failure, and used an exponential piecewise distribution for overall survival. The manufacturer considered other distributions in scenario analyses. The manufacturer calculated transition probabilities between health states as follows:

- All people started the model in the 'progression free on treatment' state.
- Transition into the 'progressive disease' state from either the 'pre-progression on treatment' or 'pre-progression off treatment' states was determined by progression-free survival.
- Transition from 'pre-progression on treatment' to 'pre-progression off treatment' was determined by the difference between progression-free survival and time to treatment failure; that is, those people whose disease had not progressed, but for whom treatment had failed because of adverse events.
- Transition from any health state to death was determined by overall survival.
- **Lenalidomide:** The manufacturer estimated progression-free survival, time to treatment failure and overall survival from the lenalidomide arm of the MM-010 dataset (see section 3.22 for further details).
- **Comparators:** The manufacturer derived a hazard ratio reflecting the effectiveness of lenalidomide compared with bortezomib. The effectiveness of lenalidomide was derived from the lenalidomide plus dexamethasone arm of the MM-010 data. The manufacturer took estimates of absolute progression-free survival and overall survival of bortezomib from the literature, as described. The manufacturer then approximated a crude hazard

ratio between lenalidomide and bortezomib by comparing the median progression-free survival or overall survival estimates between the studies, using Taverna et al. (2012) in the base case. To derive overall survival, progression-free survival and time to treatment failure curves for bortezomib, the manufacturer adjusted the lenalidomide overall survival, progression-free survival and time to treatment failure curves using these crude hazard ratios. The manufacturer used the hazard ratio calculated for progression-free survival for time to treatment failure. It employed the same approach to estimate the relative effectiveness of lenalidomide compared with bendamustine. The manufacturer assumed that chemotherapy and bendamustine were equally effective. In the 'progressive disease' health state, in the comparator arm only, people received lenalidomide as a third-line treatment in the manufacturer's initial base case. This affected overall survival in that, when the third-line treatment was lenalidomide, transition to death was based on the MM-010 trial overall survival data.

Table 1 Efficacy estimates for lenalidomide compared with bortezomib re-treatment, or bendamustine, included in the initial base-case model

Variable	Evidence source	Hazard ratio (>1 favours lenalidomide)
Re-treatment with bortezomib		
Overall survival	Taverna et al. 2012[a] (n=42; Switzerland) • Median (range) prior therapies: 2 (1–11) • Prior treatment with bortezomib: 100%	1.70
	White et al. 2013 (n=53; USA and Canada) • Number of prior therapies: 1 prior therapy 49%; 2+ prior therapies 51% • Prior treatment with bortezomib: 19%	1.42
Progression-free survival	Taverna et al. 2012[a] (as above)	1.15
	White et al. 2013 (as above)	1.76
	Hrusovsky et al. 2010 (n=60; Germany and Switzerland) • Median (range) prior therapies: 3.7 (1–14) • Prior treatment with bortezomib: 100%	1.09
	Dispenzieri et al. 2010 (n=7; USA) • Median (range) prior therapies: not reported • Prior treatment with bortezomib: 100%	1.28
	Petrucci et al. 2013 (n=50; Europe; time to progression used as a proxy for progression-free survival) • Number of prior therapies: 1 prior therapy 12%; 2+ prior therapies 88% • Prior treatment with bortezomib: 100%	1.26
	Min et al. 2007 (n=57; South Korea) • Median (range) prior therapies: not reported: 2 (1–3) • Prior treatment with bortezomib: 100%	0.84
Bendamustine (and chemotherapy agents)		
Overall survival	Damaj et al. 2012 (n=110; France) • Median (range) prior therapies: 4 (1–9) • Prior treatment with bortezomib: 100%	3.00
Progression-free survival	Damaj et al. 2012 (as above)	1.09

[a] Base-case values

n=number of people in study

3.21 'Modelled' patients entered the model having received bortezomib as a first-line treatment, and then received second-line treatment with lenalidomide plus dexamethasone, or with the comparator, until their disease progressed (modelled as the duration of the pre-progression on-treatment health state), or until stopping treatment for other reasons (modelled as the duration of the pre-progression off-treatment health state). The manufacturer used bortezomib as the comparator in its base case. The manufacturer explored in scenario analyses alternative second-line treatments, including bendamustine and other chemotherapy agents. The manufacturer's model also included the possibility that patients go on to receive third- and fourth-line treatments. For third- and fourth-line treatments, the manufacturer developed a best supportive care mix (that is, a miscellaneous 'basket' of standard chemotherapy), which depended on previous treatments and included:

- bortezomib
- dexamethasone
- melphalan
- cyclophosphamide
- cisplatin
- doxorubicin
- etoposide
- prednisolone
- prednisone
- lenalidomide.

3.22 The manufacturer estimated overall survival and progression-free survival from MM-010 rather than from combining the 2 studies, because MM-010 included a European population. The manufacturer stated that pooling the trials was not feasible because it would break randomisation. After a request for clarification by the ERG, the manufacturer provided scenario analyses that used MM-

009 only data or combined MM-009 and MM-010 results using mean outcomes weighted by the number of people in the trials. The manufacturer obtained absolute values for overall survival and progression-free survival associated with re-treating with bortezomib from the Taverna et al. (2012) study (see sections 3.13 and 3.14). The manufacturer adjusted the MM-010 patient data for factors that could influence outcomes, so that the 'modelled patients' from MM-010 better resembled the patients in the Taverna et al. study. Overall survival was adjusted for patient characteristics (concentration of beta-2 microglobulin, ECOG, and presence or absence of bone lesions) in the manufacturer's base case. However, the manufacturer did not adjust the rates of progression-free survival in MM-010 to reflect the patients in Taverna et al. re-treated with bortezomib. A scenario that included progression-free survival adjusted for duration of myeloma was provided by the manufacturer in response to a request for clarification from NICE. The manufacturer used further sources of overall survival and progression-free survival for bortezomib in scenario analyses (White et al. 2013, Petrucci et al. 2013, Hrusovsky et al. 2010, Dispenzieri et al. 2010 and Min et al. 2007; see table 1 for details). The manufacturer used the hazard ratios to adjust the extrapolated lenalidomide overall survival, progression-free survival and time to treatment failure curves (see section 3.20), to derive the comparator curves.

3.23 The manufacturer accounted for adverse events during treatment in the model, taking adverse events in the pre-progression health states from Lenalidomide for the treatment of multiple myeloma in people who have received at least one prior therapy (NICE technology appraisal guidance 171). The manufacturer estimated the event rates associated with lenalidomide plus dexamethasone from MM-010, and for bortezomib from the VISTA trial, which compared bortezomib plus melphalan and prednisone with

melphalan, and prednisone. The manufacturer used VISTA's melphalan and prednisolone arm as a proxy to estimate the adverse events rates for all other types of chemotherapy. The manufacturer estimated an average rate and applied it to each model cycle. The manufacturer assumed that people do not experience any adverse events after their disease progresses (when in the post-progression health state), but did assume that people in the comparator arm who receive lenalidomide plus dexamethasone as a third-line treatment experience adverse events. Adverse events in the model included anaemia, constipation, diarrhoea, deep vein thrombosis, hypercalcaemia, neutropenia, peripheral neuropathy, pneumonia and thrombocytopenia.

3.24 To estimate utility, the manufacturer used utility values from the literature because health-related quality-of-life data were not collected in the MM-009 and MM-010 clinical trials or in the observational studies used to describe the comparator treatments. Of the 8 studies identified by the manufacturer, only 2 (Khanna et al. 2006 [SF-36, USA], van Agthoven et al. 2004 [EQ-5D, Netherlands]) undertook primary data collection. The manufacturer modelled utility values from the study by van Agthoven et al. The value for pre-progression utility was 0.81 decreasing after 2 years to 0.77. The value for post-progression utility was 0.64. The manufacturer adjusted the utility values for age, based on published UK EQ-5D population norms. The manufacturer included decrements in utility associated with adverse events to each model cycle. The manufacturer applied these in the pre-progression health state for all treatments, as well as in the post-progression health state for patients receiving lenalidomide plus dexamethasone third line (in the comparator arm). The utility decrements associated with each treatment per model cycle were

0.013 for lenalidomide plus dexamethasone, 0.033 for bortezomib, and 0.025 for bendamustine or chemotherapy agents.

3.25 The manufacturer's model included costs associated with treatment, resource use and adverse events (see table 2). The manufacturer obtained acquisition costs for each treatment from the Department of Health's electronic Market Information Tool (eMIT) and the British National Formulary. The cost of lenalidomide included dose reductions and treatment interruptions based on the experience of patients in MM-010. Bortezomib has a patient access scheme, whereby the manufacturer of bortezomib reimburses the NHS for people whose condition does not respond to treatment. The manufacturer did not account for this in its base case; however, it did provide a scenario analysis that assumed a discount of 15%. The manufacturer based the cost of each adverse event on where a patient received treatment (inpatient, hospital day case, outpatient and general practice). The model included a weighted average of adverse events costs. Other costs were obtained from NHS reference costs.

Table 2 List of health states and associated costs in the economic model

Health state and item	Lenalidomide arm (cost per cycle)	Base-case comparator arm, bortezomib (cost per cycle)
Pre-progression		
Technology	Lenalidomide: £3773 Dex (cycles 1–4): £7.76 Dex (cycles 5+): £2.59	£4067.30
Concomitant G-CSF and administration	£473.62	n/a
Monitoring and tests	£153.34	
Administration	£161.85 (first cycle only)	£1065.76
Transport	£6.39 (first cycle only)	£17.04
Adverse events	£17.11	£29.26
Additional monitoring for lenalidomide (annual rate)	£824.26	n/a
Post-progression		
3rd line treatment	£70.20 IV administration: £69.63 Transport: £3.06 Duration: maximum 4 cycles (17.2 weeks)	£3772.88 IV administration: £0.00 Transport: £0.00 Duration: Lenalidomide PFS from MM-010 Adverse events with lenalidomide: £17.11
4th line treatment	Therapy: £2277.28 IV administration: £0.00 Transport: £0.00 Duration: maximum 4 cycles (16.8 weeks)	
Monitoring and tests	£175.86	
Terminal care	£1235 on death	
Alternative 3rd line scenarios		
Changing treatment composition after bortezomib 2nd line	n/a	Drug cost £1716.99 IV administration: £49.45 Transport: £2.20
Changing treatment composition after other 2nd line comparators	n/a	£2592.00 IV administration: £203.00 Transport: £3.25
Abbreviations: Dex, dexamethasone; G-CSF, granulocyte-colony stimulating factor; IV, intravenous; PFS, progression-free survival.		

Results of the economic analyses

3.26 The manufacturer presented initial results in its submission. However, the ERG noted several problems with the manufacturer's initial base-case modelling, and requested, and received, more analyses from the manufacturer. Later, following the concerns noted by the Committee documented in the appraisal consultation document and the consultation period, the manufacturer submitted further analyses. The analyses received pre-consultation are termed as follows: 'base case' (reflecting initial modelling); 'base-case A' (modelling after a first round of clarification); and 'base-case B' (modelling after a second round of clarification). The analyses received post-consultation are referred to as 'base-case C'. Each set of the analyses is presented below.

Manufacturer's initial analyses and results before clarification and before consultation

3.27 The manufacturer presented a deterministic base-case incremental cost-effectiveness ratio (ICER), comparing lenalidomide plus dexamethasone with bortezomib re-treatment, of £14,535 per QALY gained (incremental costs £7682, incremental QALYs 0.53), and a mean probabilistic ICER of £13,930 per QALY gained. The probabilities of lenalidomide being cost effective at £20,000 and £30,000 per QALY gained were 58.6% and 73.1% respectively. The manufacturer's deterministic sensitivity analysis showed that the ICER was most sensitive to the hazard ratio for overall survival (reflecting the relative effectiveness of lenalidomide compared with bortezomib). Other key drivers of the ICERs included which parametric function for overall survival and time to treatment failure the manufacturer chose to apply.

3.28 The manufacturer presented several scenario analyses, changing the time horizon, comparator, efficacy of lenalidomide and efficacy of the comparator. The manufacturer also used different discount

levels for the price of bortezomib to approximate the complex bortezomib patient access scheme (for which the manufacturer rebates the full cost of bortezomib for people who, after a maximum of 4 cycles of treatment, have less than a partial response). The ICERs ranged from dominant (time horizon reduced from 25 years to either 5 years or 10 years; curve fit for overall survival changed from piecewise exponential to Weibull; source of progression-free survival efficacy changed from Taverna et al. (2012) to Min et al. (2007) to £56,274 per QALY gained (log–normal curve used for progression-free survival and time to treatment failure) for lenalidomide compared with bortezomib re-treatment. Other scenarios submitted by the manufacturer included:

- Reducing the proportion of people receiving lenalidomide as a third-line treatment in the comparator arm. This increased the ICER to £38,330 per QALY gained for lenalidomide compared with bortezomib re-treatment.
- Including a patient access scheme for bortezomib. This increased the ICER to between £21,885 and £27,898 per QALY gained for lenalidomide compared with bortezomib re-treatment.
- Comparing lenalidomide plus dexamethasone with:
  - bendamustine plus prednisolone. This resulted in an ICER of £80,108 per QALY gained.
  - melphalan plus prednisone. This resulted in an ICER of £60,246 per QALY gained.
  - high-dose cyclophosphamide plus dexamethasone. This resulted in an ICER of £67,660 per QALY gained.
  - a blended comparator (that is, a mixture of bortezomib, lenalidomide, bendamustine and standard chemotherapy; the manufacturer assigned proportions based on market share). This resulted in an ICER of £32,462 per QALY gained.

Manufacturer's revised analyses after clarification round 1 (base-case A)

3.29 In its original submission, when estimating the progression-free survival hazard ratio to determine the relative effectiveness of lenalidomide and bortezomib in the model, the manufacturer had not adjusted the data from MM-010 to reflect the characteristics of the patients from Taverna et al. (2012; see section 3.20), as it had done for overall survival. Therefore, in response to the first round of clarification between the manufacturer and the ERG, the manufacturer adjusted the hazard ratio for patient characteristics, by adjusting for the duration of multiple myeloma observed in the Taverna et al. study. This reduced the hazard ratio from 1.15, which suggested that patients receiving bortezomib progress more rapidly than those receiving lenalidomide, to 0.9, which suggested that patients receiving lenalidomide progress more rapidly than those receiving bortezomib. However, these base-case A analyses showed that lenalidomide dominated bortezomib, that is lenalidomide was more effective and less costly than – bortezomib. The manufacturer explained this counterintuitive result by suggesting that people who receive bortezomib experience a longer duration of progression-free survival having previously responded to bortezomib, but die sooner than those who receive lenalidomide plus dexamethasone second line. The manufacturer stated that a person whose multiple myeloma has previously responded to bortezomib is likely to respond again (delaying progression) but that, because the patient has been exposed to treatment with bortezomib before, the benefit in progression-free survival may not translate to a benefit in overall survival.

3.30 In a scenario analysis of base-case A, the manufacturer combined the results of MM-009 and MM-010 and adjusted the progression-free survival hazard ratio for the characteristics of patients in Taverna et al. (2012). This generated an overall survival hazard ratio of 1.7 and a progression-free survival hazard ratio of 1.35, and

ICERs of £12,567 per QALY gained (MM-009 only) and £3122 per QALY gained (trials combined) for lenalidomide compared with bortezomib.

3.31 In other scenario analyses of base-case A, the manufacturer included the scenarios it presented in its original analyses, as well as fitting different curves, using trial data from patients who had 1 prior therapy only, and using MM-009 rather than MM-010 (or the trials combined) to estimate lenalidomide's clinical effectiveness. The cost-effectiveness estimates for lenalidomide plus dexamethasone compared with bortezomib re-treatment ranged from lenalidomide dominating bortezomib re-treatment to an ICER of £43,331 per QALY gained. The probabilistic results showed that lenalidomide plus dexamethasone dominated re-treating with bortezomib with a probability of cost effectiveness at £20,000 and £30,000 per QALY gained of 74.5% for lenalidomide plus dexamethasone and 85% for bortezomib. However, the manufacturer had excluded the cost of dexamethasone that is given with bortezomib.

Manufacturers revised analyses after clarification round 2 (base-case B)

3.32 During the second round of clarification between the ERG and the manufacturer, the ERG highlighted that the progression-free survival curve crossed the overall survival curve in the manufacturer's model. This resulted in the model predicting that there were more people who remained progression free than there were people alive. The manufacturer corrected this using the minimum value between overall and progression-free survival to reflect progression-free survival. So, when the model predicted that overall survival was shorter than progression-free survival, the overall survival value was used (that is, progression-free survival was the same as overall survival; all the people alive would be progression free). In addition, to prevent the curves crossing, the

manufacturer changed the function it had used to fit curves, choosing a log-logistic function for overall survival instead of a piecewise exponential curve, as in the base-case model. The manufacturer did not change the log-logistic curves it had previously fitted to progression-free survival and time to treatment failure.

3.33 In response to the second clarification in base-case B the manufacturer:

- used the log-logistic curve for overall survival, progression-free survival and time to treatment failure for the lenalidomide arm (and therefore also the comparator arm), to prevent the curves from crossing
- continued as in base-case A to adjust for patient characteristics (duration of multiple myeloma) when deriving the progression-free survival hazard ratio for bortezomib compared with lenalidomide (hazard ratio 0.9 in favour of bortezomib; see sections 3.22 and 3.29).

3.34 The revised probabilistic and deterministic base-case results from base-case B showed that lenalidomide dominated bortezomib. The probability of cost effectiveness was 100% at a threshold of £20,000 per QALY gained. The manufacturer did one-way sensitivity analyses using net monetary benefit (assigning a monetary value for costs and benefits, assuming a maximum acceptable ICER of £30,000 per QALY gained). The results were most sensitive to the hazard ratios reflecting the relative effectiveness of bortezomib and lenalidomide for progression-free survival.

3.35 For base-case B, the manufacturer conducted several scenario analyses. Lenalidomide plus dexamethasone dominated bortezomib re-treatment in the following scenarios when:

- reducing the time horizon
- including concomitant dexamethasone with bortezomib in the base case and in third- and fourth-line treatments
- using different studies to estimate the efficacy of bortezomib
- approximating the complex bortezomib patient access scheme by assuming different levels of discount.

When comparing lenalidomide plus dexamethasone with the other comparators, the manufacturer estimated ICERs that were considerably higher as follows:

- bendamustine plus dexamethasone: £23,435 per QALY gained
- melphalan plus prednisone: £28,516 per QALY gained
- high-dose cyclophosphamide plus low-dose dexamethasone: £36,718 per QALY gained.

Evidence Review Group critique of cost effectiveness

3.36 The ERG stated that the manufacturer had used appropriate search terms and databases in its systematic literature review.

3.37 The ERG reviewed the manufacturer's approach to the Markov model and highlighted several important errors. Several of these remained unresolved after the 2 rounds of clarification, and are described below.

Data extrapolation

3.38 The manufacturer adjusted progression-free survival, time to progression and overall survival curves from MM-010, using the mean of covariates method, to the means published in Taverna et al. (2012), so that the MM-010 data reflected the patient characteristics from Taverna et al. The ERG noted that the mean of covariates method may skew results, and stated that the manufacturer could have used alternative approaches. Further, the ERG stated that it was not clear why the manufacturer chose the

covariates it had chosen because some covariates, such as the ECOG score, were not statistically significant predictors of progression-free survival, time to progression or overall survival. The ERG stated that the manufacturer did not include all predictors in the model noting, for example, that the manufacturer did not include the number of prior therapies.

3.39 The ERG commented that the duration of overall survival associated with lenalidomide was likely to have been overestimated in the manufacturer's model. The ERG noted that, when the manufacturer chose a log-logistic extrapolation, 11% of people in the model were still alive after 25 years. The ERG commented that because people entered the model aged 63 years, the model predicted that 11% of people with multiple myeloma at first relapse live beyond the age of 88 years, which it considered to be an unrealistic assumption.

Comparative effectiveness

3.40 The ERG commented on the manufacturer's estimate for the hazard ratio that it used in base-cases A and B to compare progression-free survival between bortezomib and lenalidomide. The ERG noted that it had changed from 1.15 (favouring lenalidomide, in the original analyses) to 0.9 (favouring bortezomib in base-case A; see section 3.29). The ERG noted that, despite this favouring bortezomib, after this change, lenalidomide dominated bortezomib (from the original base-case ICER of about £14,500 per QALY gained). The ERG's clinical specialists advised that this hazard ratio for progression-free survival in favour of bortezomib was not plausible because, in clinical practice, multiple myeloma re-treated with bortezomib would be expected to progress more quickly than when treated with lenalidomide, despite previous response to bortezomib.

Curves crossing

3.41 The ERG considered the manufacturer's approach to resolving the problem of the progression-free survival and overall survival curves crossing in the model. It noted that the manufacturer had used the minimum value for progression-free and overall survival in its initial base-case analyses to ensure the curves did not cross. However the ERG commented that this does not address why the underpinning survival curves crossed. The ERG noted that, despite the manufacturer changing the curve fitting for overall survival from piecewise exponential to log-logistic in base-case B, the curves continued to cross in the comparator arm of the model. The ERG commented that obtaining the best fit for a curve and the natural history of disease, rather than preventing curves from crossing, should form the basis for selecting a curve.

Model structure

3.42 The ERG noted that there were differences between how the manufacturer had modelled third- and fourth-line treatments in the lenalidomide and the comparator arms of the model, as follows:

- In the lenalidomide arm, when modelling third- and fourth-line treatments, the manufacturer modelled only the costs of the treatments. It did not model disutility or costs of adverse events. However, in the comparator arm, when including lenalidomide as a third-line treatment, the manufacturer included the cost and disutility of adverse events as well as the treatment costs. In addition, overall survival, progression-free survival and time to treatment failure in the comparator arm were then determined from the lenalidomide arm of the MM-010 trial.
- The duration of third-line treatment differed between the intervention and the comparator arms. In the intervention arm the third-line treatment was a mix of chemotherapies that did not include lenalidomide, and the duration of treatment for these was

fixed. In the comparator arm, however, the model included treatment with lenalidomide, the duration of which was until treatment failure. This was derived from the progression-free survival data in the lenalidomide plus dexamethasone arm in MM-010.

3.43 The ERG noted that, for the health state defined by 'progressive disease', the manufacturer's model held patients' utility values constant. However, in this health state, patients could receive third- and fourth-line treatments that, in clinical practice, could result in a remission and an increase in utility. The ERG noted that this benefit was not captured in the manufacturer's model.

3.44 The ERG commented that the manufacturer's defined clinical treatment pathway differed from the third- and fourth-line treatments included in the model, and that several modelled treatments were no longer routinely used in clinical practice in the UK. The ERG questioned whether clinicians in the UK would offer third- or fourth-line treatments in clinical practice. Taking this and the other issues into account (see sections 3.42 and 3.43), the ERG therefore questioned whether these treatment lines should be included in the model.

Drug costs

3.45 The ERG noted several problems related to the costs used in the model:

- The manufacturer did not include treatment with dexamethasone with the comparators in the initial base case, base-case A or base-case B.
- The ERG could not determine how the manufacturer estimated the costs of bortezomib per cycle in the model.
- The manufacturer's model assumed that patients remain on bortezomib second line until disease progression or they

develop adverse effects of treatment. However, the ERG's clinical specialist suggested that bortezomib in the UK is given for a fixed number of cycles (usually a maximum of 8).

- The manufacturer assumed transportation costs for half of patients to attend clinic to receive intravenous bortezomib; however, the ERG's clinical specialists suggested that the true proportion of patients would be considerably lower.
- The ERG noted that the manufacturer's calculations for the costs of disease monitoring were not clear, particularly which costs the manufacturer attributed to each state (that is, progression or progression-free).

Health-related quality of life

3.46 The ERG noted that, in the model, the progression-free survival state had a utility value of 0.81. This value is higher than would be expected for an average member for the UK population at the same age (expected to be 0.80).

Sensitivity analysis

3.47 The ERG presented exploratory analyses using the manufacturer's model for lenalidomide plus dexamethasone compared with bortezomib excluding costs and effectiveness of third- and fourth-line treatments. The ERG assumed that 64% of people who receive bortezomib also receive dexamethasone. Given the ERG's concerns about the structure and methodology of the manufacturer's model, the ERG stated that the Committee should interpret the ERG's results with caution. In the ERG's analysis excluding third- and fourth-line treatments, lenalidomide plus dexamethasone dominated bortezomib re-treatment. The ERG presented further scenarios (based on the manufacturer's base-case B model) for lenalidomide plus dexamethasone compared with bortezomib re-treatment:

- correcting an error in the manufacturer's model related to how patients are allocated to the different health states (lenalidomide dominated)
- using a maximum duration of bortezomib re-treatment in the comparator arm of 8 cycles and correcting a mistake in the values for disutility for adverse events
- generating an ICER that incorporated all these changes of £54,369 per QALY gained.

Manufacturer response to the appraisal consultation document

3.48 The manufacturer submitted additional analyses to address concerns noted by the Committee in the appraisal consultation document. The additional analyses included:

- more evidence relating to the efficacy of lenalidomide compared with bortezomib
- more evidence relating to the efficacy of lenalidomide compared with chemotherapy
- scenarios varying health-related quality of life
- scenarios varying costs (including limiting the number of cycles of bortezomib re-treatment)
- updated cost-effectiveness analyses.

Efficacy of lenalidomide compared with bortezomib

3.49 The manufacturer presented a mixed treatment comparison comparing lenalidomide with first-time (but not first-line) bortezomib treatment for people who had at least 1 prior therapy. The manufacturer stated that the common comparator dexamethasone was used in a connected network of pairwise treatment comparisons. The mixed treatment comparison included the trials MM-009 and MM-010 (see section 3.2), as well as:

- APEX, a phase III multicentre randomised controlled trial of 669 people with multiple myeloma who had received between 1 and 3 prior treatments, which compared bortezomib plus dexamethasone with dexamethasone monotherapy. No one in the trial had received prior bortezomib treatment.
- DOXIL, a phase III multicentre randomised controlled trial of 646 people with multiple myeloma who had received at least 1 prior therapy, which compared bortezomib plus pegylated liposomal doxorubicin with bortezomib monotherapy.

The manufacturer used the mixed treatment comparison to derive hazard ratios for time to progression and overall survival (table 3) using Bucher and Bayesian methods. However, the manufacturer did not use these hazard ratios in its base-case analyses; instead, it provided them as scenario analyses (see section 3.28).

Table 3 Mixed treatment comparison results

	Bucher	**Bayesian**
Time to progression HR [95% CI]	0.63 [0.42 to 0.92]	0.64 [0.41 to 0.95]
Overall survival HR [95% CI]	0.68 [0.41 to 1.10]	0.72 [0.49 to 1.01]
Abbreviations: CI, confidence interval; HR: hazard ratio.		

3.50 The manufacturer presented further evidence that compared lenalidomide with bortezomib after 1 prior treatment with bortezomib using an analysis of subsequent treatments given after the VISTA trial (see section 3.23). The manufacturer did not present a cost-effectiveness analysis that used these data. The proportions responding to lenalidomide, bortezomib and thalidomide estimated by the manufacturer were:

- response to lenalidomide: 73% (16 out of 22 people)
- response to bortezomib re-treatment: 41% (9 out of 22 people)
- response to thalidomide: 37% (23 out of 63 people).

3.51 The manufacturer provided updated survival analyses using pooled data from MM-009 and MM-010 to produce survival models for overall survival, time to progression, progression-free survival and time to treatment failure. The manufacturer identified the Gompertz (overall survival) and gamma (progression-free survival and time to treatment failure) curves as having the best visual fit and clinical plausibility. The manufacturer chose these curves for its updated base-case analysis. The manufacturer noted that the survival curves still crossed (see section 3.32). The models were adjusted using the mean of covariates method in the base-case analysis, and the corrected group prognosis method in a scenario analysis.

3.52 The manufacturer clarified how it used the lenalidomide curves to extrapolate the survival estimates. The manufacturer adjusted the lenalidomide curves to match the study-level patient characteristics of the comparators (for bortezomib, the manufacturer used Taverna et al. (2012) in the base case, in which only the duration of myeloma was available to adjust the pooled data). The adjusted data generated predicted median survival outcomes for lenalidomide, which the manufacturer then compared with the observed median survival outcomes for bortezomib from Taverna et al. to calculate 'crude' hazard ratios. The manufacturer also derived what the manufacturer referred to as an 'unadjusted' hazard ratio (but was, as the manufacturer stated, adjusted for prognostic factors, but not patient characteristics) by comparing pooled MM-009 and MM-010 data with the observed survival outcomes for bortezomib. The manufacturer used a hazard ratio for overall survival for the base case of its updated model of 1.89 (range 1.57–2.14); and 1.11 (range 0.90–2.06) for progression-free survival (where a value greater than 1 favours lenalidomide, that is the figure shows the comparison of bortezomib compared with lenalidomide).

Efficacy of lenalidomide compared with chemotherapy

3.53 The manufacturer widened its original systematic review and found 2 further studies that estimated the effectiveness of chemotherapy. In its original submission, the manufacturer assumed that the efficacy of chemotherapy was the same as bendamustine (see section 3.20):

- Petrucci et al. (1989), a single-arm trial of melphalan plus prednisolone that included 34 patients with multiple myeloma whose disease had relapsed or was refractory to standard chemotherapy.
- Celesti et al. (1997), a non-randomised trial including 28 people with advanced multiple myeloma that compared high-dose cyclophosphamide plus low-dose dexamethasone with low-dose cyclophosphamide plus dexamethasone.

The manufacturer used these studies to derive hazard ratios for overall survival in the same manner as for bortezomib (see section 3.51). The manufacturer used the same hazard ratio for both overall survival and progression-free survival because the studies did not report on progression-free survival. The overall survival hazard ratio for lenalidomide compared with melphalan plus prednisolone was 4.66 (range 3.85–5.25) and this was used for the base-case analysis. The manufacturer used high-dose cyclophosphamide plus low-dose dexamethasone and low-dose cyclophosphamide plus dexamethasone as comparators in scenario analyses. The manufacturer did not include any other chemotherapeutic agents, such as vincristine and doxorubicin, because of a lack of evidence.

3.54 The manufacturer stated that chemotherapy was not a relevant comparator because:

- People do not receive standard chemotherapy in clinical practice because of advances in treatment, including proteasome inhibitors, immunomodulatory agents and broader use of existing treatments.
- In clinical practice, people for whom bortezomib re-treatment is not appropriate would receive lenalidomide rather than chemotherapy through the Cancer Drugs Fund.
- A survey of 7 haematology consultants performed by the manufacturer estimated that 5–10% of patients receive chemotherapy for second-line treatment.

Health-related quality of life

3.55 The manufacturer submitted 2 additional scenarios for health-related quality of life. The first scenario assumed that health-related quality of life depends on the time since the start of treatment, rather than progression status. The scenario included a range of values from 0.52, for up to 1 month since treatment, to 0.71 for 36 months or longer since treatment. The utility values were sourced from Bortezomib and thalidomide for the first-line treatment of multiple myeloma (NICE technology appraisal 228), which used mapped values (from the EORTC-C30 to the EQ-5D). The second scenario used the utility values in the manufacturer's original model (that is, were based on progression status rather than time since treatment) but incorporated a lower utility value (0.59), sourced from MM-003 (a randomised open-label trial of pomalidomide plus dexamethasone compared with dexamethasone alone for the treatment of multiple myeloma), for fourth-line treatment onwards.

Updated costs

3.56 The manufacturer submitted updated cost information after consultation. The manufacturer:

- Added the cost of dexamethasone given with bortezomib to the base case for 64.3% of patients (based on estimates from Taverna et al. 2012).
- Updated the method it used to calculate the number of people who survive long enough to become eligible for the patient access scheme for lenalidomide when given third line in the comparator arm (now 15.52% of patients; original estimate not stated).
- Included additional scenario analyses to explore the impact of fewer bortezomib treatment cycles.

3.57 The manufacturer stated that bortezomib is given until treatment failure, noting that up to 19 cycles of bortezomib were given in both Taverna et al. (2012) and Sood et al. (2009) (a study of 32 people with multiple myeloma who had more than 1 treatment bortezomib). It also noted that people whose multiple myeloma responded to treatment received a median of 10 cycles bortezomib in the APEX trial.

Updated cost-effectiveness results – base-case C

3.58 After consultation, the manufacturer submitted 2 new base-case analyses (hereafter both referred to as base-case C), depending on whether re-treating with bortezomib was appropriate or not; both contained additional scenario analyses. The manufacturer made the following changes to the updated base case; it:

- pooled MM-009 and MM-010 data
- corrected the patient flow sheets
- added the cost of concomitant dexamethasone for 63% of people in the bortezomib arm
- changed the method used to estimate the proportion of people surviving long enough to receive the third-line lenalidomide patient access scheme

- added the complex patient access scheme (see section 3.25) available for bortezomib to the comparator arm in the base case.

The manufacturer assumed that not all clinicians prescribing bortezomib for patients eligible for the patient access scheme would successfully claim it. Therefore the manufacturer assumed in the base case a discount of 15% on the cost of bortezomib for 55% of modelled patients.

Manufacturer's base-case C results after consultation

3.59 Lenalidomide dominated bortezomib (deterministic analysis: incremental costs −£13,634, incremental QALYs 0.81; probabilistic analysis: incremental costs −£14,067, incremental QALYs 0.82). The manufacturer's one-way sensitivity analysis using net monetary benefit showed a 100% chance of cost effectiveness at £20,000 per QALY gained, with the ICER most sensitive to the hazard ratio for progression-free survival for lenalidomide compared with bortezomib. In a scenario analysis in which bortezomib re-treatment was given for a specific number of cycles, the ICERs were £15,409 per QALY gained (19 cycles) and £31,999 per QALY gained (8 cycles). In another scenario analysis in which the hazard ratio from the manufacturer's mixed treatment comparison was used for overall survival, lenalidomide dominated bortezomib. When the hazard ratio from the manufacturer's mixed treatment comparison was used for progression-free survival, the ICER for lenalidomide was £8936 per QALY gained compared with bortezomib.

3.60 When re-treating with bortezomib was not appropriate, and therefore chemotherapy was the comparator, the base-case ICER was £54,898 per QALY gained (incremental costs £54,414, incremental QALYs 0.99). The probabilistic ICER was £53,686 per QALY gained (incremental costs £53,629, incremental QALYs 1.00), with a 0.1% chance of cost effectiveness at £30,000 per

QALY gained. The sensitivity analysis showed that the ICER was most sensitive to time to treatment failure.

Evidence Review Group critique of additional analyses presented by manufacturer after consultation

3.61 The ERG critiqued the additional base-case C analyses submitted by the manufacturer after consultation. The critique focused on:

- clinical effectiveness of bortezomib
- clinical effectiveness of chemotherapy
- model structure
- costs
- utility values.

Clinical effectiveness of bortezomib

3.62 **Survival – unadjusted curves:** The ERG stated that it was not clear why the manufacturer had selected the curves it chose to extrapolate the unadjusted (see section 3.51) pooled MM-009 and MM-010 data for overall survival, progression-free survival and time to treatment failure. The ERG noted that the manufacturer used the Gompertz curve in its base case to extrapolate overall survival, but commented that the gamma curve appeared to fit the trial data equally well based on visual inspection. For progression-free survival and time to treatment failure, the ERG noted that the manufacturer had changed its choice of curve from log-logistic to gamma, but that it was not clear why the manufacturer chose this particular curve because other curves may have been appropriate based on visual inspection and AIC and BIC statistics.

3.63 The ERG noted that the manufacturer had provided a piecewise exponential curve for overall survival in a scenario analysis. However, the ERG noted that the methodology used by the manufacturer to extrapolate the piecewise exponential curve

resembled a 'broken curve' approach, wherein the extrapolated curve is fitted to a small amount of data at the right tail of the dataset. Consequently, the extrapolated portion is sensitive to the point in the curve where extrapolation begins. The ERG noted that the manufacturer had chosen to extrapolate from 100 weeks, without clearly justifying this time point. The ERG also noted that the manufacturer's approach to modelling the piecewise exponential curve differed from that used in the manufacturer's original base case, when the manufacturer used a traditional piecewise approach (that is, fitted separate parametric models to different time periods, allowing the hazard to change over time).

3.64 **Prognostic factors chosen by the manufacturer to adjust curve:** The ERG raised several concerns with the manufacturer's process of selecting which variables it would include to adjust curves for the treatment arm of the combined MM-009 and MM-010 trial to reflect the characteristics of the patients in the Taverna et al. (2012) study. These included:

- Concerns with the transparency of the process related to which variables to select:
  - The manufacturer had noted that prognostic factors and outcomes were taken from observational analyses of MM-009 and MM-010, but the ERG did not find any related analyses in the clinical study reports.
  - Some of the variables (for example, the presence or absence of lytic bone lesions) included by the manufacturer from the literature (for example, Dimopoulos [2009]) were not statistically significant predictors of clinical outcomes when analysed in MM-009 and MM-010.
  - The ERG noted discrepancies between p values presented by the manufacturer in its submission and in its model.

- The ERG noted that the manufacturer did not present some measures of statistical significance (for example, p values) either in its submission or in its model.
- The ERG had concerns with the manufacturer's rule used to select prognostic factors, that is, only those variables statistically significant ($p<0.05$) associated with both overall survival and either progression-free survival or time to treatment failure. The ERG noted that:
  - this introduced bias by eliminating statistically significant predictors of progression-free survival and time to treatment failure
  - different prognostic factors should be used to adjust the different overall survival, progression-free survival and time to treatment failure curves.

3.65 **Survival models – adjusted intervention curves:** the ERG commented that the overall survival estimated by the Gompertz curve seemed reasonable because the model estimated that fewer than 1% of patients who take lenalidomide second line would be alive after 12 years. However, the ERG noted that the manufacturer's model remained flawed because the curve for overall survival still crossed the curves for progression-free survival and time to treatment failure. The ERG noted that survival curves using the Gompertz rather than the piecewise exponential curve now crossed even earlier in the model. For example, in the original base case, in the intervention arm the curves for overall survival and progression-free survival crossed at about 20 years. However, in the manufacturer's base-case C, the curves crossed at after about 10 years

3.66 **Mean of covariates method:** the ERG noted that not all the relevant prognostic factors had median and mean values reported in the comparator studies and that the method chosen by the

manufacturer needed the mean. For example, Taverna et al. (2012) reported only the median duration of myeloma. The ERG commented that this implied that the manufacturer therefore had to make assumptions, but that the manufacturer did not explicitly define these assumptions.

3.67 **Corrected group prognosis method:** the ERG noted that the manufacturer had used the corrected group prognosis method as an alternative to the mean of covariate methods to adjust the lenalidomide trial population in scenario analyses. The ERG commented that the manufacturer's calculations were unclear, particularly about the way it derived weights for each patient.

3.68 **Hazard ratio calculation:** to estimate the effectiveness of lenalidomide and its comparators indirectly, the manufacturer calculated 'crude' hazard ratios. The ERG noted that, to derive the hazard ratio, the manufacturer had assumed that progression and mortality occurred at a constant rate. However, the ERG considered the manufacturer's approach flawed because the parametric models (Gompertz, gamma) fitted to the MM-010 and MM-009 pooled data do not assume constant hazards. The ERG noted that an appropriate approach is to assume constant hazards only in an exponential model. The ERG also noted a mistake in the manufacturer's calculation of the crude hazard ratio which, once corrected, changed the progression-free survival hazard ratio from 1.11 to 1.19 (using the mean of covariate method). The ERG corrected the error and changed the reference such that a hazard ratio of less than 0 favoured lenalidomide. This resulted in the manufacturer's hazard ratios for lenalidomide compared with bortezomib as follows:

- overall survival: 0.53 (mean of covariates method); 0.61 (corrected group prognosis method) (range 0.47–0.64)

- progression-free survival: 0.84 (mean of covariates); 0.95 (corrected group prognosis method) (range 0.42–1.08).

Clinical effectiveness of chemotherapy

3.69 The ERG sought the opinion of a clinical specialist who stated that chemotherapy was likely to be less effective than lenalidomide, and that only about 5% of people with multiple myeloma receive chemotherapy as a second-line treatment.

3.70 **Hazard ratios:** the ERG noted that the manufacturer used the same methodology to derive the efficacy of lenalidomide compared with chemotherapy as it used to compare lenalidomide with bortezomib re-treatment. Therefore, the ERG's same concerns apply (see section 3.67). In addition, the ERG noted it was likely that the manufacturer's estimates of efficacy used for chemotherapy underestimated the true effect. In the papers used to derive the efficacy of chemotherapy (Celesti et al. 1997 and Petrucci et al. 1989), all patients had received prior chemotherapy, and the manufacturer had assumed that prior treatment did not affect the efficacy of chemotherapy. However, a clinical specialist advising the ERG questioned this assumption. The ERG also questioned the validity of assuming that lenalidomide was equally effective for delaying progression, time to treatment failure and death (applying the same hazard ratio to all), given that MM-009 and MM-010 did not support this.

Model structure

3.71 **Third- and fourth-line treatments:** the ERG understood that the manufacturer had included in its model the possibility that patients receive therapy third and fourth line, reflecting the complex multiple myeloma care pathway. The ERG questioned the value of including these treatments in this model because:

- The data available did not allow evaluation of the effectiveness and quality of life resulting from further treatment options in the lenalidomide arm of the model.
- The effectiveness of lenalidomide third line was assumed to be the same as second line.
- The manufacturer had not dealt with several issues in its model (see sections 3.38–3.46), including:
  - utility of progressive disease
  - inconsistency in treatment arms
  - outdated drugs in the 'treatment basket'.

Costs

3.72 **Bortezomib patient access scheme:** a clinical specialist advising the ERG believed that the bortezomib patient access scheme was rarely implemented in clinical practice in England.

3.73 **Duration of bortezomib re-treatment:** the ERG noted that the manufacturer had cited both the Taverna et al. (2012) and Sood et al. (2009) studies to support that people with multiple myeloma may receive up to 19 cycles of treatment with bortezomib. However, the ERG noted that most studies and clinical experience suggest treatment lasts between 1 and 6 cycles:

- In the Taverna et al. study, although a maximum number of 19 cycles was given, 90.4% of people in the study received fewer than 6 cycles, and only 2% of patients received more than 10 cycles.
- In Sood et al., 50% of patients received 5 cycles, with a range of 1–12 cycles.
- In APEX, 85% of people in the trial received 8 cycles or fewer.
- In Hrusovsky et al. (2010), 43% of patients received 1–3 cycles, 41% of patients received 4–6 cycles and only 6% received more than 10 cycles (maximum 14 cycles).

- The ERG noted that the bortezomib summary of product characteristics states that a maximum number of 8 cycles should be given.

3.74 **Transportation costs:** the ERG noted that, in the model, half of people receiving intravenous bortezomib treatment needed transport to a hospital or clinic. For those who needed more than 1 treatment per week, the manufacturer assumed that patients were hospitalised for up to 1 week. A clinical specialist advising the ERG stated that this was not realistic, and the ERG determined that the manufacturer was likely to have overestimated the costs of bortezomib.

Utility values

3.75 **Utility values dependent on time:** the ERG reviewed the utility scenario in which utility values depend on time since treatment. The ERG commented that the manufacturer's approach changed from Bortezomib and thalidomide for the first-line treatment of multiple myeloma (NICE technology appraisal 228), in which utility was a function of disease state, to this appraisal, in which utility values are a function of time. The ERG also noted that the alternative utility values were from a younger population than that being considered in this appraisal.

3.76 **Utility decrement for fourth-line treatment:** the ERG noted that the manufacturer had applied a utility decrement to fourth-line treatments, but stated that this did not solve the problem that the model did not include a utility value for 'disease not yet progressed' before a utility for 'progressed disease' (see section 3.43). The ERG also noted that the manufacturer did not derive its utility estimates for fourth-line treatment from the clinical trial population and there was no information about the age of patients.

Results of Evidence Review Group exploratory analyses

3.77 The ERG presented several results and scenarios for the 2 different subpopulations, that is, people who can receive bortezomib again and people who cannot. The scenarios presented included:

- correcting for errors in hazard ratio calculations
- different duration of bortezomib re-treatment
- whether third- and fourth-line treatment were included
- whether the mean of covariates or corrected group prognosis method of adjustment was used.

The ERG stated that the Committee should interpret all results with caution because the individual and cumulative impacts of the problems noted with the model were unclear.

3.78 The ERG presented updated results for lenalidomide compared with bortezomib in people who could receive bortezomib re-treatment. The ERG corrected for an error in the manufacturer's base case relating to the hazard ratio calculation. The updated analyses showed that lenalidomide dominated bortezomib. Further scenarios for lenalidomide compared with bortezomib re-treatment, which also accounted for this error, included:

- When bortezomib re-treatment was a maximum of 19 cycles, the ICER was £38,871 per QALY gained (incremental costs £36,673, incremental QALYs 0.94).
- When bortezomib re-treatment was a maximum of 8 cycles, the ICER was £59,856 per QALY gained (incremental costs £56,472, incremental QALYs 0.94).
- When using alternative utility values, lenalidomide dominated bortezomib re-treatment.

- When third- and fourth-line treatments were excluded, the ICER was £11,325 per QALY gained (incremental costs £10,684, incremental QALYs 0.94).
  - When the bortezomib patient access scheme and third- and fourth-line treatments were excluded, the ICER was £5428 per QALY gained (incremental costs £5121, incremental QALYs 0.94).
- When assuming a hazard ratio of 1 and a maximum of 19 cycles of bortezomib re-treatment, the ICER was greater than £1,000,000 per QALY gained.

3.79 The ERG presented the following results for lenalidomide compared with chemotherapy in people who could not receive bortezomib:

- When correcting for an error in the hazard ratio calculation, the ICER was £54,898 per QALY gained (incremental costs £54,414, incremental QALYs 0.99).
- When third- and fourth-line treatments were excluded, and the error was accounted for, the ICER was £43,708 per QALY gained.

3.80 Full details of all the evidence are in the manufacturer's submission and the ERG report.

4 Consideration of the evidence

4.1 The Appraisal Committee reviewed the data available on the clinical and cost effectiveness of lenalidomide, having considered evidence on the nature of multiple myeloma and the value placed on the benefits of lenalidomide by people with the condition, those who represent them, and clinical specialists. It also took into account the effective use of NHS resources.

4.2 The Committee heard about the experience of patients with multiple myeloma. It heard from a clinical specialist that survival rates improved after the introduction of thalidomide, bortezomib and lenalidomide, but multiple myeloma remains an incurable disease and can be associated with renal failure and anaemia. The Committee heard how important it was for people who are unable to take thalidomide first line to have options for treatment after first relapse with bortezomib; this is true especially for people whose disease has not responded well to bortezomib, or for whom bortezomib re-treatment is not appropriate, such as people who experience adverse reactions with bortezomib. The Committee understood that, for this patient group, in the absence of lenalidomide, the treatment options are limited to standard chemotherapy and bendamustine. The Committee also heard that, in the opinion of the patient expert, using a novel agent earlier in the pathway may provide more benefit than using it later in the pathway. Finally, the Committee recognised that patients value oral treatments, such as lenalidomide.

4.3 The Committee considered the treatment pathway for the population in this appraisal, that is, people with multiple myeloma who are not eligible for stem cell transplantation and cannot receive thalidomide first line because of contraindications, and who therefore have had 1 prior treatment with bortezomib. The Committee heard from a clinical specialist that although lenalidomide and thalidomide are structurally similar, there are people for whom thalidomide treatment is not appropriate but who could take lenalidomide. The Committee heard from the clinical specialist that this population may receive bortezomib re-treatment at first relapse after initial treatment with bortezomib. However, the clinical specialist stated that this may not be appropriate for more than half of this population either because their condition does not respond to bortezomib or because of adverse reactions. The

clinical specialist explained that for these people, in the absence of lenalidomide, current treatment options would be limited to standard chemotherapy and bendamustine, so clinicians would value having alternative treatments to offer. The clinical specialist and patient experts stated that bendamustine was rarely offered to patients because it is not licensed for second-line treatment. They noted that it is licensed for first-line treatment but is not used in this way, that it is available through the Cancer Drugs Fund for relapsed multiple myeloma where other treatments are not appropriate, and that it is used in clinical practice as fourth- and fifth-line treatment. The clinical specialist added that, even when bortezomib re-treatment is appropriate, clinicians may prefer to use a drug with a different mode of action after relapse, such as lenalidomide. The Committee understood that lenalidomide as second-line treatment is available through the Cancer Drugs Fund. The Committee concluded that there are 2 subpopulations depending on whether bortezomib re-treatment can or cannot be used, and that these 2 populations have different treatment options. For people who could be treated again with bortezomib, the treatment options include bortezomib re-treatment, lenalidomide and standard chemotherapy. For people for whom bortezomib re-treatment is not an option, treatment options include lenalidomide and standard chemotherapy. Based on the opinion of the clinical specialist, the Committee agreed that it was appropriate not to consider bendamustine as a second-line treatment in this appraisal.

Clinical effectiveness

4.4 The Committee discussed the evidence for lenalidomide compared with placebo and considered it robust because the data were sourced from 2 randomised controlled trials. The Committee agreed that MM-009 and MM-010 had demonstrated the effectiveness of lenalidomide plus dexamethasone compared with placebo plus dexamethasone for progression-free survival, overall

survival and time to progression (see sections 3.3 and 3.4), but recognised that not offering treatment (placebo) was not a relevant comparator for the purpose of this appraisal. The Committee also recognised that the population in the trial did not match the population set out in the decision problem for this appraisal because:

- only 2 out of 353 patients had received 1 prior treatment with bortezomib
- the trials' inclusion criteria did not specify that thalidomide treatment was inappropriate, contraindicated or could not be tolerated
- the trial patients were younger than the multiple myeloma population addressed in this appraisal
- the trials included a high proportion of people who had had 2 or more prior therapies.

The Committee heard from the clinical specialist that, based on her experience, despite these differences the study was generalisable to the population of interest, and that the younger patient age was unlikely to affect the generalisability of the study results. The Committee concluded that the MM-009 and MM-010 trial data show that lenalidomide is more effective than no therapy for the treatment of multiple myeloma after 1 prior therapy.

4.5 The Committee considered the evidence for clinical effectiveness of lenalidomide in people for whom bortezomib re-treatment is appropriate. The Committee noted that the manufacturer estimated the effectiveness of bortezomib re-treatment from small observational studies that did not include control arms. The Committee noted that, after consultation, the manufacturer provided indirect evidence from a mixed treatment comparison, and direct evidence using follow-up data from the bortezomib VISTA trial (see section 3.23). The Committee acknowledged that the

mixed treatment comparison appeared to favour lenalidomide over bortezomib for progression-free survival, but not for overall survival. The Committee noted that the VISTA trial investigated bortezomib used first line, and that the follow-up data reflected the post-progression period when bortezomib re-treatment, lenalidomide or other treatments could be used second line. The Committee noted that the follow-up data from the VISTA trial showed higher response rates for lenalidomide than for bortezomib retreatment. The Committee recognised that the patient populations in the mixed treatment comparison and the VISTA trial were not directly relevant to the decision problem. The Committee agreed that the limited evidence base means there is significant uncertainty about the relative effectiveness of lenalidomide compared with bortezomib re-treatment for the treatment of multiple myeloma after 1 prior therapy. When questioned during the Committee meeting, the manufacturer could not provide any information about whether there were ongoing studies directly addressing the decision problem.

4.6 The Committee considered the clinical-effectiveness evidence for lenalidomide in people for whom bortezomib re-treatment was not appropriate (that is, compared with standard chemotherapy). The Committee recognised that there was no direct evidence (see section 3.52), and the manufacturer instead had estimated a crude hazard ratio for overall survival comparing standard chemotherapy with lenalidomide that ranged from 3.85 to 5.19 in favour of lenalidomide. It heard from the clinical specialist that, despite the lack of published efficacy data, in her experience it was more effective to use lenalidomide rather than standard chemotherapy, a view the ERG's clinical specialist agreed with. The Committee concluded that, although there was significant uncertainty in the evidence, given the significant size of effect in favour of lenalidomide indicated by the hazard ratios, and the opinion of the

2 clinical specialists, lenalidomide is likely to be more effective than standard chemotherapy for the treatment of multiple myeloma in people who have had 1 prior therapy.

Cost effectiveness

4.7 The Committee considered the manufacturer's economic models (initial base case, base-case A, base-case B and base-case C [the latter provided after the consultation]), and the ERG's critique.

4.8 The Committee considered the manufacturer's choice of parametric curves for extrapolating overall survival, noting that it chose a piecewise exponential curve for its initial base case and base-case A, a log-logistic curve for base-case B and a Gompertz curve for base-case C. The Committee recognised that all the curves provided before and after consultation presented fundamental problems because the overall survival and progression-free survival curves crossed, fitted poorly to the Kaplan–Meier data, or generated clinically implausible survival estimates. Therefore, after considering the additional evidence provided during consultation, the Committee remained concerned about the validity of the extrapolation curves used.

4.9 The Committee discussed the data used in the model to derive the hazard ratio for comparing lenalidomide with bortezomib. It understood that the evidence of effectiveness in all 4 base-case analyses for bortezomib did not come from a clinical trial with clearly defined end points, but from a review of patients' medical records. The Committee agreed that it was not clear whether the definitions of response were the same in the lenalidomide trials as the bortezomib review of medical records, and therefore whether they were comparable. The Committee also understood that the manufacturer had not included a study by Hrusovsky et al. (2010) (which, like Taverna et al. 2012, was a retrospective review of

patients' medical records) in the base case because 26% of patients in the study received concomitant treatment. However, the Committee understood that the inclusion criteria for the 2 studies were similar. The manufacturer explained that the results in the Hrusovsky et al. study were similar to those in the Taverna et al. study, so including them was unlikely to affect the results. The Committee noted that the Hrusovsky et al. study included more patients than the Taverna et al. study but was concerned that the studies may have included the same Swiss patients, so pooling them may have been inappropriate. Further, the Committee was aware of the mixed treatment comparison and data from the VISTA trial provided by the manufacturer during consultation (see section 4.4). It noted that the hazard ratios derived from the mixed treatment comparison showed that lenalidomide was less effective compared with bortezomib re-treatment than when the relative effectiveness was estimated from the Taverna et al. study, and that these hazard ratios were used in scenario analyses of base-case C by the manufacturer. The Committee noted that the mixed treatment comparison was for a different patient population. The Committee recognised that there were limited data available to the manufacturer for this population (see section 4.5). Therefore significant uncertainty remained in the relative effectiveness of lenalidomide compared with bortezomib retreatment in all 4 base cases. The Committee agreed that the additional data provided by the manufacturer after consultation, including the mixed treatment comparison, had not reduced this uncertainty. It concluded that this uncertainty needed to be taken into account in the decision making.

4.10 The Committee discussed how the manufacturer calculated the hazard ratios comparing the effectiveness of lenalidomide with bortezomib. It noted that the crude hazard ratios had been calculated by comparing the medians from each distribution in all base cases. The Committee understood that the assumptions

necessary to calculate hazard ratios in this way only hold when using exponential distributions, but that the manufacturer extrapolated progression-free survival and overall survival using a range of curves that were not exponential, including log-logistic, Gompertz and gamma. The Committee therefore concluded that the manufacturer's approach was not appropriate. The manufacturer acknowledged these limitations, explaining that it chose this approach because the exponential curve was not a good fit. The Committee concluded that this further added to its concerns about the validity of the model.

4.11 The Committee discussed how the manufacturer had adjusted the lenalidomide data to derive the crude hazard ratios. The Committee understood that, when deriving the hazard ratios that compared overall survival, progression-free survival and time to progression of lenalidomide with bortezomib, the manufacturer had adjusted the lenalidomide data to include factors that predicted outcomes, and also to allow for differences in patient characteristics between the studies. The Committee noted the large effect that adjusting for patient characteristics could have, for example, the hazard ratio in base-case A for progression-free survival changed from favouring lenalidomide (without adjustment) to favouring bortezomib (when adjusted). However, for all 4 base cases, it was unclear which variables the manufacturer had chosen, and how it had adjusted the data. The Committee heard from the manufacturer at the second Committee meeting that its original submission stated that variables associated with overall survival, and 1 or the other or both of progression-free survival and time to progression were chosen. However, this had been incorrectly described because only prognostic factors associated with overall survival were chosen. The Committee concluded that this meant that the manufacturer was likely to have adjusted progression-free survival and time to progression for factors that were not associated with these

outcomes, and omitted risk factors that were in all 4 base-case analyses. Further, in all 4 base cases, the manufacturer had identified prognostic factors significant for overall survival using data from MM-010 and MM-009, for which no information was reported in Taverna et al. (2012), and had therefore assumed the patient characteristics were the same in Taverna at el. as MM-010 and MM-009. The Committee noted that, following consultation, the manufacturer had provided hazard ratios that were described as 'unadjusted', but that although these hazard ratios were not adjusted for patient characteristics, they were adjusted for prognostic factors. The Committee agreed that it was not clear whether there was any value in adjusting the data, given the absence of necessary information, and questioned whether the manufacturer had adjusted the data appropriately. It concluded that these adjustments increase the noise in the manufacturer's analyses, deepening the uncertainty of the results.

4.12 The Committee considered the manufacturer's approach to modelling third- and fourth-line treatment. The Committee agreed it was important to consider treatments that would follow second-line treatment, and to include both their costs and effectiveness. It noted that the manufacturer had modelled the costs and effects differently in the lenalidomide and the comparator arms in all 4 base cases presented, and agreed that this was not appropriate and instead costs and effectiveness should be modelled consistently in the lenalidomide and comparator arms. The Committee agreed that it was not clinically plausible that, as someone's disease progresses, his or her utility value would remain constant despite receiving third- and fourth-line treatments which is likely to increase quality of life (see section 3.43). In addition, the Committee noted that, by including third- and fourth-line treatments, lenalidomide treatment was included in the comparator arm in all 4 base cases, meaning that the model

compared lenalidomide with bortezomib followed by lenalidomide. The Committee agreed that this contributed substantially to the costs incurred in the comparator arm, and has an impact on the incremental cost-effectiveness ratio (ICER) in favour of lenalidomide. The Committee also recognised that the included third- and fourth-line treatments were not representative of current clinical practice in the UK. The Committee noted that the manufacturer had explored the impact of third- and fourth-line treatment by providing scenarios that excluded these treatments in the model; this changed the ICER in base-case C to £3900 per quality-adjusted life year (QALY) gained (from lenalidomide dominating bortezomib re-treatment). The ERG presented the same scenario, after making corrections to the model, which resulted in an ICER of £11,300 per QALY gained for lenalidomide compared with bortezomib re-treatment. The Committee agreed that it is appropriate to include third- and fourth-line treatments in the economic analysis, but agreed that here this had introduced further noise into the model. The Committee concluded that the ICERs were sensitive to modelling of third- and fourth-line treatments, and it would need to take this into account when making its decisions.

4.13 The Committee discussed the number of cycles of bortezomib that people receive when re-treated with bortezomib. The Committee noted that, in all 4 manufacturer's base cases, modelled patients receive bortezomib until disease progression which, the Committee heard from the manufacturer, was for a median of 12 cycles. The Committee was aware that the marketing authorisation for bortezomib limits the number of cycles to 8. It also heard from the clinical specialist that in clinical practice in the UK, the number of bortezomib cycles offered to patients varies from 6 cycles to until treatment progression. However, the Committee noted that, in Taverna et al. (2012), the median number of bortezomib cycles was

3, that 90% of people received 6 or fewer cycles, and, as stated by the manufacturer in the second Committee meeting, only 1 person received 19 cycles. The Committee agreed that the manufacturer's base-case model markedly overestimated the number of cycles of bortezomib used in clinical practice in the UK, and that the number of cycles had a substantial impact on the cost effectiveness of lenalidomide. The Committee noted that, when the manufacturer limited the number of cycles of bortezomib to 19 in a scenario analysis provided after consultation (base-case C), the ICER changed from lenalidomide dominating bortezomib re-treatment to £15,400 per QALY gained. The Committee noted that reducing the maximum number of cycles of bortezomib to 8 increased the ICER to about £32,000 per QALY gained for lenalidomide compared with bortezomib in the manufacturer scenario analyses (base-case C) (see section 3.58). The Committee concluded that, although there may be people in clinical practice who receive more than 8 cycles of bortezomib, there are also people who receive fewer, and that, of all the scenarios presented, the one that had a maximum of 8 cycles of bortezomib re-treatment most plausibly reflected clinical practice.

4.14 The Committee noted 1 comment received during consultation that the patient access scheme for bortezomib is rarely implemented in clinical practice. It understood that patient access schemes are pricing agreements between the manufacturer and the Department of Health. The Committee acknowledged that there may be some uncertainty about the actual costs of bortezomib paid in the NHS, but that it had no authority to assume a different cost than that agreed between the manufacturer and the Department of Health. The Committee noted comments from consultation that bortezomib can be administered as a subcutaneous injection, but that the manufacturer's model included only intravenous administration. The Committee agreed that, if some patients received bortezomib

subcutaneously rather than intravenously, the costs of administration and transportation, and the disutility associated with intravenous therapy would fall, and the cost effectiveness of lenalidomide relative to bortezomib would likely worsen.

4.15 The Committee discussed the sources of the utility values and utility decrements used by the manufacturer in its model. The Committee noted that the source of the utility values used by the manufacturer in all its base cases was van Agthoven et al. (2004), derived from a 2002 PhD thesis which, to the Committee's knowledge, has never been published in a peer-reviewed journal. The Committee also noted that the utility values were derived from a younger population, and were higher than the average population of the same age. In addition, the manufacturer took the utility decrements for adverse events from several different sources, which used different methods, were from other countries, and included people with different types of cancers. The Committee agreed that the utility scenarios provided by the manufacturer after consultation were of limited value because of the way in which they were applied in the model. The Committee concluded that, although the utility values from all base cases had been used in previous appraisals for multiple myeloma, significant uncertainty remained in how utility and disutility values affected the model outcomes.

4.16 The Committee discussed whether it could determine a preferred model version and most plausible ICERs. It acknowledged that there were limited data available to the manufacturer, and that the manufacturer had adjusted the model in an attempt to make it more realistic in all 4 base cases. However, the Committee agreed that this did not reduce the high levels of uncertainty associated with survival extrapolation, adjusting median survival outcomes, and modelling third- and fourth-line treatments. The Committee also

agreed this uncertainty had not been resolved by the additional analyses provided by the manufacturer at consultation. In considering the cost effectiveness of:

- **Lenalidomide compared with bortezomib re-treatment:** the Committee considered that the most useful base case for decision-making was one in which the manufacturer had corrected an error in the hazard ratio calculation, and for which the ERG provided scenario analyses assessing the effect of different curves (base-case C). It noted that the ICER presented by the manufacturer (which included a maximum of 8 cycles of bortezomib) was £32,000 per QALY gained for lenalidomide compared with bortezomib re-treatment, and that when incorporating the ERG corrections this ICER increased to £60,000 per QALY gained. The Committee noted that the ICER was sensitive to applying different extrapolation models that could increase or decrease the ICER. The Committee also noted that taking into account the uncertainties relating to third- and fourth-line treatments, hazard ratios and bortezomib costs would increase the ICER, noting that the ICER was higher than £1 million per QALY gained when the effectiveness of lenalidomide and bortezomib were assumed to be the same.
- **Lenalidomide compared with standard chemotherapy:** The Committee noted that the most recently presented ICER (base-case C) was about £55,000 per QALY gained for lenalidomide compared with standard chemotherapy (see section 3.59), but agreed that this value was very uncertain.

The Committee concluded that the ICERs for lenalidomide compared with bortezomib re-treatment or with standard chemotherapy were more than £30,000 per QALY gained. It therefore concluded that it could not recommend lenalidomide as a cost-effective use of NHS resources for people with multiple

myeloma for whom thalidomide treatment is not appropriate, and who have received 1 prior treatment with bortezomib.

4.17 The Committee discussed whether lenalidomide is innovative in making a significant and substantial impact on health-related benefits. It agreed that, as an oral treatment, lenalidomide would be convenient, and could save time and resources for people with multiple myeloma who had had 1 prior treatment with bortezomib, and that this may not be included in the ICERs presented. However, the Committee also appreciated that bortezomib can be administered subcutaneously, lessening the difference in convenience between treatments. The Committee concluded that it was unclear how these aspects could be modelled, and that it was unlikely to make a substantial difference to its conclusions considering the high level of uncertainty.

4.18 The Committee considered whether lenalidomide treatment meets the end-of-life criteria for people with multiple myeloma who have had 1 prior treatment with bortezomib, and for whom thalidomide treatment and stem cell transplant are not appropriate. It noted that the manufacturer had not presented data to support lenalidomide considered as an end-of-life therapy. It noted that the number of life years estimated in the comparator arm of the model was greater than 24 months, and therefore concluded that, based on the data available, the end-of-life criteria had not been met. The Committee agreed that it did not need to discuss the remaining criteria.

Summary of Appraisal Committee's key conclusions

	Appraisal title:	Section
Key conclusion		
Lenalidomide in combination with dexamethasone is not recommended within its marketing authorisation for treating multiple myeloma in people: • whose condition has relapsed for the first time, **and** • who have had 1 prior treatment with bortezomib, **and**		1.1

• for whom thalidomide is contraindicated or cannot be tolerated **and** • for whom stem cell transplantation is not appropriate. The Committee concluded that there was uncertainty in the relative effectiveness (progression-free survival and overall survival) of lenalidomide compared with bortezomib re-treatment because direct trial data were not available, the effectiveness of bortezomib re-treatment was estimated from single-arm trials, data were conflicting in terms of whether it favoured lenalidomide or bortezomib, and indirect data were not from an appropriate population. The Committee concluded that lenalidomide was likely to be more effective than standard chemotherapy because, although the data were taken from single-arm studies, the data consistently favoured lenalidomide with a large benefit, and this was supported by clinical specialist opinion. The Committee concluded that the incremental cost-effectiveness ratios (ICERs) for lenalidomide compared with bortezomib re-treatment or with standard chemotherapy were more than £30,000 per quality-adjusted life year (QALY) gained.		4.3 4.15
Current practice		
Clinical need of patients, including the availability of alternative treatments	The Committee heard how important it was for people who are unable to take thalidomide first line to have options for treatment after first relapse with bortezomib, especially for people whose disease has not responded well to bortezomib, or for whom bortezomib re-treatment is not appropriate, such as people who experienced adverse reactions with bortezomib.	4.2
The technology		
Proposed benefits of the technology How innovative is the technology in its potential to make a significant and substantial impact on health-related benefits?	The Committee agreed that, as an oral treatment, lenalidomide would be convenient and could save time and resources. However, the Committee also noted that bortezomib can be administered subcutaneously, lessening the difference between treatments in terms of convenience.	4.16
What is the position of the treatment in the pathway of care for the condition?	Lenalidomide in combination with dexamethasone has a marketing authorisation for treating multiple myeloma in adult patients who have received at least 1 prior therapy.	2.1

Adverse reactions	The summary of product characteristics includes the following adverse effects for lenalidomide: neutropenia, anaemia and thrombocytopenia. Because lenalidomide is structurally related to thalidomide, a known human teratogen that causes severe birth defects, a risk minimisation plan has been developed and agreed with the Medicines and Healthcare products Regulatory Agency to avoid fetal exposure to lenalidomide.	2.2
Evidence for clinical effectiveness		
Availability, nature and quality of evidence	The manufacturer presented evidence from randomised clinical trials, MM-009 and MM-010, to show the effectiveness of lenalidomide. However, the trials were placebo controlled, and as no treatment was not a comparator in the decision problem, only data from the lenalidomide arm of the trials were used. The Committee noted that the manufacturer estimated the effectiveness of bortezomib re-treatment from small observational studies, and that these studies did not include a control arm, preventing any direct or indirect comparisons. After consultation, the manufacturer provided indirect evidence using a mixed treatment comparison and direct evidence using follow-up data from the bortezomib VISTA trial. However, the populations of these were not directly relevant to the decision problem. The Committee agreed that limited data exist for lenalidomide compared with bortezomib in this setting. The Committee recognised that the manufacturer had presented no direct evidence comparing lenalidomide with standard chemotherapy. The Committee concluded that there was significant uncertainty in the evidence.	3.23 3.55 4.4 4.6 4.5 3.3 3.4
Relevance to general clinical practice in the NHS	The Committee recognised that the population in the MM-009 and MM-010 trials did not match the population set out in the decision problem for this appraisal. However, the Committee heard from the clinical specialist that, based on her experience, despite these differences, the trials were generalisable to the population of interest.	4.6

Uncertainties generated by the evidence	There was uncertainty in the relative effectiveness (progression-free survival and overall survival) of lenalidomide compared with bortezomib re-treatment or standard chemotherapy. This was determined by estimating crude hazard ratios that used data from different trials for each treatment because no direct or indirect evidence was available. In addition, these hazard ratios were calculated using median values. The assumptions needed to use median values in this way only hold when using an exponential distribution, however, the exponential distribution was not used. These hazard ratios were therefore highly uncertain.	4.4 4.5 4.6 4.10
Are there any clinically relevant subgroups for which there is evidence of differential effectiveness?	The Committee concluded that there were 2 subpopulations depending on whether bortezomib re-treatment was or was not appropriate, and that these 2 populations had different treatment options.	4.3
Estimate of the size of the clinical effectiveness including strength of supporting evidence	Based on the evidence provided by the manufacturer, the Committee could not conclude that lenalidomide second line was more effective than bortezomib re-treatment for the treatment of multiple myeloma after 1 therapy. The Committee concluded that, although there was significant uncertainty in the evidence, lenalidomide was likely to be more effective than standard chemotherapy for the treatment of multiple myeloma in those who have had 1 prior therapy.	4.4 4.5 4.6
How has the new clinical evidence that has emerged since the original appraisal (TA171) influenced the current (preliminary) recommendations?	The same clinical trials as in TA171 (MM-009 and MM-010) were presented to demonstrate the effectiveness of lenalidomide plus dexamethasone and placebo plus dexamethasone. The evidence presented in this appraisal included an extended follow-up of overall survival, which was not included in NICE technology appraisal guidance 171. This added further support to the evidence that lenalidomide was clinically effective compared with placebo. Additional non-comparative observational studies have been published for the comparators.	4.4 4.5 4.6

Evidence for cost effectiveness		
Availability and nature of evidence	The manufacturer's Markov model compared lenalidomide with bortezomib re-treatment and with chemotherapy. The manufacturer presented its original base case in the manufacturer's submission. Following the clarification process with the ERG, the manufacturer submitted base-case A and then base-case B, each with different extrapolation methods. Following additional concerns noted by the Committee, the manufacturer provided further analyses, base case-C, after consultation.	3.19 3.26
Uncertainties around and plausibility of assumptions and inputs in the economic model	The Committee acknowledged that there were limited data available to the manufacturer, and that the manufacturer had adjusted the model in an attempt to make it more realistic. However, the Committee agreed that this did not reduce the high levels of uncertainty, noting concerns around: • survival extrapolation • bortezomib effectiveness estimates • adjustments • modelling of third- and fourth-line treatments • the number of bortezomib re-treatment cycles • cost associated with administration of bortezomib • utility values.	4.8 4.9 4.10 4.11 4.12 4.13 4.14 4.15

Incorporation of health-related quality-of-life benefits and utility values Have any potential significant and substantial health-related benefits been identified that were not included in the economic model, and how have they been considered?	The Committee noted that the source of the utility values used in all the base cases was van Agthoven et al. (2004), derived from a 2002 PhD thesis that, to the Committee's knowledge, had not been published in a peer-reviewed journal. The Committee also noted that the utility values were derived from a younger population andwere higher than the average population of the same age. In addition, the manufacturer took the utility decrements for adverse events from several different sources, which used different methods, were from other countries and included people with different types of cancers. The Committee concluded that, although the values had been used in previous appraisals for multiple myeloma, there was significant uncertainty in the values of utility and disutility used in the modelling. The Committee agreed that the benefit of oral treatment was not included in the analysis. However, the Committee concluded that it was unclear how this could be modelled, and that it was unlikely to make a substantive difference to its conclusions considering the high level of uncertainty.	4.11 4.14 4.16
Are there specific groups of people for whom the technology is particularly cost effective?	See above for subgroups.	4.3
What are the key drivers of cost effectiveness?	The Committee noted that the key drivers of cost effectiveness were the hazard ratio for the relative effectiveness of lenalidomide compared with bortezomib or standard chemotherapy, the extrapolation and the adjustment of the curves applied for progression-free survival and overall survival, the way in which third- and fourth-line treatments had been modelled, and the number of cycles of bortezomib re-treatment.	4.15
Most likely cost-effectiveness estimate (given as an ICER)	The Committee concluded that, for both groups, the ICERs for lenalidomide compared with bortezomib re-treatment or treating with standard chemotherapy were more than £30,000 per QALY gained. It therefore concluded that lenalidomide could not be recommended as a cost-effective use of NHS resources for people with multiple myeloma for whom thalidomide treatment and stem cell transplant were not appropriate, and who had received 1 prior treatment with bortezomib.	4.15

How has the new cost-effectiveness evidence that has emerged since the original appraisal (NICE technology appraisal guidance 171) influenced the current (preliminary) recommendations?	The manufacturer identified no new health economic studies from the literature, but presented 4 iterations of a new Markov model, and data on the effectiveness of the comparators from small observational studies. The preliminary recommendations for lenalidomide for treating multiple myeloma after its first relapse have not changed from TA171.	3.19 4.15
Additional factors taken into account		
Patient access schemes (PPRS)	Not applicable.	
End-of-life considerations	The Committee noted that the manufacturer had not presented data to support lenalidomide qualifying as an end-of-life therapy. The Committee recognised that the population size was likely to be small. The Committee noted that the number of life years estimated in the comparator arm of the model were greater than 24 months, and therefore concluded that, based on the data available, the end-of-life criteria had not been met.	4.17
Equalities considerations and social value judgements	Not applicable.	

5 Implementation

Section 7(6) of the National Institute for Health and Care Excellence (Constitution and Functions) and the Health and Social Care Information Centre (Functions) Regulations 2013 requires clinical commissioning groups, NHS England and, with respect to their public health functions, local authorities to comply with the recommendations in this appraisal within [insert number] months of its date of publication. The normal period of compliance, of 3 months, has been extended for this technology because [insert reason]. This extension is made under Section 7(5) of the Regulations.

5.1 NICE has developed tools [link to www.nice.org.uk/guidance/TAXXX] to help organisations put this guidance into practice (listed below). [NICE to amend list as needed at time of publication]

- Slides highlighting key messages for local discussion.
- Costing template and report to estimate the national and local savings and costs associated with implementation.
- Implementation advice on how to put the guidance into practice and national initiatives that support this locally.
- A costing statement explaining the resource impact of this guidance.
- Audit support for monitoring local practice.

6 Recommendations for further research

6.1 The Committee heard from the manufacturer that registry data were available for multiple myeloma in the USA, and that a study had been recently completed comparing lenalidomide plus low-dose dexamethasone with high-dose dexamethasone in people who had multiple myeloma.

6.2 The Committee noted that, to reduce uncertainty, it would be valuable to have a trial comparing lenalidomide with bortezomib and standard chemotherapy for the treatment of multiple myeloma in people whose multiple myeloma has relapsed for the first time, who have received 1 prior treatment with bortezomib, for whom thalidomide is contraindicated or cannot be tolerated and for whom bone marrow transplantation is not appropriate.

7 Related NICE guidance

Details are correct at the time of consultation and will be removed when the final guidance is published. Further information is available on the NICE website.

Published

- Bortezomib for induction therapy in multiple myeloma before high-dose chemotherapy and autologous stem cell transplantation. NICE technology appraisal guidance 311 (2014).
- Bortezomib and thalidomide for the first-line treatment of multiple myeloma. NICE technology appraisal guidance 228 (2011).
- Lenalidomide for the treatment of multiple myeloma in people who have received at least one prior therapy. NICE technology appraisal guidance 171 (2009).
- Bortezomib monotherapy for relapsed multiple myeloma. NICE technology appraisal guidance 129 (2007).
- Improving outcomes in haematological cancer. NICE cancer service guidance (2003).

Under development

- Myeloma: diagnosis and management of myeloma. NICE clinical guideline, publication expected January 2016.

8 Proposed date for review

8.1 NICE proposes that the guidance on this technology is incorporated, verbatim, into the upcoming NICE clinical guideline on myeloma, at which point an appropriate review date will also be considered. NICE welcomes views on this proposed date. The Guidance Executive will decide whether the technology should be reviewed based on information gathered by NICE, and in consultation with consultees and commentators.

Amanda Adler
Chair, Appraisal Committee
July 2014

9 Appraisal Committee members, guideline representatives and NICE project team

9.1 Appraisal Committee members

The Appraisal Committees are standing advisory committees of NICE. Members are appointed for a 3-year term. A list of the Committee members who took part in the discussions for this appraisal appears below. There are 4 Appraisal Committees, each with a chair and vice chair. Each Appraisal Committee meets once a month, except in December when there are no meetings. Each Committee considers its own list of technologies, and ongoing topics are not moved between Committees.

Committee members are asked to declare any interests in the technology to be appraised. If it is considered there is a conflict of interest, the member is excluded from participating further in that appraisal.

The minutes of each Appraisal Committee meeting, which include the names of the members who attended and their declarations of interests, are posted on the NICE website.

Dr Amanda Adler (Chair)
Consultant Physician, Addenbrooke's Hospital

Professor Ken Stein (Vice Chair)
Professor of Public Health, University of Exeter Medical School

Dr Ray Armstrong
Consultant Rheumatologist, Southampton General Hospital

Dr Jeff Aronson
Reader in Clinical Pharmacology, Nuffield University Department of Primary Health Care Sciences, University of Oxford

Professor John Cairns

Professor of Health Economics Public Health and Policy, London School of Hygiene and Tropical Medicine

Matthew Campbell-Hill

Lay member

Mark Chapman

Health Economics and Market Access Manager, Medtronic UK

Professor Imran Chaudhry

Lead Consultant Psychiatrist and Deputy Associate Medical Director, Lancashire Care NHS Foundation Trust

Dr Lisa Cooper

Echocardiographer, Stockport NHS Foundation Trust

Dr Maria Dyban

GP

Robert Hinchliffe

HEFCE Clinical Senior Lecturer in Vascular Surgery and Honorary Consultant Vascular Surgeon, St George's Vascular Institute

Dr Neil Losson

GP

Anne Joshua

Associate Director of Pharmacy, NHS Direct

Dr Miriam McCarthy

Consultant, Public Health, Public Health Agency, Northern Ireland

Professor Ruairidh Milne

Director of Strategy and Development and Director for Public Health Research at the National Institute for Health Research (NIHR) Evaluation, Trials and Studies Coordinating Centre at the University of Southampton

Dr Peter Norrie

Principal Lecturer in Nursing, DeMontfort University

Christopher O'Regan

Head of Health Technology Assessment & Outcomes Research, Merck Sharp & Dohme

Dr John Pounsford

Consultant Physician, Frenchay Hospital, Bristol

Dr Danielle Preedy

Lay member

Cliff Snelling

Lay member

Professor Andrew Stevens

Professor of Public Health, Department of Public Health and Epidemiology, University of Birmingham

David Thomson

Lay member

Dr Nicky Welton

Senior Lecturer in Biostatistics/Health Technology Assessment, University of Bristol

Dr Nerys Woolacott

Senior Research Fellow, Centre for Health Economics, University of York

9.2 NICE project team

Each technology appraisal is assigned to a team consisting of 1 or more health technology analysts (who act as technical leads for the appraisal), a technical adviser and a project manager.

Carl Prescott

Technical Lead

Melinda Goodall

Technical Adviser

Jeremy Powell

Project Manager

10 Sources of evidence considered by the Committee

A. The Evidence Review Group (ERG) report for this appraisal was prepared by Peninsula Technology Assessment Group (PenTAG):

- Bacelar M, Durand A, Cooper C et al. (2014) The clinical and cost effectiveness of lenalidomide for people who have received at least one prior therapy with bortezomib (partial review of TA171).

B. The following organisations accepted the invitation to participate in this appraisal as consultees and commentators. They were invited to comment on the draft scope, the ERG report and the appraisal consultation document (ACD). Organisations listed in I were also invited to make written submissions. Organisations listed in II and III had the opportunity to give their expert views. Organisations listed in I, II and III also have the opportunity to appeal against the final appraisal determination.

I. Manufacturer/sponsor:

- Celgene

II. Professional/specialist and patient/carer groups:

- Myeloma UK

III. Other consultees:

- Cancer Research UK
- British Society of Haematology
- Royal College of Nursing
- Royal College of Physicians
- UK Myeloma Forum
- Department of Health
- NHS England

- Welsh Government

IV. Commentator organisations (did not provide written evidence and without the right of appeal):

- Commissioning Support Appraisals Service
- Department of Health, Social Services and Public Safety for Northern Ireland
- Napp
- National Collaborating Centre for Cancer

C. The following individuals were selected from clinical specialist and patient expert nominations from the consultees and commentators. They gave their expert personal view on Celgene by attending the initial Committee discussion and providing written evidence to the Committee. They are invited to comment on the ACD.

- Dr Faith Davies, Faculty and Haematology Consultant, Institute of Cancer Research and Royal Marsden Hospital Section of Haemato-oncology, nominated by the Royal College of Physicians on behalf of NCRI/RCP/RCR/ACP/JCCO – clinical specialist
- Judy Dewinter, Chairman, Myeloma UK, nominated by Myeloma UK – patient expert
- Eric Low, Chief Executive, Myeloma UK, nominated by Myeloma UK – patient expert

D. Representatives from the following manufacturer/sponsor attended Committee meetings. They contributed only when asked by the Committee chair to clarify specific issues and comment on factual accuracy. They were also invited to comment on the ACD.

- Celgene

Meindert Boysen
Programme Director
Centre for Health Technology Evaluation
1st Floor
10 Spring Gardens
London
SW1A 2BU

21 August 2014

Dear Meindert

Myeloma UK comments on the second appraisal consultation document (ACD) on lenalidomide (post bortezomib) (partial review of TA171)

Myeloma UK welcomes the opportunity to comment for a second time on the NICE ACD on lenalidomide for treating myeloma after one prior treatment with bortezomib (part-review of TA171). We do not have any further specific comments to add on the ACD aside from those we have already provided during the earlier stages of the appraisal process and through the productive discussions we have had with Celgene and NICE as well as the Department of Health.

We fully understand the reasons why NICE have reached a second draft no in this appraisal. We also understand the difficulties associated with the appraisal given its status as a partial review of already approved guidance (TA171). However, we remain frustrated and disappointed that a solution has not yet been agreed.

In the first and second drafts of the ACD, NICE have clearly accepted the clinical benefit of lenalidomide when used in myeloma patients who have received Velcade as an initial treatment. It is therefore important that Celgene, NICE and Department of Health build on recent discussions and work together to iron out the issues and agree a workable solution.

We hope this matter is resolved as a matter of urgency. If we can provide any further assistance or information to assist with this appraisal, then please do not hesitate to contact us.

Yours sincerely

Eric Low
Chief Executive

Myeloma UK
Broughton House, 31 Dunedin Street, Edinburgh EH7 4JG Tel: 0131 557 3332 Fax: 0131 556 9785
www.myeloma.org.uk myelomauk@myeloma.org.uk
Company No.190563 Charity No. SC 026116 Myeloma UK is an Investor in People
Myeloma Infoline 0800 980 3332

Email: [redacted]

20th August 2014

To the Chair

I am writing on behalf of the UK Myeloma Forum / Royal College of Physicians / Royal College of Pathologists / NCRI

Thank you for the opportunity to comment on the recently issued second draft Appraisal Consultation Document "Lenalidomide for treating multiple myeloma after 1 prior treatment with bortezomib (part-review of TA171)" (July 2014).

We are surprised and disappointed that despite re-reviewing the evidence submitted by the manufacturer and other stakeholders that this application has not been approved and believe that the provisional recommendations are neither sound nor a suitable basis for guidance to the NHS.

This decision is a particular concern as this will have a significant impact on an increasingly large number of patients with myeloma at the time of relapse as a result of increased access to bortezomib as first line therapy for both transplant eligible and transplant ineligible patients.

We believe that the population group for whom this guidance should be applicable (i.e. Myeloma patients who require second line therapy and who received bortezomib as first line therapy) is poorly defined in the draft guidance. The current treatment paradigm for newly diagnosed patients with symptomatic myeloma is determined by their suitability to undergo autologous stem cell transplant. For transplant eligible patients NICE have recently issued guidance suggesting the use of bortezomib/dexamethasone (VD) or bortezomib/thalidomide/dexamethasone (VTD) as initial therapy prior to transplant (TA311). For patients considered transplant ineligible whereas the outdated NICE TA228 (July 2011) recommends thalidomide + alkylator + corticosteroid as the majority first line therapy, NHS England have taken more recent clinical trial data into account and supported the use of the Bortezomib/Melphalan/Prednisolone (VMP) for first line therapy. As a result, an increasingly large number of patients, both transplant eligible and transplant ineligible, treated outside of a clinical trial will receive a bortezomib based first line therapy. Thus we feel that the

population for whom this guidance is relevant has changed, and will be changing substantially from the population initially identified in the scope of the appraisal.

The wording of the appraisal fails to take into account those patients who were initially considered transplant eligible, nor any patients who received a bortezomib based treatment as first line therapy by other means e.g. on a clinical trial such as Myeloma XI or PADIMAC. The final wording of this part-review therefore appears to be applicable to only a small proportion of myeloma patients requiring second line treatment, specifically only those at first relapse who have had 1 prior treatment with bortezomib but for whom thalidomide is contraindicated or not tolerated and for whom stem cell transplant is contraindicated. Consideration for access to lenalidomide as second line therapy must be considered for all patients who have received bortezomib first line and not be restricted to the small subpopulation defined above.

As in our previous letter to appeal the decision we believe that the absence of lenalidomide as a treatment option at this disease stage reduces the potential to deliver optimal care to patients with relapsed myeloma. We believe that the available peer review published clinical evidence (grade A level 1a) supports lenalidomide as the most efficacious drug in the setting of 2nd line treatment following the use of a bortezomib based therapy at first line. It should be noted that there is no evidence to suggest cross resistance between bortezomib (a proteosome inhibitor) and lenalidomide (an immunomodulator) as they have differing mechanisms of action and we would suggest that restricting review of data only to patients who have previously received bortezomib prior to lenalidomide reduces the power of evidence presented.

The ACD appears to support the use of bortezomib retreatment, thalidomide based treatment or conventional chemotherapy including cyclophosphamide, melphalan or bendamustine for second line therapy in preference to lenalidomide. There is insufficient evidence to support these options for the clinical scenario under consideration. Bortezomib retreatment has not been examined in a phase 3 setting to confirm its superior efficacy above any other treatment approach. All published data on bortezomib retreatment including the recently published meta-analysis (Knopf et al. Clinical Lymph Myeloma Leukemia 2014) suggests that the patients who benefit from this approach are those who have previously responded well and tolerated previous bortezomib. We would therefore accept that bortezomib retreatment could be considered for those patients who have had at least a Partial Response to previous bortezomib with a response duration of at least 6 months but that this approach should not be mandated.

It is certainly inappropriate for patients who have not had a durable response to bortezomib, for whom lenalidomide is the clinically appropriate therapy. The use of either thalidomide or conventional chemotherapy treatments would not be considered clinically acceptable as there is limited positive supportive data for these approaches at 1st relapse and they represent an outdated and inferior treatment approach in comparison to lenalidomide. Similarly, as we stated in our response to the 1st draft guidance there is currently little data to support the use of bendamustine at second line as a therapeutic option and it is not currently licensed for this indication. Indeed the Committee concluded that "… lenalidomide was likely to be more effective than standard chemotherapy for the treatment of…" this group of patients.

In conclusion we ask that the Committee re-consider the evidence and revise the draft recommendation to take into account the changing treatment paradigm for these patients and to address the use of lenalidomide as second line therapy for all myeloma patients who have

previously received bortezomib. We very much hope that the Committee will support the use of lenalidomide as a second line treatment option to allow the improvements that have been observed in myeloma survival over the last 10 years to continue.

We would be very happy to engage and comment further on this consultation if this would be deemed helpful.

Yours faithfully

[redacted]

Janssen's Response to the Appraisal Consultation Document (ACD)

Lenalidomide for the treatment of multiple myeloma in people who have received at least one prior therapy with bortezomib (partial review of TA171)

Janssen is pleased to have the opportunity to provide our comments in relation to the second ACD for the partial review of TA171.

We are reassured that the ERG and Committee has recognised the impact on cost effectiveness of a more plausible treatment duration and administration cost associated with bortezomib re-treatment.

However, we remain concerned that statements in relation to the bortezomib response scheme in the ACD could cause confusion around the nature of this complex patient access scheme (PAS).

- As previously commented, Janssen is concerned that the use of the word '*discount*' (paragraph 3.25 and 3.58) to describe the manufacturer's approach to proxy the value of the complex PAS may be understood by third parties as an endorsement of the figure as an actual discount from bortezomib list price. This concern is compounded by the ACD highlighting only the 15% 'discount' in the manufacturer's basecase.
- At paragraph 4.14, the ACD refers to the complex PAS in the context of a '*pricing agreement*' between manufacturers and the Department of Health, and notes that NICE '*had no authority to assume a different **cost***' (emphasis added). We are concerned that the specific wording in this paragraph could be misinterpreted that the complex PAS is an adjustment to the bortezomib price.

Thus, we request that further consideration is given to the wording of the ACD with respect to the bortezomib acquisition cost. We feel strongly that the final NICE guidance must more accurately reflect that the bortezomib response scheme is a clinical outcome based scheme, rather than a discount to list price.

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Technology appraisals

Patient access scheme submission template

October 2009

1 Introduction

The 2009 Pharmaceutical Price Regulation Scheme (PPRS) (www.dh.gov.uk/en/Healthcare/Medicinespharmacyandindustry/Pharmaceuticalpriceregulationscheme/2009PPRS) is a non-contractual scheme between the Department of Health and the Association of the British Pharmaceutical Industry. The purpose of the 2009 PPRS is to ensure that safe and cost-effective medicines are available on reasonable terms to the NHS in England and Wales. One of the features of the 2009 PPRS is to improve patients' access to medicines at prices that better reflect their value through patient access schemes.

Patient access schemes are arrangements which may be used on an exceptional basis for the acquisition of medicines for the NHS in England and Wales. Patient access schemes propose either a discount or rebate that may be linked to the number, type or response of patients, or a change in the list price of a medicine linked to the collection of new evidence (outcomes). These schemes help to improve the cost effectiveness of a medicine and therefore allow the National Institute for Health and Care Excellence (NICE) to recommend treatments which it would otherwise not have found to be cost effective. More information on the framework for patient access schemes is provided in the 2009 PPRS (www.dh.gov.uk/en/Healthcare/Medicinespharmacyandindustry/Pharmaceuticalpriceregulationscheme/2009PPRS.

Patient access schemes are proposed by a pharmaceutical company and agreed with the Department of Health, with input from the Patient Access Schemes Liaison Unit (PASLU) within the Centre for Health Technology Evaluation at NICE.

2 Instructions for manufacturers and sponsors

This document is the patient access scheme submission template for technology appraisals. If manufacturers and sponsors want the National Institute for Health and Care Excellence (NICE) to consider a patient access scheme as part of a technology appraisal, they should use this template. NICE can only consider a patient access scheme after formal referral from the Department of Health.

The template contains the information NICE requires to assess the impact of a patient access scheme on the clinical and cost effectiveness of a technology, in the context of a technology appraisal, and explains the way in which background information (evidence) should be presented. If you are unable to follow this format, you must state your reasons clearly. You should insert 'N/A' against sections that you do not consider relevant, and give a reason for this response.

Please refer to the following documents when completing the template:

- 'Guide to the methods of technology appraisal' (http://publications.nice.org.uk/guide-to-the-methods-of-technology-appraisal-2013-pmg9)
- 'Specification for manufacturer/sponsor submission of evidence' (http://www.nice.org.uk/aboutnice/howwework/devnicetech/singletechnologyappraisalsubmissiontemplates.jsp) and
- Pharmaceutical Price Regulation Scheme 2009 (www.dh.gov.uk/en/Healthcare/Medicinespharmacyandindustry/Pharmaceuticalpriceregulationscheme/2009PPRS).

For further details on the technology appraisal process, please see NICE's 'Guide to the single technology appraisal (STA) process' and 'Guide to the multiple technology appraisal (MTA) process' (http://www.nice.org.uk/aboutnice/howwework/devnicetech/technologyappraisalprocessguides/technology_appraisal_process_guides.jsp). The 'Specification for manufacturer/sponsor submission of evidence' provides details on disclosure of information and equality issues.

Make the submission as brief and informative as possible. Only mark information as confidential when absolutely necessary. Sufficient information must be publicly available for stakeholders to comment on the full content of the technology appraisal, including details of the proposed patient access scheme. Send submissions electronically to NICE in Word or a compatible format, not as a PDF file.

Appendices may be used to include additional information that is considered relevant to the submission. Do not include information in the appendices that has been requested in the template. Appendices should be clearly referenced in the main submission.

When making a patient access scheme submission, include:

- an updated version of the checklist of confidential information, if necessary
- an economic model with the patient access scheme incorporated, in accordance with the 'Guide to the methods of technology appraisal' (http://publications.nice.org.uk/guide-to-the-methods-of-technology-appraisal-2013-pmg9).

If you are submitting the patient access scheme at the end of the appraisal process, you should update the economic model to reflect the assumptions that the Appraisal Committee considered to be most plausible. No other changes should be made to the model.

3 Details of the patient access scheme

3.1 Please give the name of the technology and the disease area to which the patient access scheme applies.

REVLIMID®▼ (lenalidomide) for the treatment of multiple myeloma (MM) in adult patients who have received a prior therapy with bortezomib.

3.2 Please outline the rationale for developing the patient access scheme.

The same Patient Access Scheme (PAS) is in operation for TA171[1] which currently restricts use of lenalidomide to after 2 prior therapies.

In England to date, lenalidomide after 1 prior therapy has been funded via the Cancer Drugs Fund. Because a PAS existed with TA171, Celgene voluntarily agreed to provide lenalidomide through the 'Revlimid Option Scheme' free of charge after 26 cycles of use. This scheme has been in operation since 2011 and has included patients backdated to June 2010. Therefore since the appraisal is a part review of existing guidance for which the same PAS is in place, Celgene have offered the scheme.

The overall aim of the scheme is to improve the cost-effectiveness of lenalidomide after bortezomib therapy and ensure value-for-money for the NHS.

3.3 Please describe the type of patient access scheme, as defined by the PPRS.

The patient access scheme is a complex, finance-based scheme. As described above, the scheme is already in operation for MM patients after two prior therapies (TA171) and also for patients with myelodysplastic syndromes associated with an isolated deletion 5q cytogenetic abnormality (TA322).

3.4 Please provide specific details of the patient population to which the patient access scheme applies. Does the scheme apply to the whole licensed population or only to a specific subgroup (for example, type of tumour, location of tumour)? If so:

- How is the subgroup defined?

- If certain criteria have been used to select patients, why have these have been chosen?
- How are the criteria measured and why have the measures been chosen?

The treatment position in the part-review of TA171 is specifically looking at patients who have received bortezomib as their initial treatment according to NICE TA228[2]. There are no further restrictions on the patient population eligible to benefit from the PAS. To date, these patients are either re-challenged with bortezomib or receive lenalidomide through the CDF (if prescribed prior to November 2015). A small amount 5-10% of patients will receive a chemotherapy agent such as melphalan in combinations with steroids.

3.5 Please provide details of when the scheme will apply to the population specified in 3.4. Is the scheme dependent on certain criteria, for example, degree of response, response by a certain time point, number of injections? If so:

- Why have the criteria been chosen?
- How are the criteria measured and why have the measures been chosen.

For any patient who stays on lenalidomide for more than 26 cycles, the treatment costs post cycle 26 will be met by Celgene Ltd.

I.e., the scheme is a capping scheme which limits costs to 26 cycles and from cycle 27 onwards, Celgene reimburse the NHS via either credit note, BACS rebate or by providing free stock.

3.6 What proportion of the patient population (specified in 3.4) is expected to meet the scheme criteria (specified in 3.5)?

A revised base case cost-effectiveness model was provided at the ACD stage. This used survival estimates from the mean of covariates survival approach and predicted that 23.9% of patients will reach a 27th treatment cycle, at which point lenalidomide will be received free-of-charge under the scheme. Using the corrected group prognosis method for covariate adjustment, this proportion is

estimated to be 24.7%. These estimates are therefore consistent in predicting that almost 1 in 4 patients is expected to meet the scheme criteria.

Since the ERG and the committee had serious concerns around the crossing of the overall survival and progression free survival curves in the model, these have been re-analysed using a multi-state modelling (MSM) approach (explained below) and using this updated analysis, the model still predicts that 24.7% of patients will reach cycle 27.

3.7 Please explain in detail the financial aspects of the scheme. How will any rebates be calculated and paid?

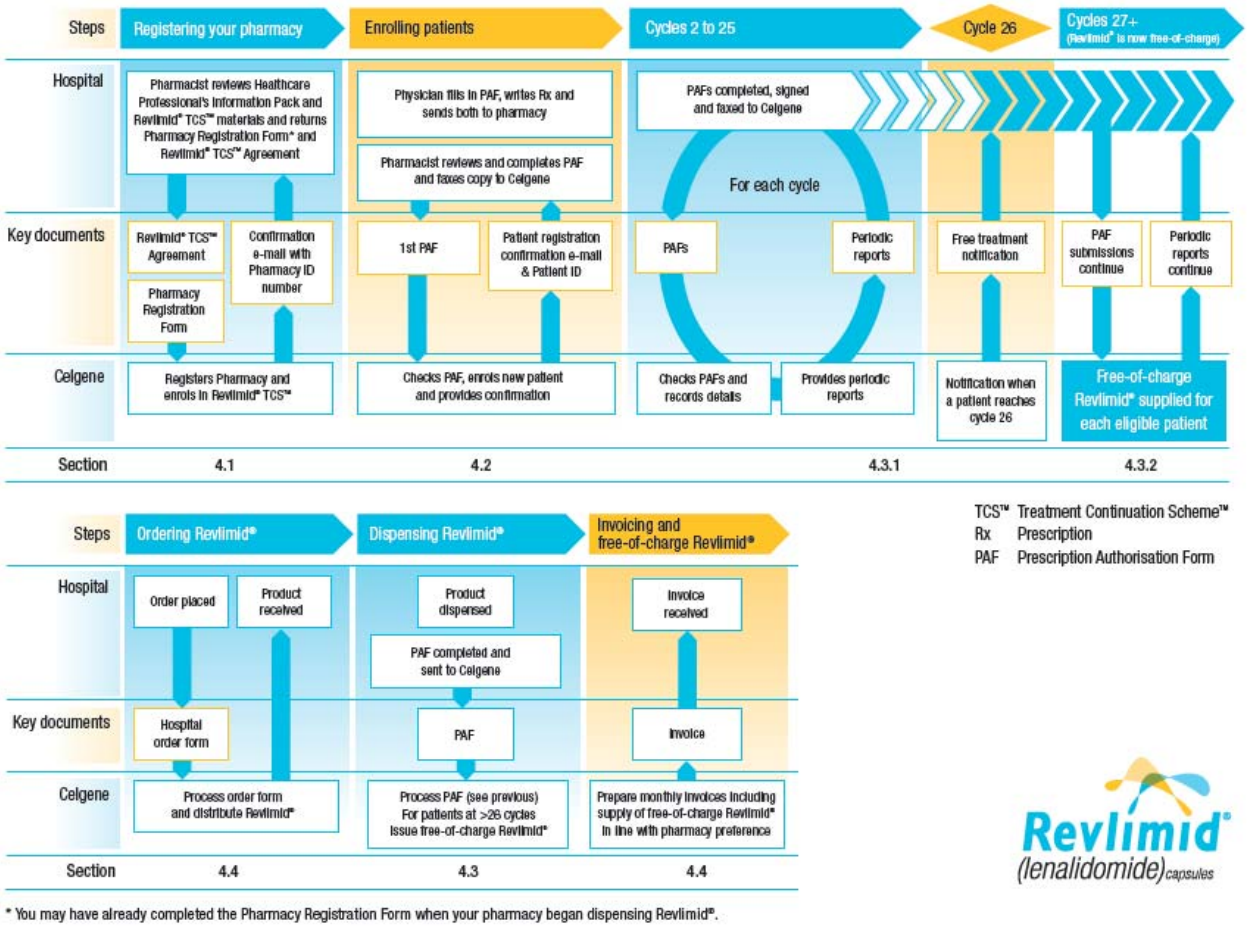
Celgene offers 3 options with the scheme and NHS organisations can either choose to receive a credit note, free stock or a rebate through BACS. With the free stock option, when Celgene receives a PAF for the 27th and beyond cycle, it simply supplies free stock. With the credit note, the NHS organisation can use this against any invoice for stock already used or being ordered.

3.8 Please provide details of how the scheme will be administered. Please specify whether any additional information will need to be collected, explaining when this will be done and by whom.

Upon a patient reaching their 27th cycle, a letter is sent to the Chief Pharmacist or other designated individual each time a credit note, free stock unit, or BACs rebate is issued. This letter will detail the qualifying patients and the cycle numbers to which the free treatment cycles relate. A copy of the letter is also included with each credit note and free stock delivery in order to allow cross reference. This is automatic from when a PAF is received at Celgene so a request/claim is not required.

3.9 Please provide a flow diagram that clearly shows how the scheme will operate. Any funding flows must be clearly demonstrated.

Figure 1 – PAS Flow diagram



3.10 Please provide details of the duration of the scheme.

The Scheme will run until such time that the NICE guidance is withdrawn or reviewed, or potentially if there is change in UK policy and PPRS to allow for differential pricing across indications.

3.11 Are there any equity or equalities issues relating to the scheme, taking into account current legislation and, if applicable, any concerns identified during the course of the appraisal? If so, how have these been addressed?

None have been identified.

3.12 If available, please list any scheme agreement forms, patient registration forms, pharmacy claim forms/rebate forms, guides for pharmacists and physicians and patient information documents. Please include copies in the appendices.

The following forms are relevant to the PAS and the latest versions are included in the appendices below:

- TCS™ agreement form.
- Pharmacy Registration form.

3.13 In the exceptional case that you are submitting an outcome-based scheme, as defined by the PPRS, please also refer to appendix B.

4 Cost effectiveness

4.1 If the population to whom the scheme applies (as described in sections 3.4 and 3.5) has not been presented in the main manufacturer/sponsor submission of evidence for the technology appraisal (for example, the population is different as there has been a change in clinical outcomes or a new continuation rule), please (re-)submit the relevant sections from the 'Specification for manufacturer/sponsor submission of evidence' (particularly sections 5.5, 6.7 and 6.9). You should complete those sections both with and without the patient access scheme. You must also complete the rest of this template.

N/A

4.2 If you are submitting the patient access scheme at the end of the technology appraisal process, you should update the economic model to reflect the assumptions that the Appraisal Committee considered to be most plausible. No other changes should be made to the model.

As described in the text for section 4.2, there have been changes made to the model to reflect the assumptions which the Appraisal Committee considered to be most plausible and to address areas where they felt there were high levels of uncertainty. The systematic literature review has also been updated to identify relevant information since the last run; as a significant amount of time has passed. The changes will be described below and the new base-case explained and justified.

In the scenario analyses, the previous assumptions will also be presented and tested.

The main areas identified by the Appraisal Committee and evidence review group (ERG) as causing high levels of uncertainty or not being supported by clinical practice.

The following areas were identified as leading to high levels of uncertainty or not-being aligned to clinical practice in the UK as defined by the Appraisal Committee's interpretation of the advice of the clinical expert and the views of the ERG:

- Relationship between OS and progression-free survival (PFS) – the OS and PFS curves in the model should not cross.
- Covariate adjustment – The evidence review group preferred the corrected group prognosis method to the mean of covariates method.
- Overall survival (OS) and use of medians – Use of median values suggest use of an exponential distribution and proportional hazards between treatments, but the exponential distribution fitted the data poorly and the validity of assuming proportional hazards was unknown.
- Drug costs – Bortezomib costs should be capped at 8 cycles of treatment.
- Resource use costs – transportation costs were hard to quantify in clinical practice.
- 3rd line inclusion of lenalidomide in the comparator arm – Celgene disagree that excluding lenalidomide from the comparator arm would reflect clinical practice (please see section 4.11 below).

Updates Since the last Appraisal Committee Meeting

Updated systematic review

An updated systematic review was conducted in December 2015 to identify additional evidence for multiple myeloma patients treated with the comparator therapies identified as relevant during previous appraisal committee meetings. These were:

- Bortezomib retreatment ± steroids

- Chemotherapy ± steroids (including melphalan and cyclophosphamide)
- Bendamustine ± steroids

The update search covered the time period between August 2013 and December 2015; search terms and inclusion/exclusion criteria were consistent with the searches provided in the original submission (see Appendix C: Systematic Review Methods).

Following completion of the search, one paper of potential relevance was identified across all interventions included in the original searches (see Figure 2). Narrowing down to the relevant comparator therapies gave 11 papers of potential interest. These papers were reviewed to determine whether the information contained within the papers would add to the weight of evidence available for both comparators based upon the following criteria:

- Provision of Kaplan–Meier data for overall survival.
- Provision of Kaplan–Meier data for progression-free survival.
- Sample size.

Data extraction for the relevant studies can be found in Table 1. The provision of Kaplan–Meier data for survival was considered of primary importance as previously Kaplan–Meier data were only available for one comparison (versus chemotherapy), with no paper providing anything other than median survival to allow comparison to bortezomib retreatment.

Two additional papers were identified providing overall survival Kaplan–Meier data (Ahn et al. 2014[3]; San-Miguel et al. 2015[4]). The Ahn paper additionally provided Kaplan–Meier data for progression-free survival and was therefore selected for inclusion within the economic model.

Information was not available within the San-Miguel paper on patient characteristics in the relevant patient subgroup (those who received prior bortezomib) and additionally all patients had received prior lenalidomide. This paper was therefore not included within the economic model.

Ahn 2014 (used for a scenario analysis below) presents results for 30 Korean relapsed, refractory multiple myeloma (RRMM) patients. Patients received retreatment with BOR in combination with cyclophosphamide and thalidomide for a median of 6 (2-12) cycles. All patients had received prior therapy, with 57%, 20%, 13% and 10% receiving 2, 3, 4 and 5 prior lines of therapy, respectively. In addition to this, 50% of patients had also received an SCT. The median age of patients was 62 (51-81) years of age, and the median time from diagnosis was 43.6 (16.9-249.6) months. Median TTP was 5.8 (2.6-9) months, median PFS was 5.8 (4.2-6.8) months, and median OS was 17.6 (14.4–23.5) months. It was noted by the author that although the data represented a Korean cohort, OS and PFS estimates were consistent with Caucasian data.

Figure 2 - Preferred Reporting Items for Systematic Reviews and Meta-Analyses diagram

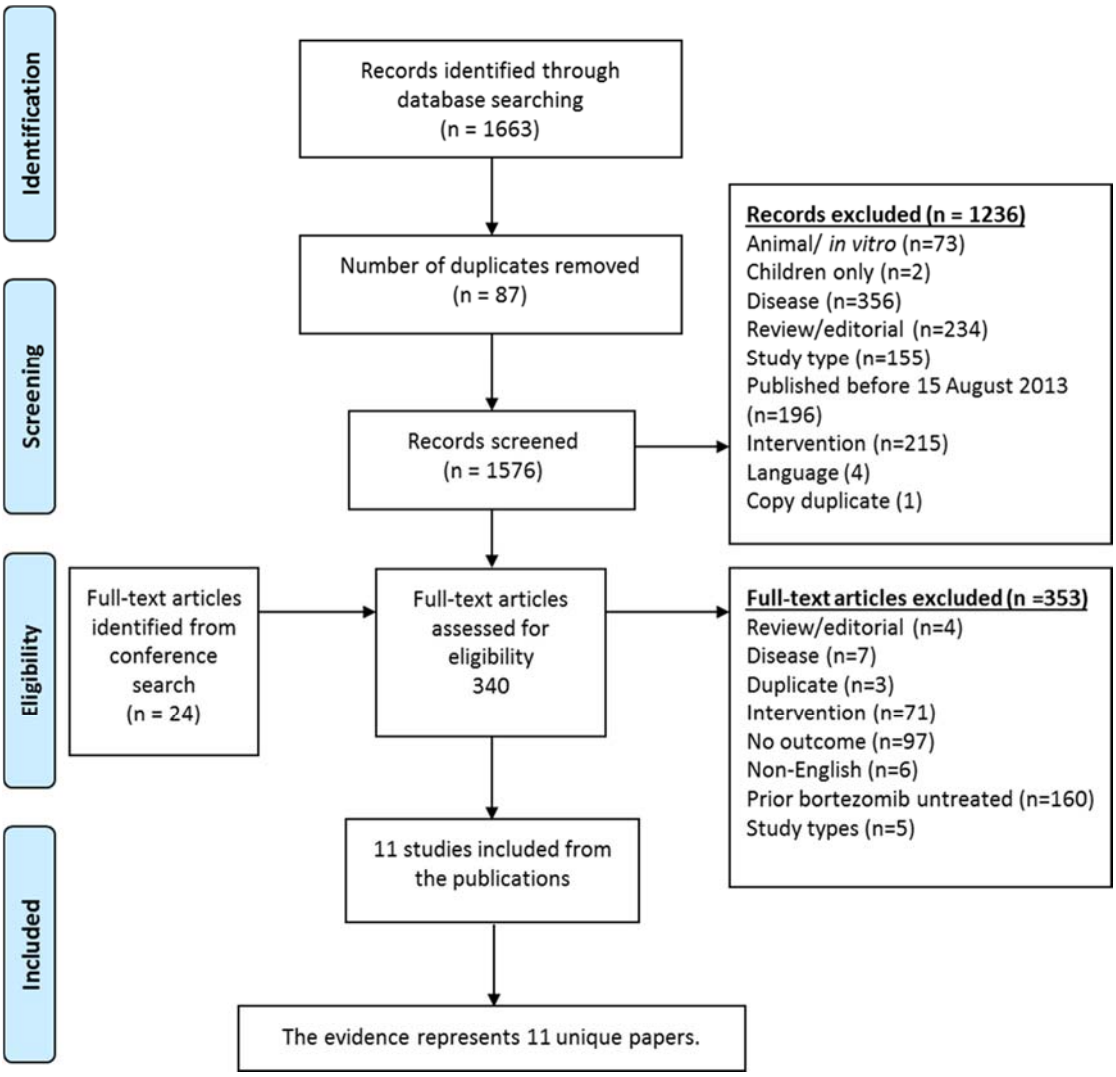


Table 1: Data extraction

Publication details	San-Miguel 2015[4]	Hulin 2013[5]	Ahn 2014[3]	Berenson 2014[6]	Kim 2015[7]	Mao 2014[8]	Orio 2014[9]	Cerchione 2014[10]	Reece 2014[11]	Jagannath 2015[12]	Cerchione 2013[13]
Study type	RCT	POS	ROS	ROS	ROS	ROS	ROS	ROS	ROS	ROS	ROS
Treatment regimen of interest	BOR	BOR	BOR	BOR	CTD	BOR	BOR	BEND	CyBorP/D	BOR	BEND
Geographical region	Worldwide	Worldwide	Korea	NR	Korea	China	EU	NR	NR	NR	NR
N (on arm of interest)	51*	96	30	58	6	20	35	24	98	69	16
% previous use of DOX	NR	NR	NR	NR	NR	100	11	NR	NR	NR	NR
Lines of previous therapy (median or mean no. / % with 1 prior therapy)	NR	NR	2- 57%, 3 -20%, 4-13%, 5 - 10%	2 (2-10)	75% (1 prior)	4 (2-11)	100% (1 prior)	6.3 (4-8)	2 (1-6)	2	5.7 (4-8)
ISS stage	NR	II - 19	I -15/25, II - 5/25, III - 6/25	NR	NR	NR	NR	ISS was equally distributed	NR	III - 46.9%	ISS was equally distributed
Beta 2M	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
% worsening extramedullary disease	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR

Age (years)	NR	62 (34–80)	67 (51-81)	64 (31-84)	75.1 (66.7-78.8)	63 (39-72)	63 (34-84)	66 (48-83)	64 (36-88)	68	62.6 (39-82)
Time from diagnosis (months)	NR	NR	43.6 (16.9-249.6)	NR	NR	NR	0.8 (0-3)	NR	NR	NR	NR
% receiving concomitant steroids	NR	41	76.7	NR	100	100	51	8%	100	NR	NR
% receiving other relevant concomitant therapies (chemo-therapy)	NR	NR	100	NR	0	100	81	100	NR	NR	100
Previous SCT (%)	NR	NR	50	NR	0	5	15	58	75	46.8	69
Cycles of therapy received	NR	4	6 (2-12)	NR	15 (4-23)	2 (2-4)	4 (3-13)	4.3 (2-9)	5 (-47)	NR	4.7 (2-6)
Response rate	NR	NR	73.3	NR	16.7	50	NR	62.5	68	NR	68
Median OS (95% CI)	19.5 (14.1-32.5)	17.6 (14.4-23.5)	13.4 (6.1-20.7)	NR	Not reached	NR	NR	6.7 (2-19)	24.2 (20.2-28.1)	NR	3.6 (2-6)
OS KM included (y/n)	Y	N	Y	NR	N	N	N	N	N	N	N
Median TTP (95% CI)	NR	NR	5.8 (2.6-9	NR	NR	NR	NR	NR	NR	NR	NR

TTP KM included (y/n)	N	N	Y	N	N	N	N	N	N	N	N
Median PFS (95% CI)	N	6.9 (4.6–8.2)	5.8 (4.2-6.8)	NR	not reached	20 (6-32	15 (1-40)	NR	14.9 (12-17.9)	7.3 (4.2-10.8)	NR
TTP/PFS KM included (y/n)	N	N	Y	NR	N	N	Y	N	N	N	N
Median (or mean) time on treatment (95% CI or SD)	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
Time on treatment KM included (y/n)	N	N	N	N	N	N	N	N	N	N	N

Key: BOR, bortezomib; BEND, bendamustine; CTD, Cyclophosphomide, Thalidomide, Dexamethasone; CyBorP, Cyclophosphomide, bortezomib, prednisone; D, dexamethasone; POS, prospective observational study; RCT, randomised control trial; ROS, retrospective observational study.

Note: *, Sub-population who had been previously treated with bortezomib and an IMiD; outcomes were not reported separately for patients only previously treated with bortezomib.

Updates to the model – MSM approach

Research was conducted to identify a suitable method to produce a combined model of OS and PFS. A multi-state Markov modelling approach[14] was selected for this analysis as it allowed both:

- Simultaneous modelling of the transitions between progression and death which prevents the survival cures for PFS and OS crossing.
- Covariate adjustment of survival curves – required in order to be able to use all the data from the MM-009 and MM-010 trials to estimate outcomes in a second line patient population and to be able to produce comparative effectiveness estimates

Simultaneous modelling of the transitions between progression and death accounts for patients who 'skip' a state. For the previously adopted method, patients who died without progression being observed would have had their death time used as the date of event for both PFS and OS. The multi-state model estimates the probability of moving directly from pre-progression to death without observing progression. As multi-state modelling can model all possible transitions, the resultant transition probability matrices provide a more realistic model for the movement between states.

Although multi-state models use the Markov assumption (the future evolution only depends on the current state), the hazard of moving between states and therefore transition probabilities matrices can be calculated such that the probability of moving between states varies based upon the time elapsed since randomisation. The time point at which hazards are allowed to change will be based upon observed change in Kaplan–Meier data; this is considered to represent a similar approach to using piecewise models to extrapolate survival data. Models were implemented in the R package msm[15].

Methods

A three-state multi-state Markov model was fitted to the combined data from the MM-009 and MM-010 trials. Patients were defined as being in one of three states; pre-progression, post progression and death. Progression was

determined using the central review assessment of progression. Models were developed according to the following procedure:

1) Unadjusted models were fitted to determine whether the hazard of moving between states should be constant over time, or allowed to vary based upon the time elapsed since randomisation: a time homogeneous or inhomogeneous model, respectively. This was performed by comparing the Kaplan–Meier curves for PFS and OS to the survival probabilities estimated using transition probability matrices estimated from the model. This assessment was based upon a visual assessment of the Kaplan Meier curves and log cumulative hazard plots and by using a likelihood ratio test to compare the time inhomogeneous and homogeneous models.

2) Potentially relevant prognostic covariates were identified in the same manner submitted in response to the first ACD using the publications, clinical study reports (CSRs) from the MM-009 and MM-010 trials and clinician feedback.

3) Forward selection regression analysis was performed to inform the selection of covariates into the final model. Covariates were included in the model if the p-value obtained from a likelihood ratio test was <0.05.

4) Using covariate adjusted models, covariate adjusted transition probability matrices were derived and used to estimate the probability of remaining in each state (pre-progression, post progression and death) over time for each sub-group of patients. Uncertainty in the estimation of the probability transition matrices was quantified by creating a 100 bootstrap samples of the transition matrices using the boot.msm function in R.

Step 1: Unadjusted models

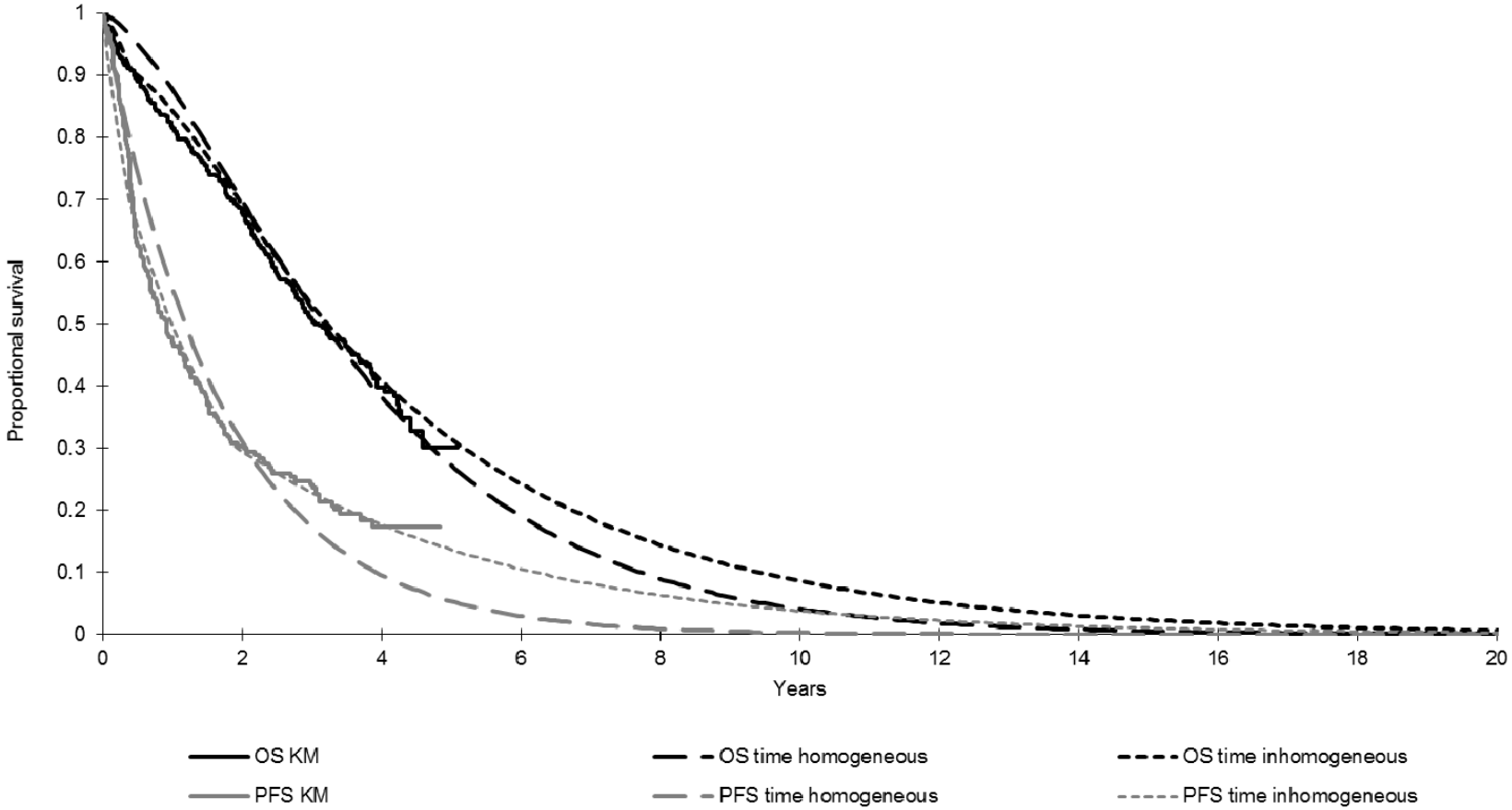
Figure 3 presents the Kaplan–Meier graphs of PFS and OS for the combined data from the MM-009 and MM-010 trials. Superimposed on to the Kaplan–

Meier curves are survival probabilities estimated from time homogeneous and in-homogenous, multi-state Markov models.

Visual inspection of survival probabilities estimated using the time homogeneous Markov model suggested that the fit of the model could be improved by allowing the hazard of moving between states to vary over time: a time in-homogeneous model.

Based upon the Kaplan-Meier and log-cumulative hazard plots for OS, a change was observed in the hazard of death in the first 6 months and for PFS after approximately 2 years. Based upon this information the hazard of moving between states was allowed to vary at 168 and 728 days in the statistical model. These time points were selected to coincide with the 28 day cycle used within the economic model. There was a statistically significant improvement in model fit using this time inhomogeneous model using the likelihood ratio test ($p<0.01$).

Figure 3: Survival probabilities estimated from time homogeneous and time in-homogeneous multi-state Markov models



Key: KM, Kaplan–Meier; OS, overall survival; PFS, progression-free survival.

Step 2: Identification of relevant covariates

The variables identified as potentially relevant determinants of clinical outcomes, based upon analysis presented in the CSRs for MM-010 and MM-009 and associated publications, are presented in Table 2. The CSRs were searched first, and additional variables were added from relevant publications. Clinical experts were also consulted via personal communication as part of the previous response to NICE to inform the relevant covariates, and all variables highlighted by clinicians were included.

Table 2: Variables identified for testing as potentially relevant prognostic factors

Source for inclusion	Variable	Definition
MM-009 CSR (2007a)	Prior melphalan	Yes or no
MM-010 CSR (2007b)	Prior thalidomide	Yes or no
	Prior doxorubicin	Yes or no
	Prior bortezomib	Yes or no
	Worsening extramedullary plasmacytoma disease	Yes or no
	Baseline bone marrow cellularity-aspirate/biopsy	Normal, hyperplasia, hypoplasia or missing
Dimopoulos et al (2009)	Baseline beta-2 microglobulin	>2.5 or ≤2.5 mg/L
	Duration of multiple myeloma	Years
	Lytic bone lesions present at baseline	Yes or no
	Prior high-dose therapy or stem cell transplant	Yes or no
	Prior dexamethasone	Yes or no
	Baseline ISS score	1, 2 or 3
	Baseline plasma cell %	High or low [0-3% vs 3+]
Stadtmauer et al (2009)	Number of prior anti-myeloma therapies	1 or 2+ Note: required to produce estimates for the population of interested (1 prior therapy)
	Age	Years
Clinical advice	Sex	Male or female

Key: CSR, clinical study report; ISS, International Staging System.

Step 3: Stepwise regression to inform selection of covariates

Forward selection regression analysis was performed for the full intent-to-treat (ITT) population data, using all covariates identified as being potential prognostic factors (listed in Table 2). Prior therapy was added to the model at step 1, as this covariate is required to present outcomes for a second-line population. Variables with the lowest p-value were added to the model in turn until no variable resulted in an improved model (p-value<0.05). The results of the forward selection regression analyses are provided in Appendix D: Distribution of patients' covariate combinations and transition probability matrices.

Baseline beta-2 microglobulin [>2.5 or ≤2.5 mg/L], prior doxorubicin [Yes or No] and worsening extramedullary plasmacytoma disease [Yes or no] were found to be statistically significant prognostic factors in the three-state Multi-state Markov model. These covariates, along with number of prior therapies (one or two or more) were selected for inclusion in the final adjusted survival model. Transition probability matrices for each possible covariate combination are provided in Appendix D: Distribution of patients' covariate combinations and transition probability matrices.

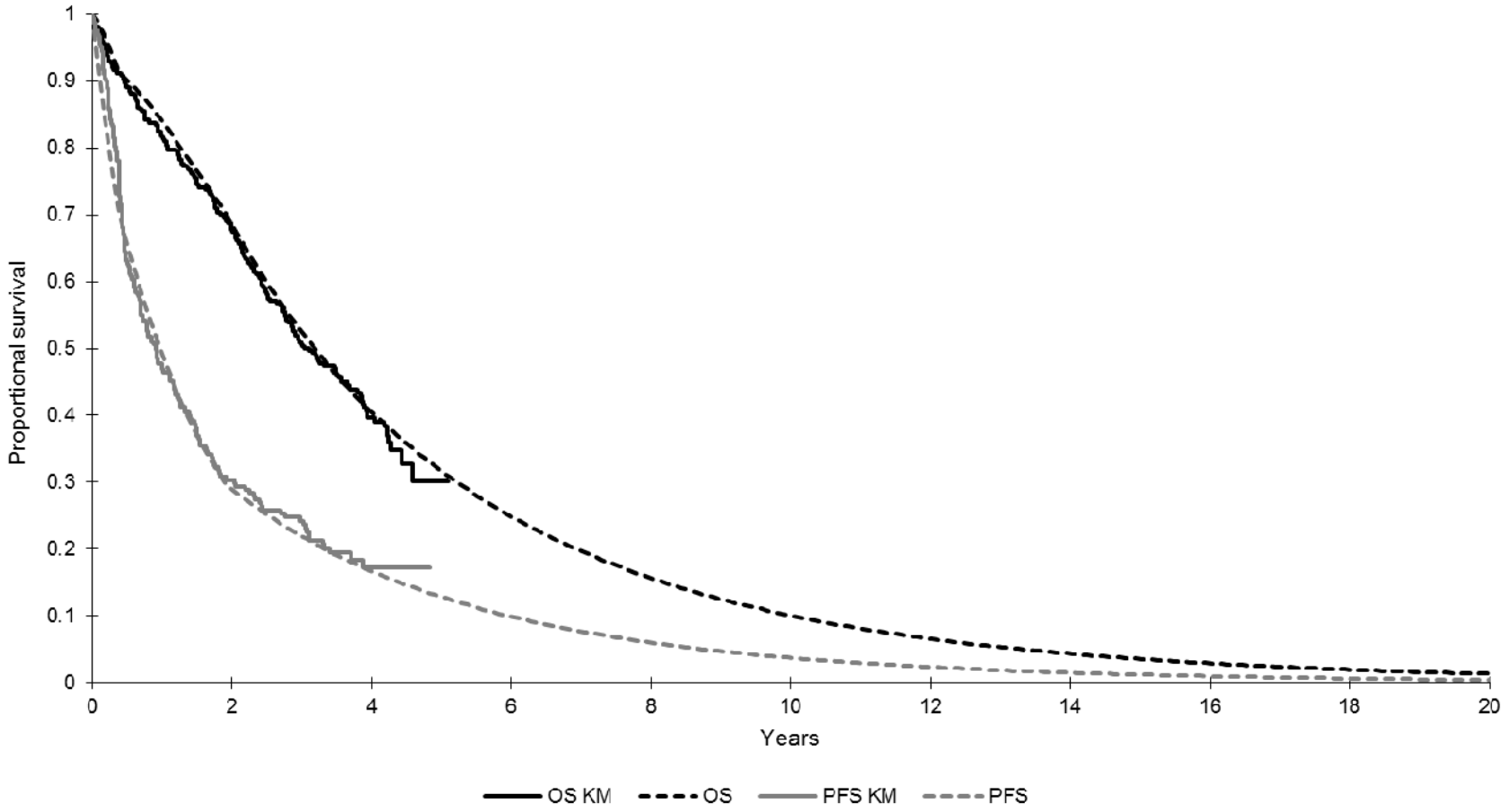
Step 4: Production of adjusted survival models

Using the covariate adjusted multi-state Markov model, predictions of PFS and OS were performed using the corrected group prognosis method. The corrected group prognosis method was previously highlighted by the ERG as being a potentially superior method to the mean of covariates method used to estimate OS and PFS in the original submission. Rather than applying a mean covariate value to all patients, this involves predicting the relevant clinical outcome for each 'possible patient' and calculating the average survival (weighted by the chance of each possible patient occurring).

For example, 70.8% of pooled MM-010 and MM-009 patients had a beta-2 microglobulin level of >2.5mg/L. Application of this value (0.708) is sufficient for a mean of covariates model, by assuming that the 'average patient' has a 70.8% likelihood of having beta-2 microglobulin of >2.5mg/L. The corrected group prognosis approach would instead estimate survival for a patient who presents

with beta-2 microglobulin >2.5mg/L and every possible combination of other covariates, and then do the same for a patient who presents with beta-2 microglobulin level of ≤2.5mg/L. The resulting array of survival estimates are then weighted by the likelihood of each type of patient presenting and summed to give a corrected group prognosis survival estimate. Figure 4 presents predictions of PFS and OS using the corrected group prognosis method where the hazard of moving between states is free to vary at 168 and 728 days. Figure 5 presents predictions of PFS and OS by prior therapy; one or two or more prior therapies at baseline.

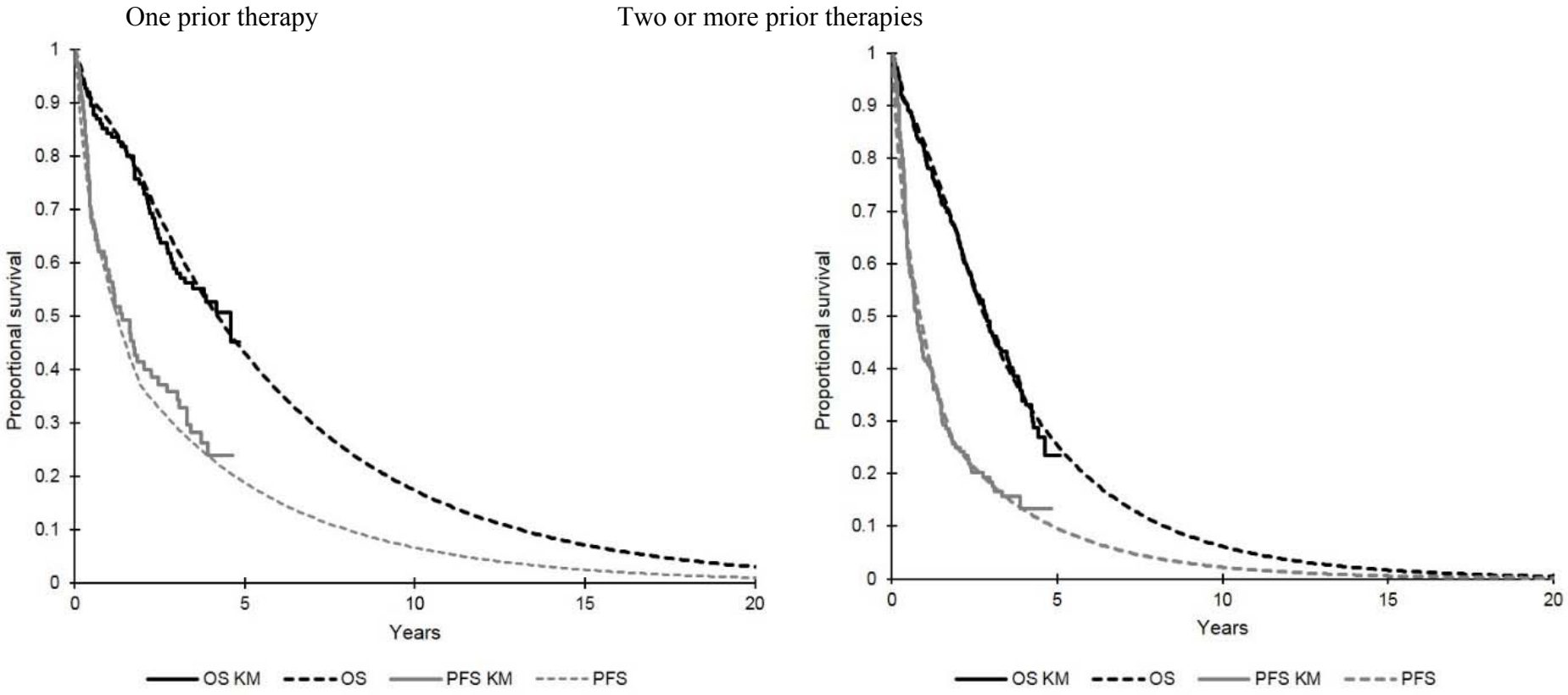
Figure 4: Survival probabilities estimated from covariate adjusted, time-inhomogeneous* multi-state Markov models



Key: KM, Kaplan–Meier; OS, overall survival; PFS, progression-free survival.

Notes: *The hazard of moving between states is free to vary at 168 and 728 days in the statistical model

Figure 5: Survival probabilities estimated from covariate adjusted, time-inhomogeneous* Multi-state Markov models – by prior therapy



Key: KM, Kaplan–Meier; OS, overall survival; PFS, progression-free survival.

*The hazard of moving between states is free to vary at 168 and 728 days in the statistical model

Updates to the model – Other updates impacting the base-case

The following additional updates have been made to model:

- Incorporation of the patient access scheme.
- Correction of the ERG's overestimation of the costs for bortezomib (in the last ERG model, the ERG limited bortezomib to 8 cycles, however, the limit was set at 8, 4 week model cycles, not 8 bortezomib treatment cycles of 3 weeks).
- Full implementation of correction to the calculation of pre-progression QALYs (the ERG correction was only implemented in the first cycle of the comparator flow sheets, not throughout the pre-progression state).
- Removal of transport costs for patients from the model base case.
- Limiting bortezomib retreatment to 8 cycles in the base case.

4.3 Please provide details of how the patient access scheme has been incorporated into the economic model. If applicable, please also provide details of any changes made to the model to reflect the assumptions that the Appraisal Committee considered most plausible.

The cost of lenalidomide has been capped at 26 cycles. In the health-economic model this can be seen in the PF lenalidomide sheet.

The updates to the model to reflect the views of the Appraisal Committee are explained in section 4.2 above.

4.4 Please provide the clinical effectiveness data resulting from the evidence synthesis and used in the economic model which includes the patient access scheme.

The clinical effectiveness data remains unchanged with the exception of the additional data presented for bortezomib retreatment in the scenario analysis.

4.5 Please list any costs associated with the implementation and operation of the patient access scheme (for example, additional pharmacy time for stock management or rebate calculations). A

suggested format is presented in Table 3. Please give the reference source of these costs. Please refer to section 6.5 of the 'Specification for manufacturer/sponsor submission of evidence'

Table 3 - Costs associated with the implementation and operation of the patient access scheme (PAS)

	Calculation of cost	Reference source
Stock management	N/A	
Administration of claim forms	N/A	
Staff training	N/A	
Other costs…	N/A	
…		
…		
Total implementation/ operation costs	N/A	

There are no incremental costs associated with the operation of the patient access scheme at this line. This was discussed in depth with PASLU. As the PAS is built around the monitoring requirements of the Pregnancy Prevention Program (PPP) and is already in place as the REVLIMID Options Scheme™ for patients with existing funding through the Cancer Drugs Fund, there will be no impact on management in clinical practice.

4.6 Please provide details of any additional treatment-related costs incurred by implementing the patient access scheme. A suggested format is presented in Table 4. The costs should be provided for the intervention both with and without the patient access scheme. Please give the reference source of these costs.

Table 4 - Additional treatment-related costs for the intervention both with and without the patient access scheme (PAS)

	Intervention without PAS		Intervention with PAS		Reference source
	Unit cost (weighted average cycle cost) (£)	Total cost (£)	Unit cost (weighted average cycle cost) (£)	Total cost (£)	
Interventions	3,772	90,945	3,772	61,856	Health Economic model
Monitoring and diagnostic tests	various	12,115	various	12,115	Health Economic model
Administration	162	162	162	162	Health Economic model
Total treatment-related costs		110,976		81,887	Health Economic model

As stated above, there are no costs associated with the operation of the patient access scheme at this line.

1.1 Summary results

1.1.1 Base-case analysis

4.7 Please present in separate tables the cost-effectiveness results as follows.[1]

- the results for the intervention without the patient access scheme
- the results for the intervention with the patient access scheme.

- A suggested format is shown below (Table 5).

[1] For outcome-based schemes, please see section 5.2.8 in appendix B.

Table 5 - Base-case cost-effectiveness results without PAS

	Intervention	Bortezomib retreatment	Melphalan plus Prednisone
Intervention cost (£)	90,945	19,036	926
Other costs (£)	20,031	39,672	41,684
Total costs (£)	110,976	58,708	42,610
Difference in total costs Lenalidomide vs comparator (£)	-	52,268	68,366
LYG	5.867	3.685	3.153
LYG difference Lenalidomide vs comparator	-	2.182	2.714
QALYs	3.546	2.374	1.880
QALY difference Lenalidomide vs comparator	-	1.172	1.666
ICER Lenalidomide vs comparator (£) per QALY	-	**44,605**	**41,030**

LYG: life-year gained; QALY: quality-adjusted life-year; ICER: incremental cost-effectiveness ratio.

Table 6 - Base-case cost-effectiveness results with PAS

	Intervention	Bortezomib retreatment	Melphalan plus Prednisone
Intervention cost (£)	61,856	19,036	926
Other costs (£)	20,031	39,672	41,684
Total costs (£)	81,887	58,708	42,610
Difference in total costs Lenalidomide vs comparator (£)	-	23,179	39,277
LYG	5.867	3.685	3.153
LYG difference Lenalidomide vs comparator	-	2.182	2.714
QALYs	3.546	2.374	1.880
QALY difference Lenalidomide vs comparator	-	1.172	1.666
ICER Lenalidomide vs comparator (£) per QALY	-	**19,781**	**23,572**

LYG: life-year gained; QALY: quality-adjusted life-year; ICER: incremental cost-effectiveness ratio.

4.8 Please present in separate tables the incremental results as follows. [2]

- the results for the intervention without the patient access scheme
- the results for the intervention with the patient access scheme.

List the interventions and comparator(s) from least to most expensive. Present the incremental cost-effectiveness ratios (ICERs) in comparison with baseline (usually standard care), and the incremental analysis ranking technologies in terms of dominance and extended dominance. A suggested format is presented in table 4.

Table 7 - Updated Base-case incremental results without PAS

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£) versus MP (QALYs)	ICER (£) incremental (QALYs)
Melphalan and Prednisone	42,610	3.153	1.880	-	-	-	-	-
bortezomib retreatment	58,708	3.685	2.374	16,098	0.533	0.494	32,557	32,557
Lenalidomide plus dexamethasone	110,976	5.867	3.546	52,268	2.182	1.172	**41,030**	**44,605**

LYG: life-year gained; QALY: quality-adjusted life-year; ICER: incremental cost-effectiveness ratio.

[2] For outcome-based schemes, please see section 5.2.9 in appendix B.

Table 8 - Updated Base-case incremental results with PAS

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£) versus MP (QALYs)	ICER (£) incremental (QALYs)
Melphalan and Prednisone	42,610	3.153	1.880	-	-	-	-	-
bortezomib retreatment	58,708	3.685	2.374	16,098	0.533	0.494	32,557	32,557
Lenalidomide plus dexamethasone	81,887	5.867	3.546	23,179	2.182	1.172	**23,572**	**19,781**

LYG: life-year gained; QALY: quality-adjusted life-year; ICER: incremental cost-effectiveness ratio.

1.1.2 Sensitivity analyses

4.9 Please present deterministic sensitivity analysis results as described for the main manufacturer/sponsor submission of evidence for the technology appraisal. Consider using tornado diagrams.

Figure 6: Tornado Diagram vs bortezomib

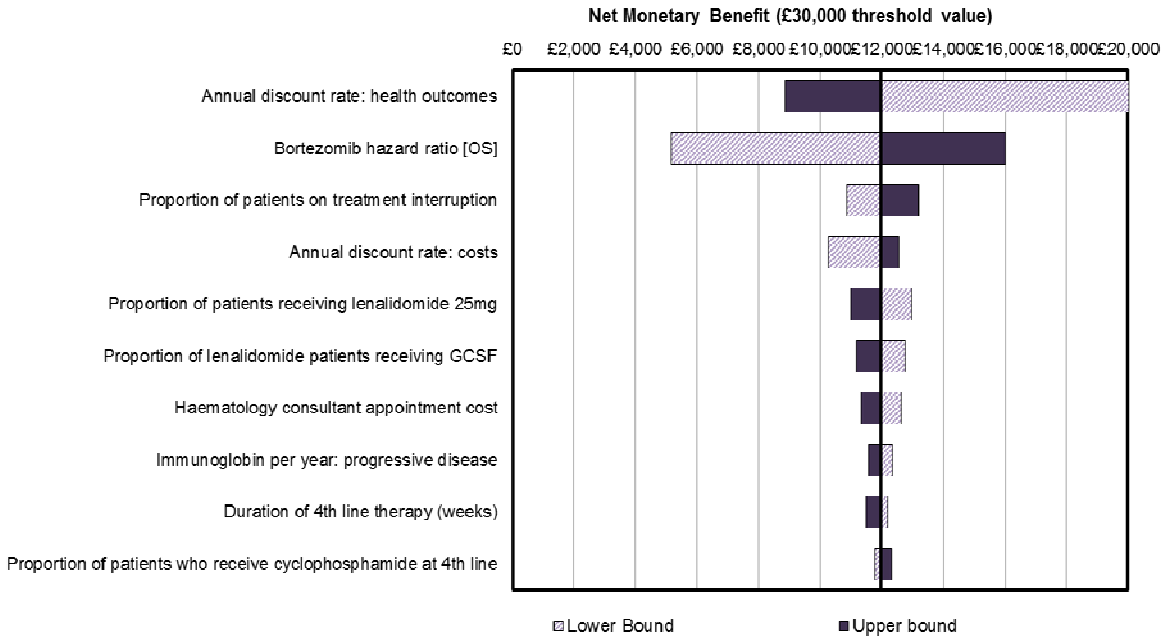


Table 9:Deterministic Sensitivity Analysis vs bortezomib

	Parameter	Lower Bound	Upper bound	Difference	LB ICER	UB ICER
1	Annual discount rate: health outcomes	£21,633	£8,864	£12,770	£15,518	£21,702
2	Bortezomib hazard ratio [OS]	£5,164	£16,007	£10,843	£24,545	£17,834
3	Proportion of patients on treatment interruption	£10,868	£13,179	£2,311	£20,726	£18,753
4	Annual discount rate: costs	£10,267	£12,531	£2,264	£21,238	£19,306
5	Proportion of patients receiving lenalidomide 25mg	£12,980	£10,993	£1,987	£18,923	£20,619
6	Proportion of lenalidomide patients receiving GCSF	£12,767	£11,153	£1,613	£19,105	£20,482
7	Haematology consultant appointment cost	£12,628	£11,320	£1,308	£19,223	£20,339
8	Immunoglobin per year: progressive disease	£12,356	£11,569	£787	£19,455	£20,127
9	Duration of 4th line therapy (weeks)	£12,196	£11,491	£705	£19,592	£20,194
10	Proportion of patients who receive cyclophosphamide at 4th line	£11,787	£12,324	£537	£19,941	£19,483

Figure 7: Tornado Diagram vs MP

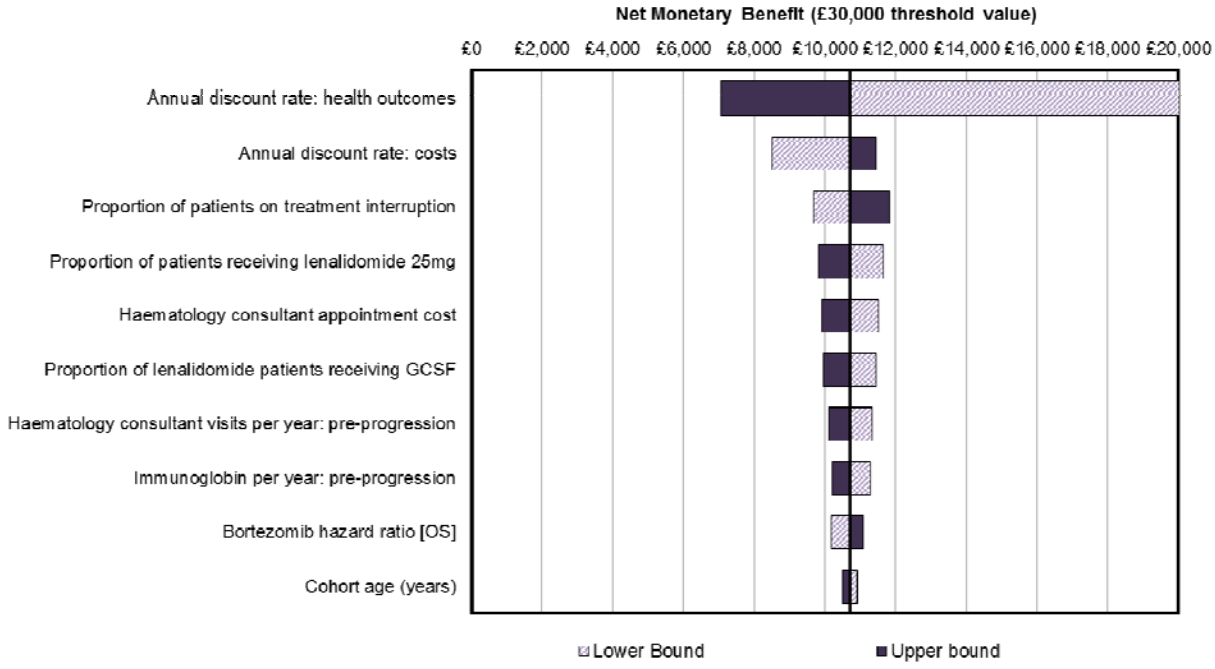


Table 10:Deterministic Sensitivity Analysis vs MP

	Parameter	Lower Bound	Upper bound	Difference	LB ICER	UB ICER
1	Annual discount rate: health outcomes	£21,873	£7,060	£14,813	£19,269	£25,429
2	Annual discount rate: costs	£8,505	£11,457	£2,952	£24,896	£23,124
3	Proportion of patients on treatment interruption	£9,680	£11,832	£2,153	£24,191	£22,899
4	Proportion of patients receiving lenalidomide 25mg	£11,648	£9,796	£1,851	£23,010	£24,121
5	Haematology consultant appointment cost	£11,508	£9,913	£1,595	£23,093	£24,051
6	Proportion of lenalidomide patients receiving GCSF	£11,449	£9,945	£1,504	£23,129	£24,031
7	Haematology consultant visits per year: pre-progression	£11,322	£10,099	£1,222	£23,205	£23,939
8	Immunoglobin per year: pre-progression	£11,293	£10,201	£1,092	£23,223	£23,878
9	Bortezomib hazard ratio [OS]	£10,182	£11,086	£904	£23,577	£23,579
10	Cohort age (years)	£10,902	£10,497	£405	£23,482	£23,673

4.10 Please present any probabilistic sensitivity analysis results, and include scatter plots and cost-effectiveness acceptability curves.

Figure 8: Scatter plot vs bortezomib

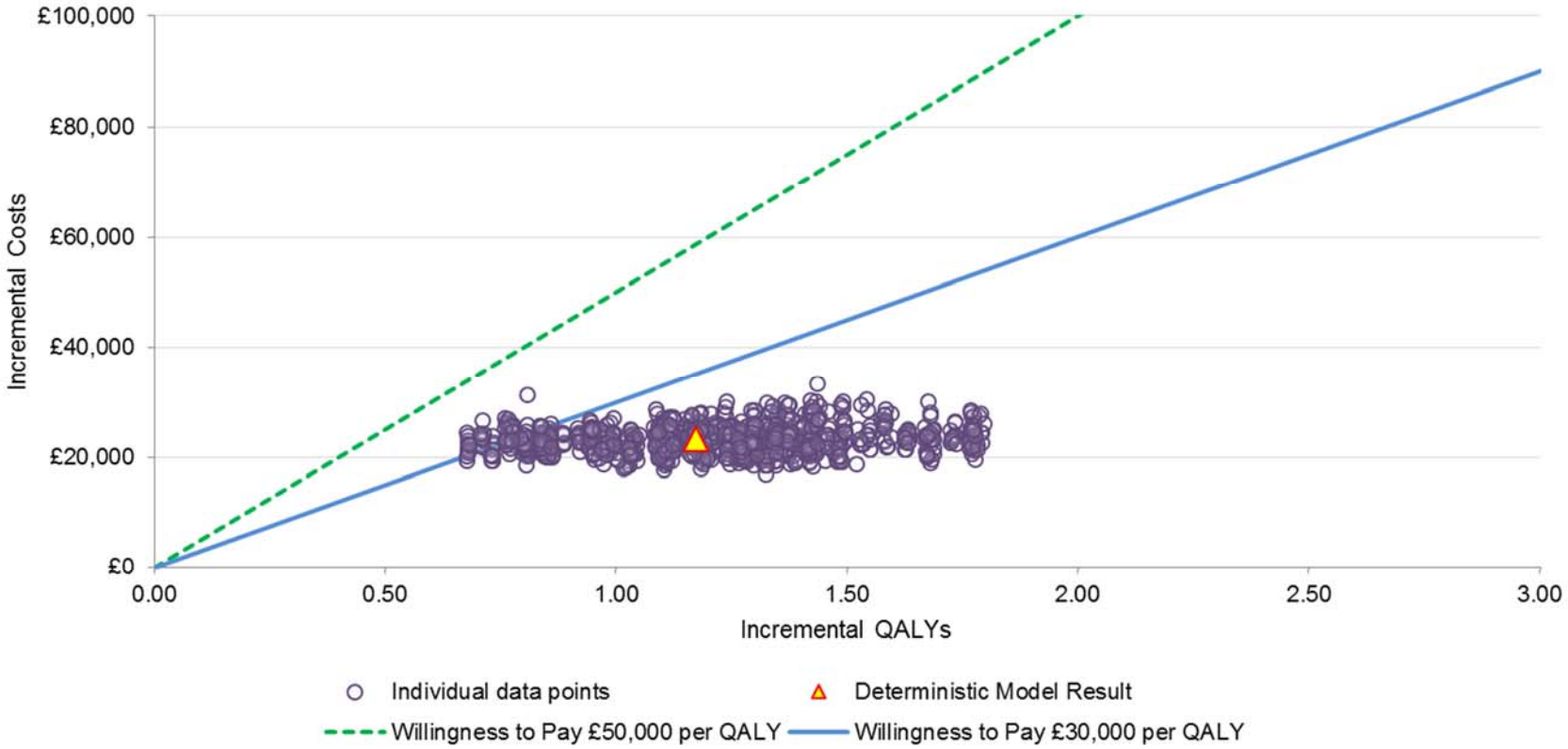
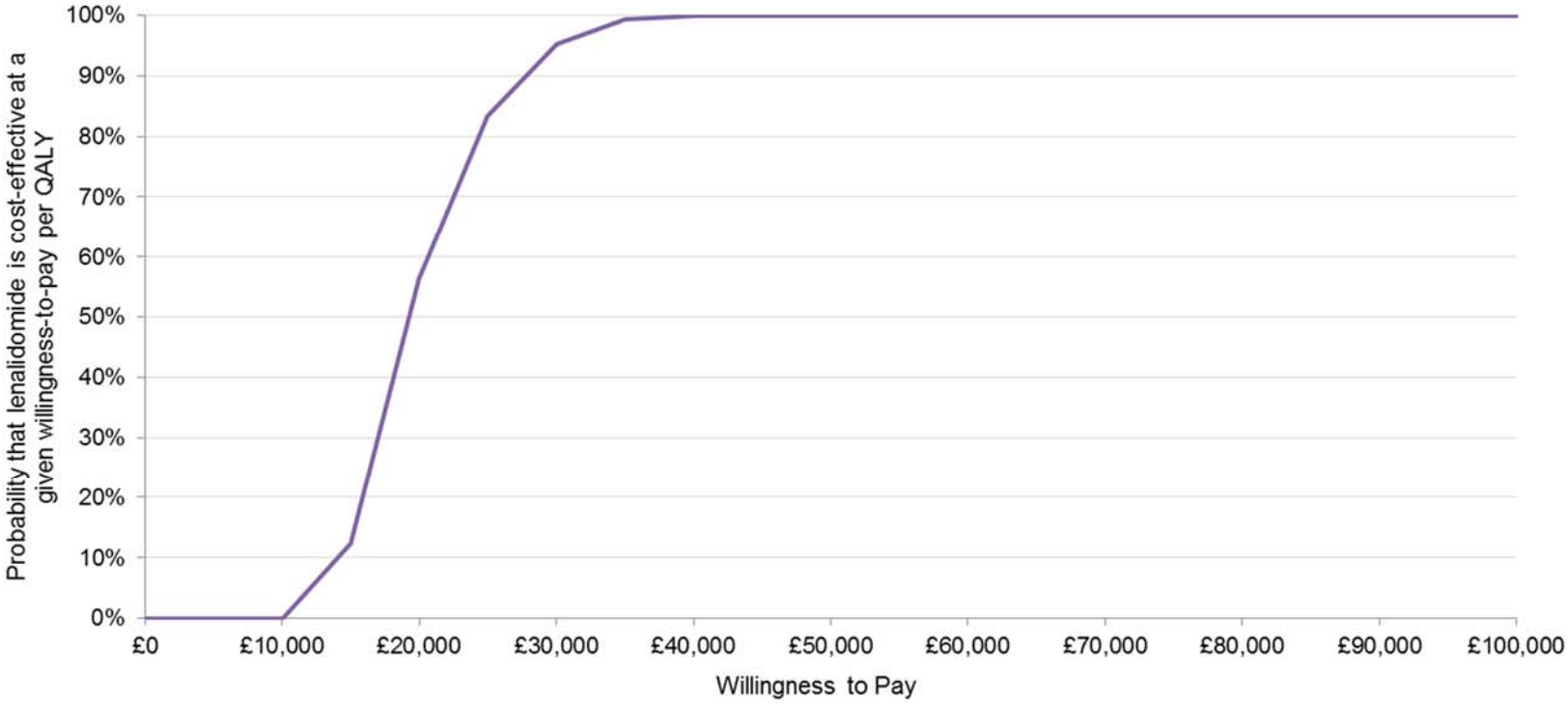


Figure 9: CEAC vs bortezomib



The probabilistic sensitivity analysis against bortezomib retreatment shows that the probability of lenalidomide being cost-effective at a willingness to pay threshold of £20,000 per QALY is 56.5% and at £30,000 per QALY it is 95.2%.

Figure 10: Scatter plot vs MP

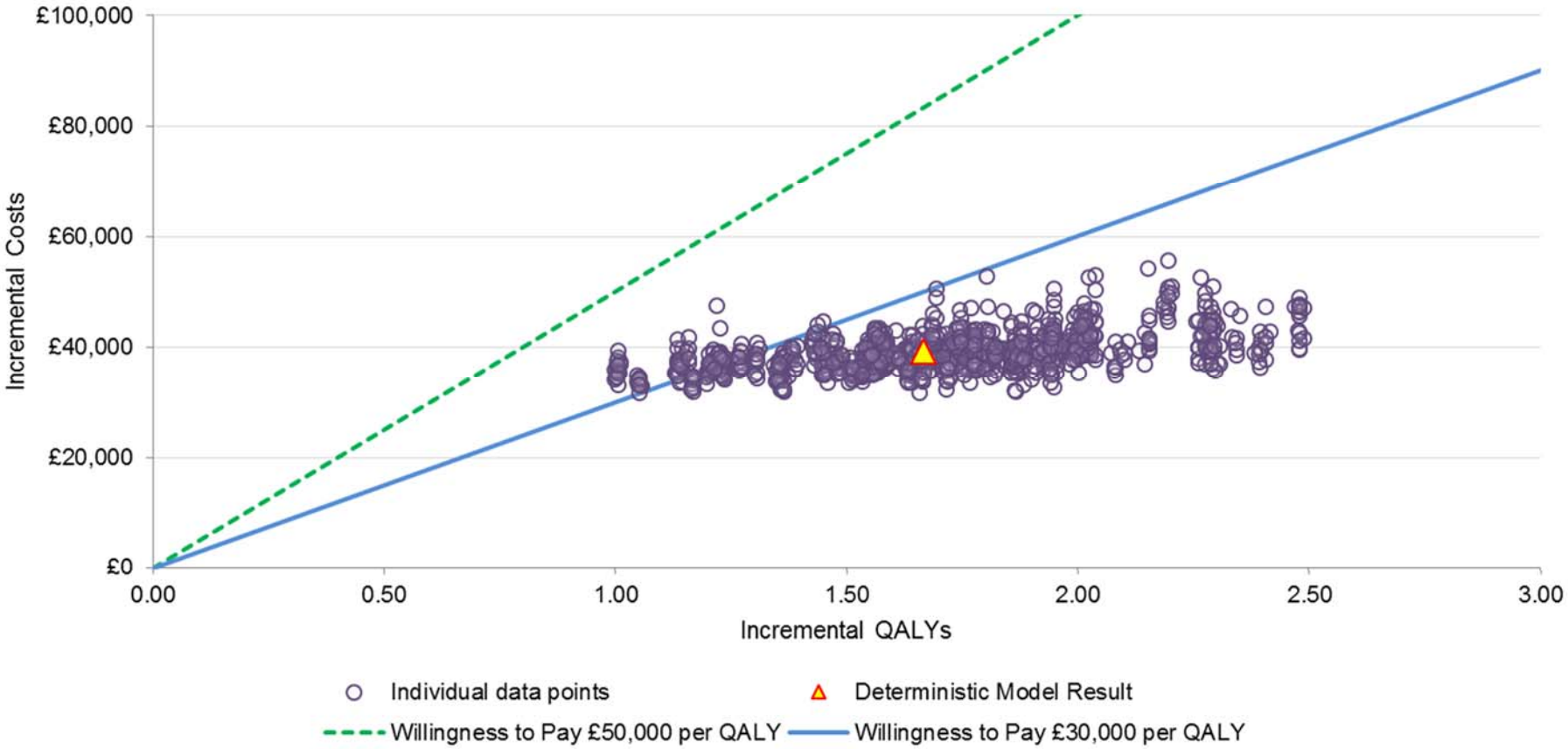
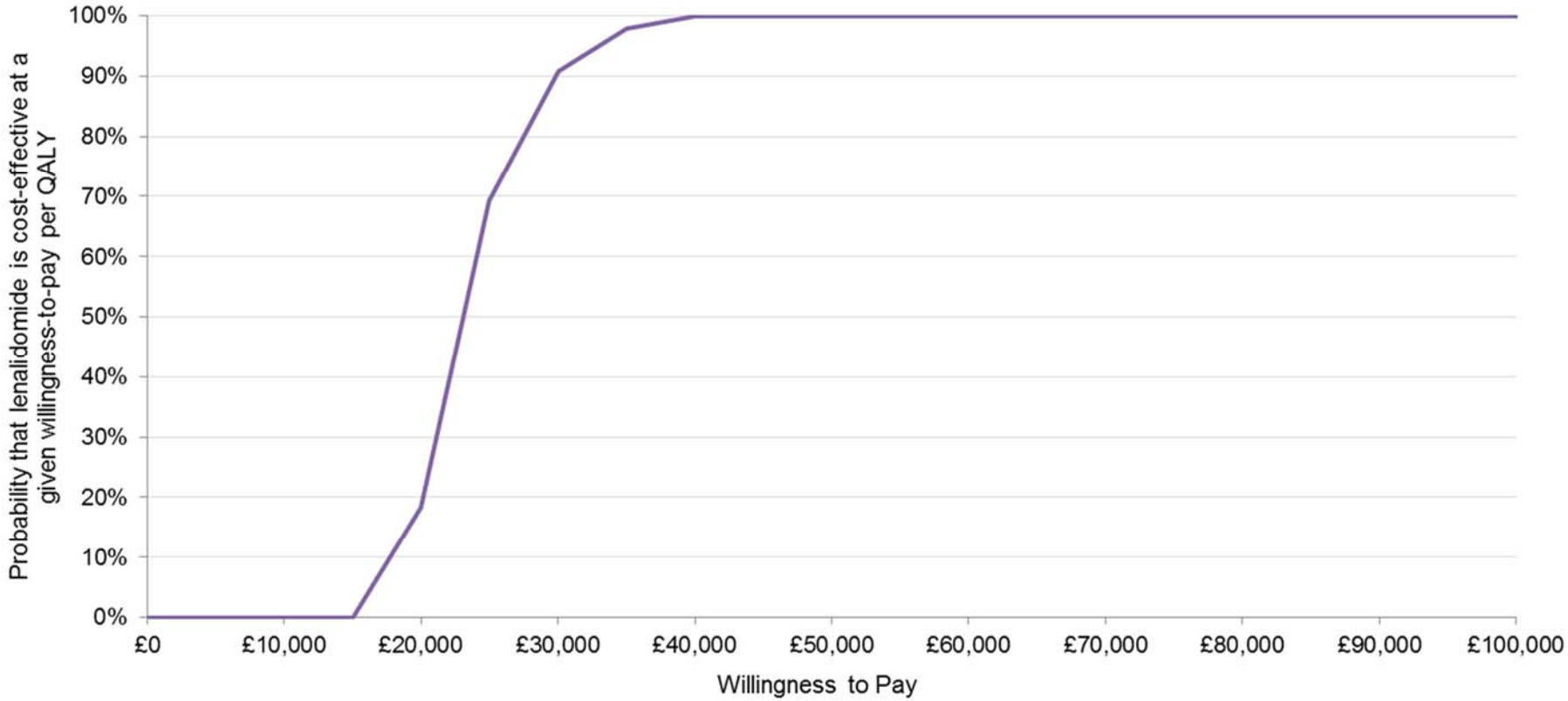


Figure 11: CEAC vs MP



The probabilistic sensitivity analysis against MP shows that the probability of lenalidomide being cost-effective at a willingness to pay threshold of £20,000 per QALY is 18.3% and at £30,000 per QALY it is 90.8%.

4.11 Please present scenario analysis results as described for the main manufacturer/sponsor submission of evidence for the technology appraisal.

Scenario Analyses

Taking the base case as scenario 1, the following additional scenarios were evaluated (please note, these differ to those presented originally in the manufacturer submission, but we believe are more informative to the Appraisal Committee at this stage of the review):

- Scenario 2 – MOC.
- Scenario 3 – CGP.
- Scenario 4 – MSM with uncapped bortezomib treatment duration.
- Scenario 5 – MSM using adjusted parametric survival curves.
- Scenario 6 – MSM using TTF based on HR to PFS curve.
- Scenario 7 – MSM with no 3rd line lenalidomide in comparator arm.

Scenario 2 – MOC

The results incorporating the PAS using the MOC method of covariate adjustment are presented in Table 11 below.

Table 11: Results using MOC

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£) versus MP (QALYs)	ICER (£) incremental (QALYs)
Melphalan and Prednisone	45,652	3.015	1.868	-	-	-	-	-
bortezomib retreatment	57,490	3.256	2.200	11,837	0.241	0.332	35,678	35,678
Lenalidomide plus dexamethasone	79,998	4.693	3.050	22,508	1.437	0.851	**29,041**	**26,452**

Celgene believe that the MSM analysis is methodologically more valid and helps reduce the uncertainty around the comparative efficacy and extrapolation

of survival data. This scenario is presented to allow an evaluation of the impact of the PAS in isolation of the methodological improvements.

Scenario 3 – CGP

The results incorporating the PAS using the MOC method of covariate adjustment are presented in Table 12 below.

Table 12: Results using CGP

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£) versus MP (QALYs)	ICER (£) incremental (QALYs)
Melphalan and Prednisone	45,390	3.032	1.869	-	-	-	-	-
bortezomib retreatment	55,695	3.417	2.312	10,305	0.385	0.443	23,238	23,238
Lenalidomide plus dexamethasone	80,460	4.826	3.108	24,765	1.409	0.796	**28,299**	**31,119**

Celgene believe that the MSM analysis is methodologically more valid and helps reduce the uncertainty around the comparative efficacy and extrapolation of survival data. The MSM does use a CGP methodology for covariate adjustment.

As with the MOC results, this scenario is presented to allow an evaluation of the impact of the PAS in isolation of the methodological improvements.

Scenario 4 – MSM with uncapped bortezomib treatment duration

Celgene agree that in UK clinical practice the majority of patients treated at second line with bortezomib will receive a maximum of 8 cycles. However, there will be some who are treated for longer and within the health economic model, the efficacy data for bortezomib based upon a dataset where some patients received over 8 cycles of treatment.

Table 13: Results using MSM with uncapped bortezomib treatment duration

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£) versus MP (QALYs)	ICER (£) incremental (QALYs)
Melphalan and Prednisone	42,610	3.153	1.880	-	-	-	-	-
bortezomib retreatment	106,690	3.685	2.374	64,080	0.533	0.494	129,597	129,597
Lenalidomide plus dexamethasone	81,887	5.867	3.546	-24,803	2.182	1.172	**23,572**	**Dominant**

Table 13 shows that for patients treated with bortezomib until progression, lenalidomide is highly cost-effective with the PAS applied.

Scenario 5 – MSM using adjusted parametric survival curves

A concern at the previous Appraisal Committee meetings was the use of median survival estimates alone to inform comparative effectiveness estimates. The use of medians to estimate comparative survival has two key limitations:

- Assume proportional hazards between treatments – this may not hold; particularly as all treatments being compared have quite different mechanisms of action.
- Application of a hazard derived from medians is technically only correct when used in conjunction with an exponential function to estimate survival, however, (as described above) the exponential function did not provide a good estimate of survival for the one available dataset (lenalidomide data for MM009/MM010)

The updated systematic literature review identified only two papers which which provide Kaplan Meier data for OS and/or PFS for bortezomib retreatment or MP:

- Petrucci 1989[16] provides Kaplan Meier data for OS for MP.
- Ahn 2014[3] provides Kaplan Meier data for OS and PFS for bortezomib retreatment.

In order to provide a direct comparison using Kaplan Meier data the following steps were followed:

- Digitisation of Kaplan Meier data and translation to simulated individual patient level data using the Guyot algorithm[17].
- Fitting of standard parametric functions to the Kaplan Meier data (see Appendix E: Digitalisation of curves; fitting for OS and PFS).
- Adjustment to derive outcomes for each parametric function for second-line patients based upon the published hazard ratio[18] of 1.181 for third-line plus patients versus second line patients:
    - Observed survival = Survival at second-line * % at second-line + Survival at second-line ^ 1.181 * (1 - % at second-line).

Comparative estimates using median survival were produced in the same manner as before by adjusting the lenalidomide survival curves produced using the MSM method to the relevant population within each clinical trial paper using the corrected group prognosis method to account for reported patient characteristics.

As no patient-level data were available, the following simplifying assumption had to be made: the likelihood of a patient presenting in one particular clinical group has no bearing on the likelihood that the patient falls within any other group. For example, the likelihood of having received one versus two or more prior therapies is independent of prior doxorubicin usage.

Results of this analysis show that the survival curves produced using hazard ratios produced from medians are generally similar / more optimistic in their projections of comparator outcomes than the survival curves produced using curves fit to Kaplan Meier data (Figure 12 and Figure 13).

Figure 12: Comparison of survival curves fit to digitised data versus hazard ratios produced by median estimates using Petrucci 1989

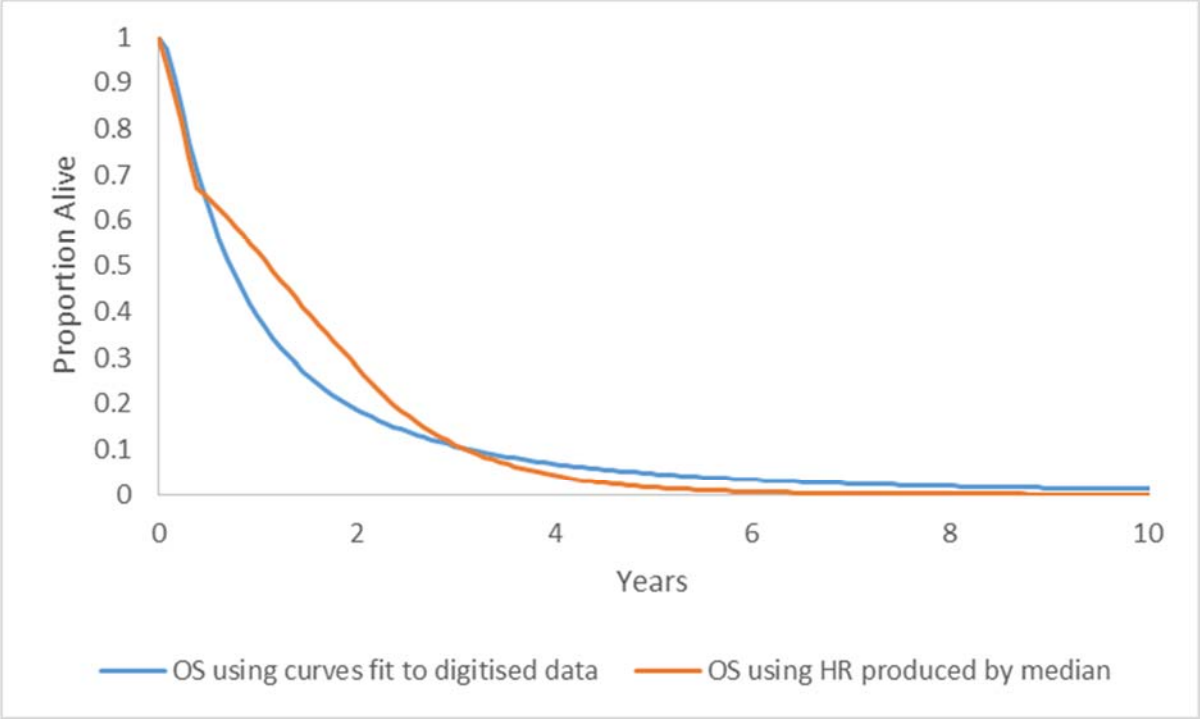


Figure 13: Comparison of survival curves fit to digitised data versus hazard ratios produced by median estimates using Ahn 2014

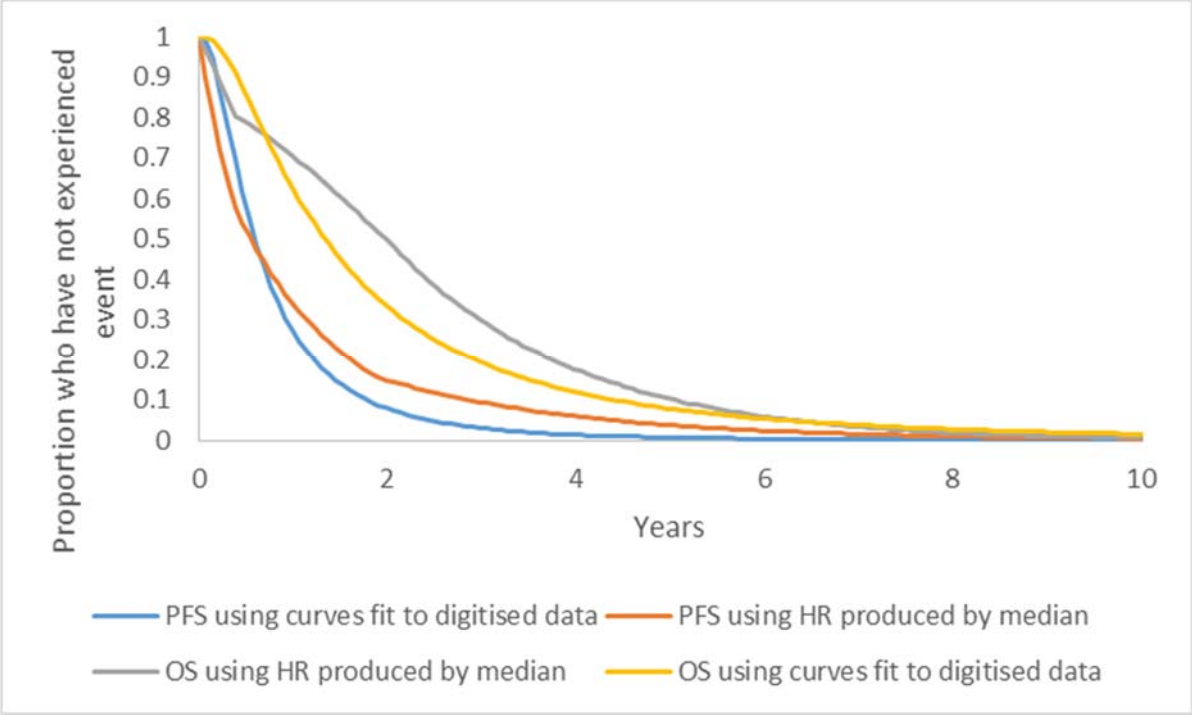


Table 14 shows the results when the digitized curves are used in the MSM model and not median values.

Table 14: Results using MSM with digitised curves not medians

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£) versus MP (QALYs)	ICER (£) incremental (QALYs)
Melphalan and Prednisone	42,499	3.147	1.878	-	-	-	-	-
bortezomib retreatment	55,999	3.250	2.020	13,500	0.103	0.142	95,386	95,386
Lenalidomide plus dexamethasone	81,887	5.867	3.546	25,889	2.617	1.526	**23,618**	**16,963**

Scenario 6 – MSM using TTF based on HR to PFS curve

To date, TTF is modelled using a gamma curve fit to patient level data for lenalidomide from MM009/MM010. No information was available for TTF (or time on treatment) from any of the published information for comparators. The same hazard ratio applied for PFS was therefore applied to TTF for comparative efficacy estimates.

- To explore assumptions around the link between TTF and PFS, a scenario analysis is presented assuming constant proportional hazards (PH) between TTF and PFS (Figure 14 demonstrates the similarity in the shape of the survival curves for the two measures). This scenario removes some of the previous uncertainty around survival curve fits for TTF.

Figure 14 Comparison of PFS and TTF in MM009/MM010

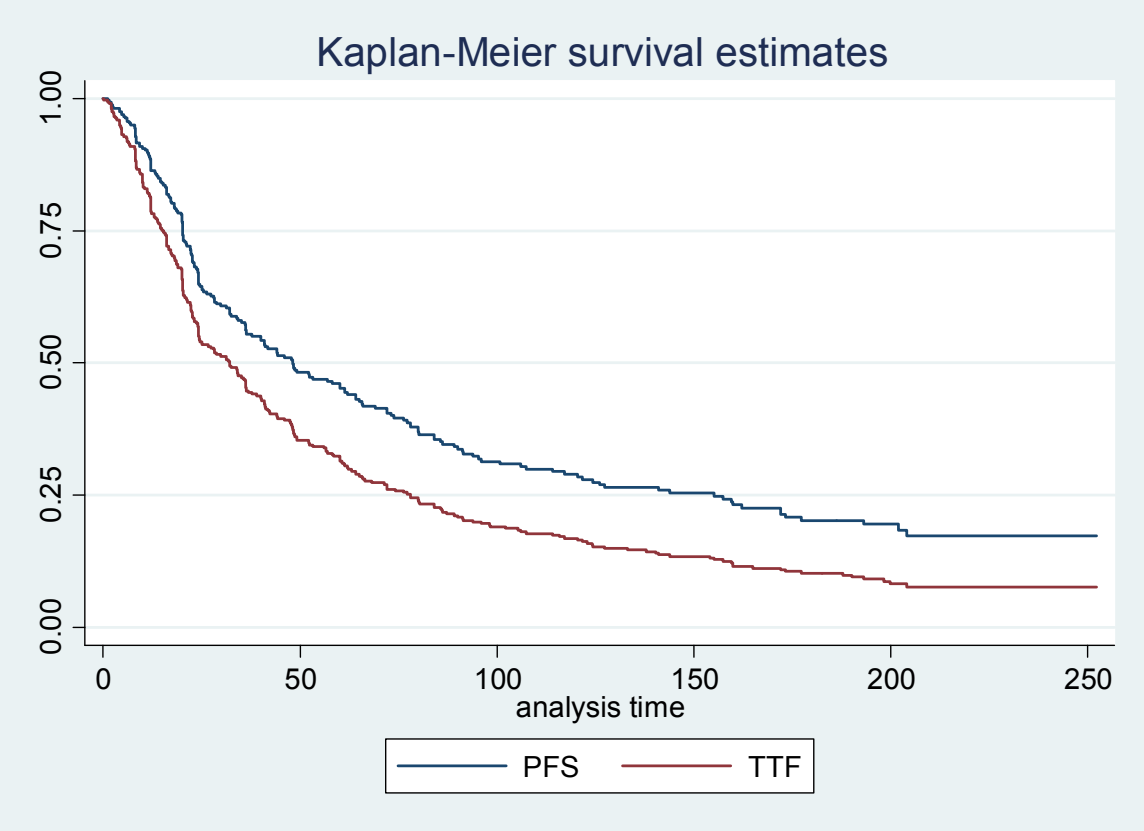


Table 15 shows the MSM results when PH is assumed between PFS and TTF.

Table 15: Results using MSM with TTF assuming PH to PFS

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£) versus MP (QALYs)	ICER (£) incremental (QALYs)
Melphalan and Prednisone	43,621	3.280	1.954	-	-	-	-	-
bortezomib retreatment	57,350	3.694	2.380	13,729	0.414	0.426	32,217	32,217
Lenalidomide plus dexamethasone	78,984	5.867	3.546	21,634	2.173	1.166	**22,211**	**18,554**

Scenario 7 – MSM with no 3rd line lenalidomide

Table 16 shows the MSM results when no 3rd line use of lenalidomide is assumed. Celgene do not believe that this is an appropriate reflection of clinical

practice in the UK. Since NICE TA171 was published, lenalidomide has been the standard of care at 3rd line, however it is not realistic to assume that patients progressing on lenalidomide at 2nd line would then go on to receive lenalidomide again at 3rd line.

When removing lenalidomide from 3rd line, a weighted average of subsequent therapies is used across both arms which is slightly different in composition and cost to the data specifically post LEN from the HMRN database. As such, the lenalidomide arm is also affected by this change and the results against each comparator must be presented in separate tables.

Table 16: Results vs bortezomib retreatment using MSM with no 3rd line lenalidomide

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£) versus baseline (QALYs)
bortezomib retreatment	37,005	2.701	1.862	-	-	-	-
Lenalidomide plus dexamethasone	81,748	5.867	3.546	44,743	3.166	1.684	**26,567**

Table 17: Results vs MP using MSM with no 3rd line lenalidomide

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£) versus baseline (QALYs)
Melphalan and Prednisone	18,941	1.142	0.782	-	-	-	-
Lenalidomide plus dexamethasone	84,568	5.867	3.546	65,627	4.725	2.764	**23,742**

4.12 If any of the criteria on which the patient access scheme depends are clinical variable (for example, choice of response measure, level of response, duration of treatment), sensitivity analyses around the individual criteria should be provided, so that the

Appraisal Committee can determine which criteria are the most appropriate to use.

N/A. The application of the PAS does not depend on clinical variables, only that the patient is receiving lenalidomide beyond 26 cycles.

1.1.3 Impact of patient access scheme on ICERs

4.13 For financially based schemes, please present the results showing the impact of the patient access scheme on the ICERs for the base-case and any scenario analyses. A suggested format is shown below (see table 5). If you are submitting the patient access scheme at the end of the appraisal process, you must include the scenario with the assumptions that the Appraisal Committee considered to be most plausible.

Table 18 - Results showing the impact of patient access scheme on ICERs

	Against comparator:			
	bortezomib		Melphalan and prednisone	
	Without PAS	With PAS	Without PAS	With PAS
Scenario 1 (base-case)	44,605	**19,781**	41,030	**23,572**
Scenario 2 – MOC	55,664	**26,452**	50,057	**29,041**
Scenario 3 - CGP	66,619	**31,119**	51,096	**28,299**
Scenario 4 – MSM with uncapped bortezomib treatment duration	3,658	**Dominant**	41,030	**23,572**
Scenario 5 – MSM using adjusted parametric survival curves	36,022	**16,963**	41,060	**23,618**
Scenario 6 – MSM using TTF based on HR to PFS curve	45,363	**18,554**	41,844	**22,211**
Scenario 7 – MSM with no 3rd line lenalidomide	43,840	**26,567**	34,265	**23,742**

Key: PAS, patient access scheme; MOC, Mean of Covariates; CGP, Corrected Group Prognosis; MSM, multi stage modelling; HR, Hazard Ratio; PFS, progression free survival

Conclusion

Celgene have taken on board the feedback received at the previous Appraisal Committee meetings and in the ERG's reports and ACD. We have made every effort to reduce the levels of uncertainty in the model and present the committee with sufficient scenario analyses to help inform their decision.

With the changes made to the base case bringing down the uncertainty and the PAS incorporated into the model, Celgene believe that lenalidomide has been demonstrated to be cost-effective and clinically important treatment option.

5 Appendices

5.1 Appendix A: Additional documents

5.1.1 If available, please include copies of patient access scheme agreement forms, patient registration forms, pharmacy claim forms/rebate forms, guides for pharmacists and physicians, patient information documents

Please see below the relevant forms:

Treatment Continuation Schem

Pharmacy Registration form.pdf

5.2 Appendix B: Details of outcome-based schemes

5.2.1 If you are submitting a proven value: price increase scheme, as defined in the PPRS, please provide the following information:

- the current price of the intervention
- the proposed higher price of the intervention, which will be supported by the collection of new evidence
- a suggested date for when NICE should consider the additional evidence.

N/A

5.2.2 If you are submitting an expected value: rebate scheme, as defined in the PPRS, please provide the following details:

- the current price of the intervention (the price that will be supported by the collection of new evidence)
- the planned lower price of the intervention in the event that the additional evidence does not support the current price
- a suggested date for when NICE should consider the additional evidence.

N/A

5.2.3 If you are submitting a risk-sharing scheme, as defined in the PPRS, please provide the following details:

- the current price of the intervention (the price that will be supported by the collection of new evidence)
- the proposed relationship between future price changes and the evidence to be collected.

N/A

5.2.4 For outcome-based schemes, as defined in the PPRS, please provide the full details of the new information (evidence) planned to be collected, who will collect it and who will carry the cost associated with this planned data collection. Details of the new information (evidence) may include:

- design of the new study
- patient population of the new study
- outcomes of the new study
- expected duration of data collection
- planned statistical analysis, definition of study groups and reporting (including uncertainty)
- expected results of the new study
- planned evidence synthesis/pooling of data (if applicable)
- expected results of the evidence synthesis/pooling of data (if applicable).

N/A

5.2.5 If you are submitting a risk-sharing scheme, please specify the period between the time points when the additional evidence will be considered.

N/A

5.2.6 Please provide the clinical effectiveness data resulting from the evidence synthesis and used in the economic modelling of the patient access scheme at the different time points when the additional evidence is to be considered.

N/A

5.2.7 Please provide the other data used in the economic modelling of the patient access scheme at the different time points when the

additional evidence is to be considered. These data could include cost/resource use, health-related quality of life and utilities.

N/A

5.2.8 Please present the cost-effectiveness results as follows.

- For proven value: price increase schemes, please summarise in separate tables:
  - the results based on current evidence and current price
  - the anticipated results based on the expected new evidence and the proposed higher price.
- For expected value: rebate schemes, please summarise in separate tables:
  - the results based on the expected new evidence and the current price (which will be supported by the additional evidence collection)
  - the results based on the current evidence and the lower price (if the new evidence is not forthcoming).
- For risk-sharing schemes, please summarise in separate tables:
  - the results based on current evidence and current price
  - the results based on the expected new evidence and the current price (which will be supported by the additional evidence collection)
  - the results based on the current evidence and the lower price (if the new evidence is not forthcoming)
  - the anticipated results based on the expected new evidence and the proposed higher price.

A suggested format is shown in table 3, section 4.7.

N/A

5.2.9 Please present in separate tables the incremental results for the different scenarios as described above in section 5.2.8 for the type of outcome-based scheme being submitted.

List the interventions and comparator(s) from least to most expensive. Present the incremental cost-effectiveness ratios (ICERs) in comparison with baseline (usually standard care), and the incremental analysis ranking technologies in terms of dominance and extended dominance. A suggested format is presented in table 4, section 4.8.

N/A

5.3 Appendix C: Systematic Review Methods

The Systematic Literature Review processes conformed to the specifications laid out in the Preferred Reporting Items for Systematic Reviews and Meta-Analyses report[19]. This SLR was performed in two parts: a comprehensive and systematic search of the published literature to identify all potentially relevant studies and a systematic selection of relevant studies based on explicit inclusion and exclusion criteria. There was a two stage review process (primary and secondary). At both primary and secondary screening level, abstracts were evaluated by one reviewer for inclusion/exclusion criteria. The decisions were then reviewed by a second independent reviewer.

5.3.1 Literature search

The review included searches of the following electronic databases from 15 August 2013 to 15 December 2015:

- MEDLINE and EMBASE (using EMBASE.com)
- MEDLINE In-Process
- CINAHL
- The Cochrane Library, including the following:
  - NHS Economic Evaluations Database (NHSEED)
  - Centre for Reviews and Dissemination (CRD)-Health Technology Assessment Database

In addition, conference proceedings from 2013 onwards were searched to identify recently completed or ongoing studies of interest. These included:

- American Society of Hematology (ASH)
- American Society of Clinical Oncology (ASCO)
- European Hematology Association (EHA)

Reference lists of previous systematic reviews/meta-analyses were hand-searched to ensure that all relevant studies were captured.

5.3.2 Study selection criteria

Potentially relevant publications were reviewed and assessed in order to collate a final set of studies, which formed the main body of clinical evidence. To determine the final set of studies eligible for the review, explicit inclusion and/or exclusion criteria were applied to the literature search results.

Table 19: Eligibility criteria used for study selection

Category	Inclusion criteria	Exclusion criteria
Population	Adult patients with rrMM with ≥1 prior treatment with bortezomib	• Newly diagnosed multiple myeloma or treatment-naïve patients • Studies that investigated both newly-diagnosed and rrMM, but did not segregate the results • Studies on children and other blood cancer • Studies in which patients had no prior bortezomib
Intervention	• Lenalidomide/dexamethasone • Lenalidomide based regimen • Bortezomib monotherapy • Bortezomib/high-dose dexamethasone • Bortezomib based regimen • Melphalan, vincristine, cyclophosphamide or doxorubicin based regimens • Bendamustine based regimen	• Lenalidomide monotherapy
Comparators	• Any other intervention	• Stem cell transplantation
Outcomes	• Progression-free survival (PFS) • Overall survival (OS) • Overall response rate (ORR) • Best response, including complete response (CR), partial response (PR), and very good partial response (VGPR). Minimal response (MR) if part of an ORR summation. • Time to next treatment (TNT) • Time to progression (TTP) • Adverse events (only Grade 3 and 4, or serious AEs) • Health related quality of life (HRQL)	Studies that did not report data on at least one of the outcomes of interest

Study design	• Randomised and non-randomised controlled trial of ≥ 5 patients • Prospective and retrospective observational studies of ≥ 5 patients • SR/MA of RCTs and non-RCTs	Letter, secondary analysis with no new/relevant data, expert opinions, commentaries, non-systematic reviews
Restriction	English language	Non-English language
Date	15 August 2013 onwards	

5.3.3 Patient population

This review focused on adult patients with relapsed/refractory multiple myeloma previously treated with bortezomib. Patients with newly diagnosed multiple myeloma or treatment-naïve patients were excluded. Studies were included at first screening stage if the line of therapy was unclear. However, at second screening stage, studies were excluded if the therapy line was unclear.

The studies assessing patients with both newly diagnosed and relapsed/refractory multiple myeloma were included if subgroup data for relapsed/refractory multiple myeloma was reported separately. Studies that reported adult and children patients were excluded if subgroup data for adult patients were not provided separately.

5.3.4 Study design

The study designs of interest to this review were randomised controlled trials, non-randomised controlled trials, prospective and retrospective observational studies. Single arm studies were also included.

5.3.5 Intervention

Studies were included if they evaluated at least one of the interventions mentioned in Table 19. Studies were included if at least one of the interventions mentioned in Table 19 was evaluated. There was no exclusion on the basis of comparator.

5.3.6 Date restriction

Studies published 15 August 2013 onwards were included.

5.4 Appendix D: Distribution of patients' covariate combinations and transition probability matrices

Table 20: Distribution of patients covariate combinations used in multi-state Markov models M-009 and MM-010

Group	Prior Therapies	Baseline beta-2 microglobulin	Prior doxorubicin	Worsening extramedullary plasmacytoma disease	Number of patients (%) N = 353
1	1	< 2.5 mg/L	Yes	Yes	1 (0.28)
2	2 or 3	<2.5 mg/L	Yes	Yes	0 (0.00)
3	1	>2.5 mg/L	Yes	Yes	3 (0.85)
4	2 or 3	>2.5 mg/L	Yes	Yes	3 (0.85)
5	1	<2.5 mg/L	No	Yes	0 (0.00)
6	2 or 3	<2.5 mg/L	No	Yes	0 (0.00)
7	1	>2.5 mg/L	No	Yes	3 (0.85)
8	2 or 3	>2.5 mg/L	No	Yes	2 (0.57)
9	1	<2.5 mg/L	Yes	No	25 (7.08)
10	2 or 3	<2.5 mg/L	Yes	No	38 (10.76)
11	1	>2.5 mg/L	Yes	No	31 (8.78)
12	2 or 3	>2.5 mg/L	Yes	No	94 (26.62)
13	1	<2.5 mg/L	No	No	19 (5.38)
14	2 or 3	<2.5 mg/L	No	No	20 (5.67)
15	1	>2.5 mg/L	No	No	42 (11.90)
16	2 or 3	>2.5 mg/L	No	No	72 (20.40)

Table 21: Transition probability matrices derived from multi-state Markov models - MM-009 and MM-010

		Time period of validity								
		0< - <=168			168 < - <=728			728 < - <=inf		
Group		Pre-prog.	Post-prog.	Death	Pre-prog	Post-prog.	Death	Pre-prog	Post-prog.	Death
1	Pre-prog.	0.90	0.10	0.01	0.92	0.07	0.00	0.96	0.04	0.00
	Post-prog.	0.00	0.95	0.05	0.00	0.98	0.02	0.00	0.97	0.03
	Death	0.00	0.00	1.00	0.00	0.00	1.00	0.00	0.00	1.00
2	Pre-prog.	0.86	0.13	0.01	0.90	0.10	0.00	0.94	0.05	0.00

	Post-prog.	0.00	0.92	0.08	0.00	0.97	0.03	0.00	0.96	0.04
	Death	0.00	0.00	1.00	0.00	0.00	1.00	0.00	0.00	1.00
3	Pre-prog.	0.83	0.14	0.03	0.88	0.11	0.00	0.94	0.06	0.00
	Post-prog.	0.00	0.91	0.09	0.00	0.97	0.03	0.00	0.96	0.04
	Death	0.00	0.00	1.00	0.00	0.00	1.00	0.00	0.00	1.00
4	Pre-prog.	0.79	0.19	0.02	0.84	0.15	0.00	0.91	0.08	0.00
	Post-prog.	0.00	0.87	0.13	0.00	0.95	0.05	0.00	0.94	0.06
	Death	0.00	0.00	1.00	0.00	0.00	1.00	0.00	0.00	1.00
5	Pre-prog.	0.92	0.08	0.01	0.94	0.06	0.00	0.97	0.03	0.00
	Post-prog.	0.00	0.92	0.08	0.00	0.97	0.03	0.00	0.96	0.04
	Death	0.00	0.00	1.00	0.00	0.00	1.00	0.00	0.00	1.00
6	Pre-prog.	0.89	0.10	0.01	0.92	0.08	0.00	0.96	0.04	0.00
	Post-prog.	0.00	0.88	0.12	0.00	0.95	0.05	0.00	0.94	0.06
	Death	0.00	0.00	1.00	0.00	0.00	1.00	0.00	0.00	1.00
7	Pre-prog.	0.86	0.11	0.02	0.91	0.09	0.00	0.95	0.05	0.00
	Post-prog.	0.00	0.87	0.13	0.00	0.95	0.05	0.00	0.94	0.06
	Death	0.00	0.00	1.00	0.00	0.00	1.00	0.00	0.00	1.00
8	Pre-prog.	0.83	0.15	0.02	0.87	0.12	0.01	0.93	0.06	0.00
	Post-prog.	0.00	0.81	0.19	0.00	0.92	0.08	0.00	0.90	0.10
	Death	0.00	0.00	1.00	0.00	0.00	1.00	0.00	0.00	1.00
9	Pre-prog.	0.96	0.04	0.00	0.97	0.03	0.00	0.99	0.01	0.00
	Post-prog.	0.00	0.97	0.03	0.00	0.99	0.01	0.00	0.99	0.01
	Death	0.00	0.00	1.00	0.00	0.00	1.00	0.00	0.00	1.00
10	Pre-prog.	0.95	0.05	0.00	0.96	0.04	0.00	0.98	0.02	0.00
	Post-prog.	0.00	0.95	0.05	0.00	0.98	0.02	0.00	0.98	0.02
	Death	0.00	0.00	1.00	0.00	0.00	1.00	0.00	0.00	1.00
11	Pre-prog.	0.92	0.05	0.02	0.96	0.04	0.00	0.98	0.02	0.00
	Post-prog.	0.00	0.95	0.05	0.00	0.98	0.02	0.00	0.98	0.02
	Death	0.00	0.00	1.00	0.00	0.00	1.00	0.00	0.00	1.00
12	Pre-prog.	0.91	0.07	0.01	0.94	0.06	0.00	0.97	0.03	0.00
	Post-prog.	0.00	0.92	0.08	0.00	0.97	0.03	0.00	0.96	0.04
	Death	0.00	0.00	1.00	0.00	0.00	1.00	0.00	0.00	1.00
13	Pre-prog.	0.97	0.03	0.00	0.98	0.02	0.00	0.99	0.01	0.00
	Post-prog.	0.00	0.95	0.05	0.00	0.98	0.02	0.00	0.98	0.02
	Death	0.00	0.00	1.00	0.00	0.00	1.00	0.00	0.00	1.00
14	Pre-prog.	0.96	0.04	0.00	0.97	0.03	0.00	0.98	0.02	0.00
	Post-prog.	0.00	0.93	0.07	0.00	0.97	0.03	0.00	0.97	0.03
	Death	0.00	0.00	1.00	0.00	0.00	1.00	0.00	0.00	1.00

15	Pre-prog.	0.94	0.04	0.02	0.97	0.03	0.00	0.98	0.02	0.00
	Post-prog.	0.00	0.92	0.08	0.00	0.97	0.03	0.00	0.96	0.04
	Death	0.00	0.00	1.00	0.00	0.00	1.00	0.00	0.00	1.00
16	Pre-prog.	0.93	0.06	0.01	0.95	0.05	0.00	0.97	0.02	0.00
	Post-prog.	0.00	0.89	0.11	0.00	0.95	0.05	0.00	0.94	0.06
	Death	0.00	0.00	1.00	0.00	0.00	1.00	0.00	0.00	1.00

Table 22: Forward selection of covariates

Step	Addition to Model	p-value
1	Prior_treatment	n/a
2	Baseline beta-2 microglobulin	9.21e-06
3	Worsening extramedullary plasmacytoma disease	0.018
4	Prior doxorubicin	0.034

5.5 Appendix E: Digitalisation of curves; fitting for OS and PFS

Table 23: Pettrucci 1989 – OS

Model	AIC	BIC
Log-normal	161.68	164.74
Log-logistic	163.00	166.05
Gompertz	165.27	168.32
Exponential	166.46	167.99
Weibull	168.42	171.47

Figure 15: Pettrucci 1989 – OS

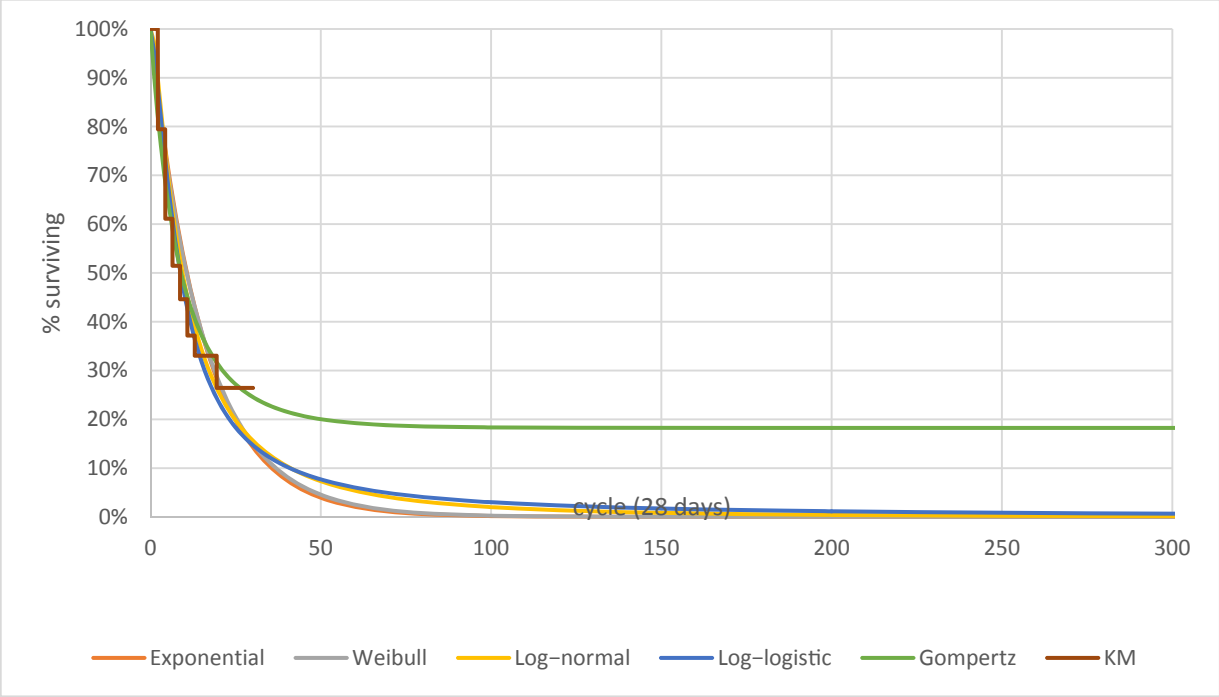


Table 24: Ahn 2014 – OS

Model	AIC	BIC
Generalised Gamma	167.58	167.58
Log-normal	169.23	172.03
Log-logistic	170.05	172.85
Exponential	173.89	175.29
Weibull	175.12	177.92
Gompertz	175.65	178.46

Figure 16: Ahn 2014 – OS

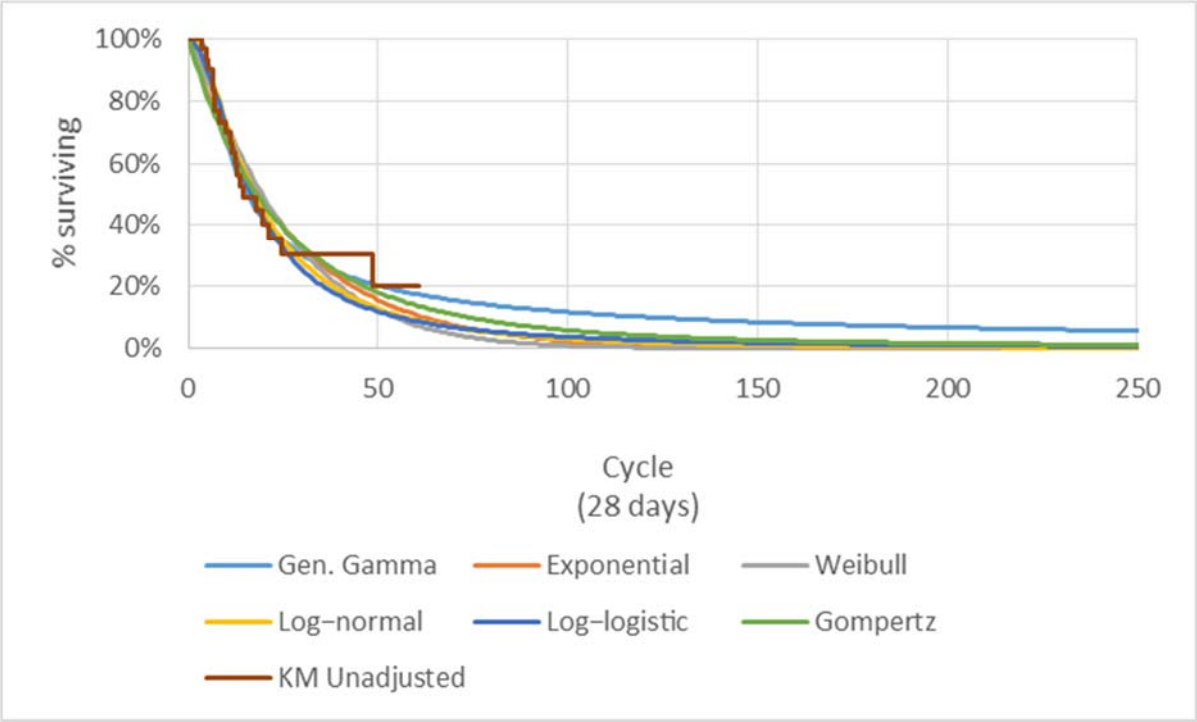
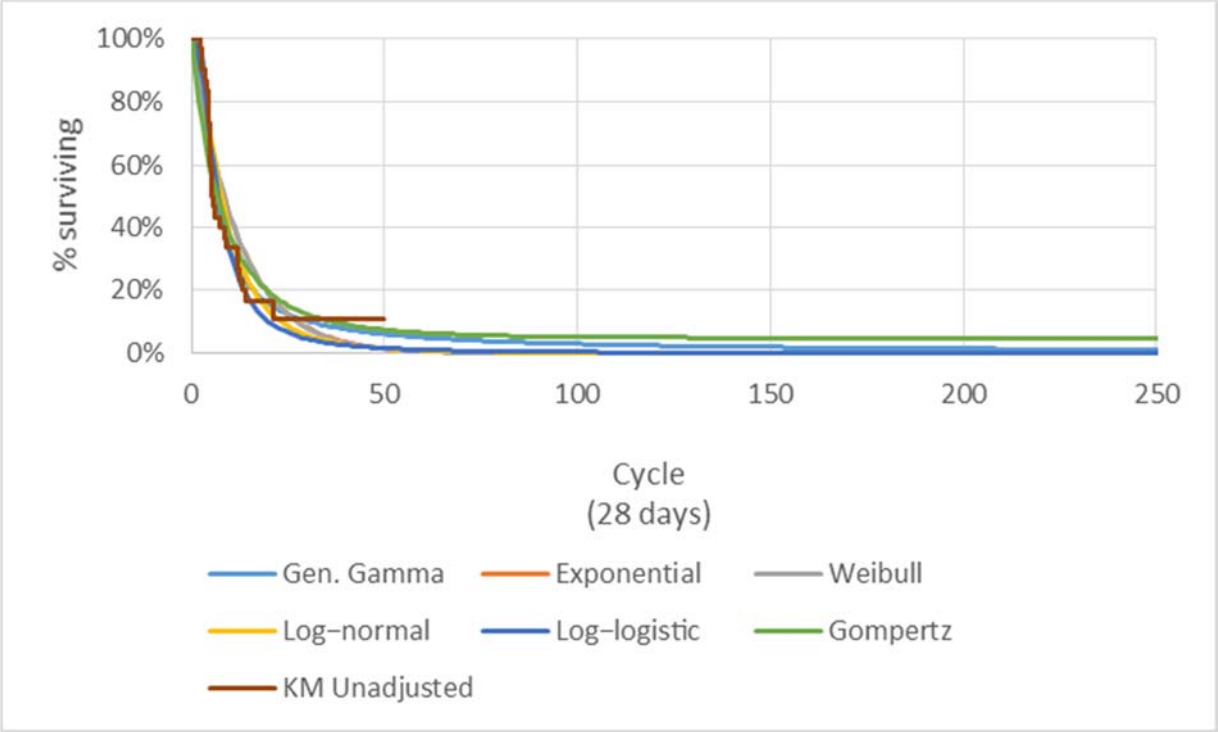


Table 25: Ahn 2014 – PFS

Model	AIC	BIC
Generalised Gamma	164.77	164.77
Log-logistic	171.43	174.24
Log-normal	173.01	175.82
Gompertz	181.53	184.33
Exponential	183.42	184.83
Weibull	185.39	188.19

Figure 17: Ahn 2014 – PFS



[1] NICE TA171. Lenalidomide for the treatment of multiple myeloma in people who have received at least one prior therapy. Available at: https://www.nice.org.uk/guidance/ta171/resources/lenalidomide-for-the-treatment-of-multiple-myeloma-in-people-who-have-received-at-least-one-prior-therapy-82598430636997, accessed on: 29/12/15

[2] NICE TA228. Bortezomib and thalidomide for the first-line treatment of multiple myeloma. Available at: https://www.nice.org.uk/guidance/ta228/resources/bortezomib-and-thalidomide-for-the-firstline-treatment-of-multiple-myeloma-82600316845765, accessed on: 29/12/15

[3] Ahn JS, Jung SH, Lee SS, et al. Clinical outcome of bortezomib retreatment in patients with relapsed or refractory multiple myeloma. Biomed Res Int. 2014: 145843.

[4] San-Miguel JF, Vania TM, Hungria, Yoon S-S, et al. Final analysis of overall survival from the Phase 3 Panorama 1 trial of panobinostat plus bortezomib and dexamethasone versus placebo plus bortezomib and dexamethasone in patients with relapsed or relapsed and refractory multiple myeloma. American Society of Hematology. Orlando, FL, USA. 5-8 December 2015. Abstract 3026.

[5] Hulin C, Rubia J, Dimopoulos M, et al. Bortezomib (BTZ) retreatment for relapsed/refractory (REL/REF) multiple myeloma (MM) in real-world medical practice: final data from the international, non-interventional, observational EVOBS study. Haematologica. Stokholm, Sweden. June 13-16 2013. B1497.

[6] Berenson J, Bitran J, Tabbara I, et al. A Retrospective Study to Evaluate the Incidence and Severity of Peripheral Neuropathy for Relapsed/Refractory Multiple Myeloma Patients upon Retreatment with Bortezomib. American Society of Hematology. San Francisco, CA, United States. December 6-9 2014. 3469.

[7] Kim HJ, Kwon J, Jung JY, et al. Metronomic chemotherapy of thalidomide, cyclophosphamide, and dexamethasone for relapsed/refractory multiple myeloma patients. Clinical Lymphoma Myeloma and Leukemia. 2015; 15:e311-e2.

[8] Mao J, Cheng F, Chen H, et al. Short-course bortezomib-based retreatment for patients with multiple myeloma who had received bortezomib-thalidomide-dexamethasone (VTD) as an initial therapy: A single-center case series. Exp Ther Med. 2014; 7(4):977-81.

[9] Oriol A, Giraldo P, Kotsianidis I, et al. Efficacy and safety of bortezomib-based retreatment at the first relapse in multiple myeloma patients: A retrospective study. Hematology. 2015; 20(7):405-9.

[10] Cerchione C, Madonna E, Pareto A, et al. Management of relapsed and refactory multiple myeloma with bendamustine. Haematologica. Milan, Itlay. June 12-15 2014. PB 1684.

[11] Reece D, Trieu Y, Masih-Khan E, et al. Treatment of relapsed/refractory multiple myeloma with weekly cyclophosphamide plus bortezomib plus prednisone or dexamethasone (CyBorP/D): Updated experience at Princess Margaret Cancer Centre (PMCC). American Society of Clinical Oncology. Chicago, IL, United States. May 30 - June 3 2014. e19568.

[12] Jagannath S, Roy A, Kish J, et al. Progression-Free Survival in Third-Line Treatment of Relapsed/ Refractory Multiple Myeloma Among Patients with Who Have Received Prior Bortezomib and Immunomodulatory Drugs. American Society of Hematology. Orlando, FL, United States. December 5-8 2015. 3311.

[13] Cerchione C, Madonna E, Marano L, et al. Bendamustine : Anew option for relapsed/refractory multiple myeloma. Haematologica. Stokholm, Sweden. June 13-16 2013. B1538.

[14] Kay R. A Markov model for analysing cancer markers and disease states in survival studies. Biometrics; 1986;42(4):855-65.

[15] Jackson C. Multi-state models for panel data: The msm package for R. Journal of Statistical Software; 2011;38(8).

[16] Petrucci T, Avvisati G, Tribalto M, et al. Intermediate-dose (25 mg/m2) intravenous melphalan for patients with multiple myeloma in relapse or refractory to standard treatment. Eur J Haematol 1989;42:233-237

[17] Guyot P, Ades AE, Ouwens JNM, et al. Enhanced secondary analysis of survival data: reconstructing the data from published Kaplan-Meier survival curves. BMC Medical Research Methodology 2012, 12:9

[18] Stadtmauer EA, Weber DM, Niesvizky R, et al. Lenalidomide in combination with dexamethasone at first relapse in comparison with its use as later salvage therapy in relapsed or refractory multiple myeloma. European Journal of Haematology. 2009; 82 (426–432)

[19] Moher D, Liberati A, Tetzlaff J, et al. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med. 2009;6(7): e1000097.

The clinical and cost-effectiveness of lenalidomide for people who have received at least one prior therapy with bortezomib (partial review of TA171)

A critique of the submission from Celgene

Addendum

Produced by

Peninsula Technology Assessment Group (PenTAG)
University of Exeter Medical School
South Cloisters
St Luke's Campus
Heavitree Road
Exeter
EX1 2LU

Authors

Nicola Huxley, Research Fellow, PenTAG, University of Exeter
Linda Long, Research Fellow, PenTAG, University of Exeter
Simon Briscoe, Information Specialist, PenTAG, University of Exeter
Martin Hoyle, Associate Professor, PenTAG, University of Exeter

Correspondence to

Nicola Huxley
Research Fellow
Peninsula Technology Assessment Group (PenTAG)
University of Exeter Medical School
South Cloisters
St Luke's Campus
Heavitree Road
Exeter
EX1 2LU
n.j.huxley@exeter.ac.uk

Date completed

07/03/2016

Source of funding

This report was commissioned by the NIHR HTA Programme as project number 13/07/01.

Declared competing interests of the authors

None

Acknowledgments We acknowledge the excellent administrative support of Sue Whiffin and Jenny Lowe (both of University of Exeter). We also acknowledge support and guidance given by Professor Chris Hyde (University of Exeter) relating to the critique of the company's clinical evidence. We also acknowledge the clinical advice provided by Claudius Rudin (Royal Devon and Exeter Hospital). We also acknowledge the support of Professor Rob Anderson (University of Exeter) who provided cover for the Director role.

Rider on responsibility for report The views expressed in this report are those of the authors and not necessarily those of the NIHR HTA Programme. Any errors are the responsibility of the authors.

This report should be referenced as follows: Huxley N., Long L., Briscoe S., Hoyle M. The clinical and cost-effectiveness of lenalidomide for people who have received at least one prior therapy with bortezomib (partial review of TA171): A Single Technology Appraisal (Addendum). Peninsula Technology Assessment Group (PenTAG), 2016.

Contributions of authors

Nicola Huxley	Provided project management, led the checking of economic model submitted by Celgene. Contributed to the writing and editing of the report.
Linda Long	Led the critique of the company's clinical effectiveness evidence and contributed to the writing and editing of the report.
Simon Briscoe	Critiqued the search strategies submitted by Celgene and contributed to the writing of the report.
Martin Hoyle	Contributed to the critique of Celgene's economic evaluation and contributed to the writing and editing of this report.

Contents

List of tables

List of figures

Abbreviations

AE	Adverse event
AFT	Accelerated failure time
ASH	American Society of Hematology
ASCT	Autologous stem cell transplantation
BOR	Bortezomib
CDF	Cancer Drugs Fund
CEAC	Cost-effectiveness acceptability curve
CI	Confidence interval
CIC	Commercial in confidence
CR	Complete response
CSR	Case study report
DEX	Dexamethasone
EBMT	European Society for Blood and Marrow Transplantation
ECOG	Eastern Cooperative Oncology Group
EMA	European Medicines Agency
EORTC	European Organisation for Research and Treatment of Cancer
ERG	Evidence review group
FDA	Food and Drug Administration
G-CSF	Granulocyte-colony stimulating factor
HR	Hazard ratio
HRQoL	Health related quality of life
ICER	Incremental cost effectiveness ratio
IDMC	Independent Data Monitoring Committee
IMiD	Immunomodulatory drug
ITT	Intention to treat
KM or K-M	Kaplan-Meier
LEN	Lenalidomide
LEN+DEX	Lenalidomide plus dexamethasone

MA	Meta-analysis
MR	Minimal response
MM	Multiple myeloma
MP	Melphalan plus prednisolone
nCR	Near complete response
NE	Not estimable
NMB	Net monetary benefit
ORR	Overall response rate
OS	Overall survival
PAS	Patient access scheme
PD	Progressive disease
PFS	Progression-free survival
PR	Partial response
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PSA	Probabilistic sensitivity analysis
QALY	Quality adjusted life year
QLQ	Quality of Life Questionnaire
rrMM	Relapsed or refractory multiple myeloma
SCT	Stem cell transplant
SD	Stable disease
SE	Standard error
SPC	Summary of product characteristics
SR	Systematic review
TTF	Time to treatment failure
TTP	Time to progression
VGPR	Very good partial response
WTP	Willingness to pay

1 Summary

This addendum critiques the additional information submitted by Celgene and received by us, PenTAG, on 2nd February 2016.

This critique should be viewed as an addendum to the evidence review group (ERG) critique submitted in 2014. We differentiate between the 2013-14 documents and the current 2016 submission as follows:

- We refer to documents by their submission year, and where necessary, submission month. We will primarily refer to the last submitted evidence from May 2014.
- For clarity, all mentions of "the ERG" refer to the work undertaken in 2013-14. The work from 2013-14 was prepared by the consultancy company Matrix in tandem with PenTAG, whereas the current critique has been prepared by PenTAG alone. PenTAG has remained responsible for the content of all ERG work submitted during this project.

1.1 Summary and critique of new clinical effectiveness evidence submitted by the company

Celgene present a reasonable update of the systematic review of the literature, although we note a lack of rigor in reporting of both systematic review methodology and the data extraction process which may have increased the possibility of bias. Celgene's searches were satisfactory and their inclusion criteria, while not fitting the Scope exactly, were appropriate. We have concluded that, despite ten studies being excluded on the basis of being non-English language, the company is unlikely to have missed any evidence.

The updated search identified 11 unique papers (both journal articles and conference papers) which included participants who had received bortezomib as their initial treatment. Seven of these papers provided evidence for bortezomib retreatment, two included bendamustine regimens, one included cyclophosphamide in combination with thalidomide and dexamethasone, and one included dexamethasone treatment and cyclophosphamide in combination with bortezomib and prednisone. The data extraction table contained typographical errors, both in data values and references, but we are satisfied that this did not affect the use of any relevant data in the economic model.

Two papers were identified providing overall survival Kaplan-Meier data (Ahn et al., 2014; San-Miguel et al., 2015). The Ahn et al. (2014) paper, a retrospective observational study investigating the clinical efficacy and safety of bortezomib retreatment in participants with relapsed or refractory multiple myeloma, additionally provided Kaplan-Meir data for progression-free survival. This was considered in a scenario analysis in the economic model. Again, we noted a number of typographical errors in the reporting of the Ahn et al. paper, but were satisfied that this did not affect the use of the correct data in the economic model. Data from the paper by San Miguel et al. was not included in the economic model due to a lack of information on patient characteristics in the relevant patient subgroup (those who received prior bortezomib).

We independently checked the remaining nine papers and were satisfied that none contained Kaplan-Meier data for overall survival and hence did not contain usable data for the economic model.

We remind the committee that all the clinical data underlying the economic analysis includes many patients that had previously been treated with thalidomide and many that had not previously been treated with BOR. Despite this, we note that the committee has previously concluded that the underlying data are generalisable to the population of interest in this HTA (patients previously treated with BOR, unsuited to thalidomide).

1.2 Summary of new cost-effectiveness evidence submitted by Celgene

We believe that prior to the current model submitted by Celgene, the most recent previous version of the model is the ERG-corrected version submitted after the appraisal consultation document (ACD) in 2014. Following the ACD, Celgene produced a new analysis, in which the ICERs changed significantly such that the lenalidomide (LEN) arm dominated bortezomib (BOR) and had an ICER of £55,000 per QALY versus melphalan plus prednisolone (MP). The ERG response, submitted to NICE in June 2014, highlighted several concerns with this model. Under the preferred analysis where no subsequent lines of treatment were included and the ERG corrections were made, the ICERs came to £11,000 per QALY for lenalidomide versus bortezomib and £44,000 for lenalidomide versus melphalan plus prednisolone. We believe this is the last version of the model that the Committee has seen. We refer to this as the "2014 model".

Celgene now provide results for the cost-effectiveness of lenalidomide compared to two comparators: bortezomib retreatment and melphalan plus prednisolone. Other comparators from the scope (namely bendamustine) are no longer considered, in line with the opinion of the appraisal committee in 2014. The changes that have had a substantial impact on the cost-effectiveness of LEN are:

- Incorporation of the PAS (where lenalidomide becomes free of charge after 26 cycles of use) for 2nd-line LEN.
- The duration of BOR treatment is now restricted to 8, 3-week cycles.
- Estimation of TTP PFS and OS for LEN is now based on multi-state modelling (MSM). PFS and OS for LEN are now substantially longer tailed than in the 2014 model.
- The estimation of the HR for PFS and OS for LEN vs. BOR and LEN vs. MP is now adjusted using the MSM method, whereas in the 2014 model, the mean of covariates method was used.
- In the comparator arms, LEN as assumed to be taken 3rd-line by all patients that survive to progression.

In Celgene's current submission the incremental cost-effectiveness ratio (ICER) for lenalidomide vs. bortezomib is £44,605 per QALY gained without the second line patient access scheme (PAS) for lenalidomide (£19,781 per QALY gained when second line lenalidomide PAS is applied). The incremental QALYs gained are 2.37 and the incremental costs are £52,268 (£23,179 with PAS). For lenalidomide versus melphalan plus

prednisolone the ICER is £41,030 per QALY gained (£23,572 with lenalidomide PAS), incremental QALYs gained are 1.88 and incremental costs are £68,366 (£39,277 with PAS).

1.3 Summary of PenTAG's critique of the new cost-effectiveness evidence

We are satisfied that the changes to the 2014 model that Celgene have declared are correctly implemented in the 2016 model.

However, we have the following serious criticisms of Celgene's current 2016 analysis:

- Celgene present TTF, PFS and OS data for LEN with maximum follow up of just 4.6 years, whereas the maximum follow-up now, March 2016, is approximately 12 years. Therefore, we encourage Celgene to use outcomes data for LEN that is far more mature than presented in their latest submission. If much more mature data were used, this would substantially reduce the uncertainty in the cost-effectiveness of LEN.
- For consistency with the underlying clinical evidence, we prefer to cost for a mean of 3.8 treatment cycles of BOR, as opposed to the 6.6 that Celgene currently assume.
- We are satisfied with the MSM adjustments to the HRs used to estimate PFS and OS for BOR and MP. However, in the 2016 model, Celgene now adjust the resulting OS for BOR and MP to yield greater estimates of OS. The difference in OS is very slight for BOR, but substantial for MP. Celgene did not make this adjustment in their 2014 model. Celgene do not discuss this important change in their recent document discussing the PAS. Furthermore, we can see no reason for this adjustment. Therefore, in our base case analysis, we remove the adjustment.
- We are concerned about the choice of treatment comparators. Current NICE guidance indicates that bortezomib monotherapy is the only treatment option for patients who have received one prior treatment (i.e. second line).[1] However, our clinical advisor indicates that retreatment with bortezomib is rarely used. They also believe that if thalidomide is not a feasible treatment as in the population identified in this HTA, then bendamustine would be their preferred treatment option. Chemotherapy based regimens would be used very rarely as second line treatment. This is a significant diversion from the original scope and from the understanding of the committee reported in the ACD. As such, we would recommend further discussion on the choice of appropriate comparators.
- In 2014, the ERG work indicated that it would be most appropriate to remove the costs of subsequent lines of treatment from the model, as the treatments used in the effectiveness studies were not reported and it was deemed important to ensure that the modelled costs and effectiveness were consistent. In their current model, Celgene assume that all patients in the comparator arms would receive 3rd-line LEN on progression. We argue that this is inconsistent with the clinical data, and it would be preferable to assume no costs for 3rd- and subsequent lines of treatment, as in the 2014 model.

The NICE appraisal committee in 2014 considered that the End of Life criteria did not apply in this assessment (ACD, p60). We believe this remains the case. Under Celgene's current base case, life expectancy is clearly greater than the 2 year threshold for BOR and MP. Under our base case, life expectancy for BOR remains clearly greater than 2 years;

therefore, the criteria do not apply for the comparison with BOR. Conversely, life expectancy for MP is now below the threshold, at 1.5 years. However, we consider that the criteria do not apply for the comparison with MP, as the clinical data is anything but robust.

1.4 Commentary on the robustness of evidence submitted by the company

1.4.1 Strengths and weaknesses

The company present a reasonable update of the systematic review of the literature, although we note a lack of rigor in reporting of both systematic review methodology and the data extraction process which may have increased the possibility of bias

No studies which included both lenalidomide and a relevant comparator arm were identified.

Celgene presented an update of their model from 2014, where they have addressed some of the important concerns highlighted in the ERG and Committee documents from 2014. They have implemented most of these changes correctly and no major wiring errors have been identified in the model.

Celgene present a scenario analysis using the study identified from their update review (Ahn et al., 2014), but give the results using the parametric survival curves, which they have not done for the base case study (Taverna et al., 2012). They have also not explained their reasons for preferring Taverna et al. (2012) as their base case.

We disagree with the following three assumptions made by Celgene. These strongly affect the estimated cost-effectiveness of LEN:

- Assumptions concerning subsequent lines of treatment.
- Treatment duration of bortezomib.
- Adjustment to OS for BOR and MP.

It is important to stress that all estimates of cost-effectiveness of LEN are highly uncertain. This is mostly because the clinical effectiveness data is not randomised and is of very poor quality for BOR and MP. Therefore, we recommend that all ICERs should be considered as informing cost-effectiveness only broadly

1.4.2 Areas of uncertainty

Important remaining uncertainties are:

- Choice of treatment comparators. As this STA began three years ago, in 2013, we feel it would be appropriate to assess whether the comparators identified then are still appropriate. Our clinical expert does not believe this is the case. Most notably it is our understanding that BOR retreatment may not occur in current practice. Instead, for the population of interest, multiple myeloma patients, post-bortezomib who are contraindicated to thalidomide, bendamustine may be the most likely treatment option in current practice.
- Quality of evidence for comparator clinical effectiveness. The MM-009 and MM-010 provide high quality evidence for the effectiveness of LEN vs. DEX. However, the evidence for the comparator effectiveness of LEN vs. BOR or LEN vs. MP is of very

low quality because it is not randomised, and because the studies of BOR and MP are very small.

- Overall survival for lenalidomide. We believe that much more mature data for OS for LEN should be available. If such data were available, this would substantially reduce the uncertainty in the extrapolation of lenalidomide survival.
- The utilities are highly uncertain.

1.5 Summary of exploratory and sensitivity analyses undertaken by PenTAG

Table 1 shows the impact of our additional analyses on the ICER. It is very important to appreciate that all ICERs in Table 1 are highly uncertain because:

1. The underlying clinical data is not randomised.
2. The quality of the clinical data used to inform PFS and OS for BOR and MP is extremely low.
3. Given that modelled TTF, PFS and OS for LEN are based on immature data, substantial extrapolation is required.
4. The utilities are highly uncertain.
5. The nature of subsequent treatments is uncertain.

Celgene currently assume that all patients take BOR intravenously. However, we believe a proportion will likely take BOR subcutaneously. If we were to model a proportion taking BOR subcutaneously, this would increase the ICER of LEN vs. BOR.

Table 1. Impact on the ICER of additional analyses undertaken by PenTAG

	LEN vs. BOR	LEN vs. MP
Celgene 2016 model	**£20,000**	**£24,000**
1: No 3rd- or 4th-line treatments	£36,000	£37,000
2: Reduce mean duration of BOR from 6.6 to 3.8 treatment cycles	£29,000	£24,000
3: OS for BOR and MP based on HR alone	£20,000	£19,000
1 & 2	£45,000	£37,000
1 & 3	£35,000	£26,000
2 & 3	£28,000	£19,000
PenTAG base case (1 & 2 & 3)	**£44,000**	**£26,000**

Key: BOR, bortezomib; HR, hazard ratio; ICER, incremental cost-effectiveness ration; LEN, lenalidomide; MP, melphalan plus prednisolone; OS, overall survival

2 Background

This addendum presents a critique of the most recent submission by Celgene as part of the NICE Single Technology Appraisal ID667 “Lenalidomide for the treatment of multiple myeloma in people who have received at least one prior therapy with bortezomib (partial review of TA171)”. This appraisal considers the use of lenalidomide for the treatment of multiple myeloma in people who have received at least one prior therapy with bortezomib and who are unable to receive thalidomide.

This submission was received by PenTAG 3nd February 2016 and is intended to provide work in addition to the original submission received November 2013 and the additional work received in 2014.

Celgene have provided an updated literature search, new modelling methods and the addition of a patient access scheme (PAS) for 2nd-line use of lenalidomide.

Here, we critique the most recent submission from Celgene. We do not provide a full critique of Celgene’s latest model in its entirety. Whilst we assume prior knowledge of this STA, on occasions, we summarise the history of this submission.

To differentiate between the last submission from Celgene and the current submission, we refer to them by their submission year, and where necessary, the submission month

For clarity, all mentions to “the ERG” refer to the work prepared by the consultancy company Matrix in tandem with PenTAG in 2013-14. The current critique has been prepared by PenTAG alone.

PenTAG has remained responsible for the content of all ERG work submitted during this project.

3 Clinical effectiveness

3.1 Critique of the methods of review(s)

Celgene state that the updated systematic review was conducted in order to identify additional evidence for multiple myeloma patients treated with the following comparator therapies which were identified as relevant during appraisal committee meetings:

- Bortezomib retreatment ± steroids
- Chemotherapy ± steroids (including melphalan and cyclophosphamide)
- Bendamustine ± steroids (see pp.11-12 Celgene submission)

The updated search identified 11 unique papers (both journal and conference papers)[2-12] published in English language between August 2013 and December 2015 and which included participants who had received bortezomib as their initial treatment according to NICE TA228.[13]

3.1.1 Searches

The clinical effectiveness search was updated in December 2015. Searches were date limited to August 2013, which is the date the searches were run for the original submission in 2014. The full search strategies were supplied on request (Celgene 2016, data on file). Celgene state that the search terms are consistent with those used in the original submission, which were deemed appropriate in the ERG report from 2014. The search terms in the current submission are not precisely the same as the original submission: some intervention terms have been removed, presumably because they are not within the scope of the current submission. However, the remaining search terms are consistent with those in the original submission.

The update search used an appropriate range of bibliographic databases, including:

- MEDLINE and EMBASE (via EMBASE.com);
- PubMed (detailed as MEDLINE-in-Process via PubMed in the submission);
- CINAHL (via EBSO.com);
- The Cochrane Library, including the CDSR, CENTRAL, HTA and DARE databases.

We do not have access to EMBASE.com and cannot verify the syntax is correct for the MEDLINE or EMBASE searches. The syntax is appropriate for the other database searches. The submission also states that NHS EED was searched, but this is not reported in the additional data we requested which details the search strategies for each database that was searched (Celgene 2016, data on file).

Conference proceedings from 2013 onwards were searched to identify recently completed and ongoing studies. There is no record of searching trials registries such as clinicaltrials.gov or the ISRCTN Registry for ongoing or recently completed studies. This is an omission, but not necessarily problematic, as an attempt was made to identify unpublished studies by searching conference proceedings. Reference lists of previous systematic reviews/meta-analyses were hand-searched.

3.1.2 Inclusion criteria

Celgene's inclusion criteria are given in Table 2, with an additional column added to the right of the table, taken from the Scope,[13] for reference and comparison.

Table 2. Scope of the literature review: PICOS criteria for study inclusion

Criteria	Celgene definition	NICE Scope definition
Population	Adult patients with rrMM with ≥ 1 prior treatment with bortezomib	Adults with multiple myeloma for whom thalidomide is contraindicated and whose disease has progressed after at least 1 prior treatment with bortezomib
Interventions / comparators	Lenalidomide/dexamethasone Lenalidomide based regimen Bortezomib monotherapy Bortezomib/high-dose dexamethasone Bortezomib based regimen Melphalan, vincristine, cyclophosphamide or doxorubicin based regimens Bendamustine based regimen Any other intervention	Intervention: Lenalidomide in combination with dexamethasone Comparators: Bortezomib monotherapy and bortezomib in combination with high dose dexamethasone Chemotherapy including regimens based on mephalan, vincristine, cyclophosphamide and doxorubicin Bendamustine
Outcomes	Progression-free survival (PFS) Overall survival (OS) Overall response rate (ORR) Best response, including complete response (CR), partial response (PR), and very good partial response (VGPR). Minimal response (MR) if part of an ORR summation. Time to next treatment (TNT) Time to progression (TTP) Adverse events (only Grade 3 and 4, or serious AEs) Health-related quality of life (HRQL)	Progression-free survival Overall survival Response rates Time to next treatment Adverse effects of treatment Health-related quality of life
Study design	Randomised and non-randomised controlled trial of ≥ 5 patients Prospective and retrospective observational studies of ≥ 5 patients SR/MA of RCTs and non-RCTs	
Restriction	English language	
Date	15 August 2013 onwards	

Key: rrMM, relapsed/refractory multiple myeloma; AEs, adverse events; SR, systematic review; MA, meta-analysis; RCT, randomised controlled trial

Source: Celgene submission, Table 19, pp 55-56 and NICE Scope[13]

Population

Celgene's population (adult patients with rrMM with ≥1 prior treatment with bortezomib) is broadly in agreement with the Scope that defines the population of interest as "*adults with multiple myeloma for whom thalidomide is contraindicated*".[13] We note that the only study identified in this updated review and included in the updated scenario analysis involves a heavily treated population group; 50% of participants had previously received autologous stem cell transplantation (ASCT) and most (93.3%) of the participants had received prior thalidomide combination therapy.[4]

3.1.2.1 Interventions / Comparators

The intervention/comparators for inclusion broadly match the Scope.[13]

We note that many of these comparators were listed as interventions in Celgene's eligibility criteria used for study selection (Celgene submission, Section 5.3.2, p.55, Table 19).

3.1.2.2 Interventions

The interventions in the report do not strictly match those in the NICE Scope. The NICE Scope specifies the intervention to be lenalidomide in combination with dexamethasone. Celgene include both lenalidomide/dexamethasone and lenalidomide based regimen in their intervention criteria. We note that Celgene additionally list bortezomib monotherapy; bortezomib/high-dose dexamethasone; bortezomib based regimen; melphalan, vincristine, cyclophosphamide or doxorubicin based regimens; and bendamustine based regimen as interventions in their eligibility criteria (Celgene submission, Section 5.3.2, Table 19, p.55).

3.1.2.3 Comparators

The comparators in the report do not strictly match those in the NICE Scope. The Scope specifies bortezomib monotherapy and bortezomib in combination with high dose dexamethasone; chemotherapy including regimens based on melphalan, vincristine, cyclophosphamide and doxorubicin; and bendamustine as comparators. Celgene specifies "*any other intervention*" (with the exception of stem cell transplantation) as a comparator.

We note that the updated review has identified two unique studies of bendamustine treatment.[9, 12] However, no bendamustine data was included in the current submission.

3.1.2.4 Outcomes

The outcomes in the company report broadly match those in the Scope. However, they have added a new outcome, time to progression (TTP), and have made the following amendments to response rates and adverse effects of treatment:

> *Best response, including complete response (CR), partial response (PR), and very good partial response (VGPR). Minimal response (MR) if part of an ORR summation.*
>
> *Adverse events (only Grade 3 and 4, or serious AEs)*

We agree with the 2014 ERG report which stated that these outcomes appeared reasonable to provide a sensible range of dimensions to assess the clinical effectiveness of lenalidomide.

3.1.2.5 Study design

The Scope did not restrict study design. However, the NICE reference case in the guide to the methods of technology appraisal 2013 (Chapter 5.2.3) recommends studies should be restricted to RCTs and when they are not available, non RCTs.[14] Studies in the company report included randomised and non-randomised controlled trials, prospective and retrospective observational studies. We are satisfied the study designs meet the reference case.

3.1.2.6 Other

Celgene applied an English language restriction to their systematic review. Ten studies were excluded on the basis of language (4 during primary screening and 6 during full text screening, Figure 1).

3.1.2.7 Study selection

Celgene's report explains the process used in study selection as follows:

"There was a two stage review process (primary and secondary). At both primary and secondary screening level, abstracts were evaluated by one reviewer for inclusion / exclusion criteria. The decisions were then reviewed by a second independent reviewer" (Celgene submission, Section 5.3, p.54).

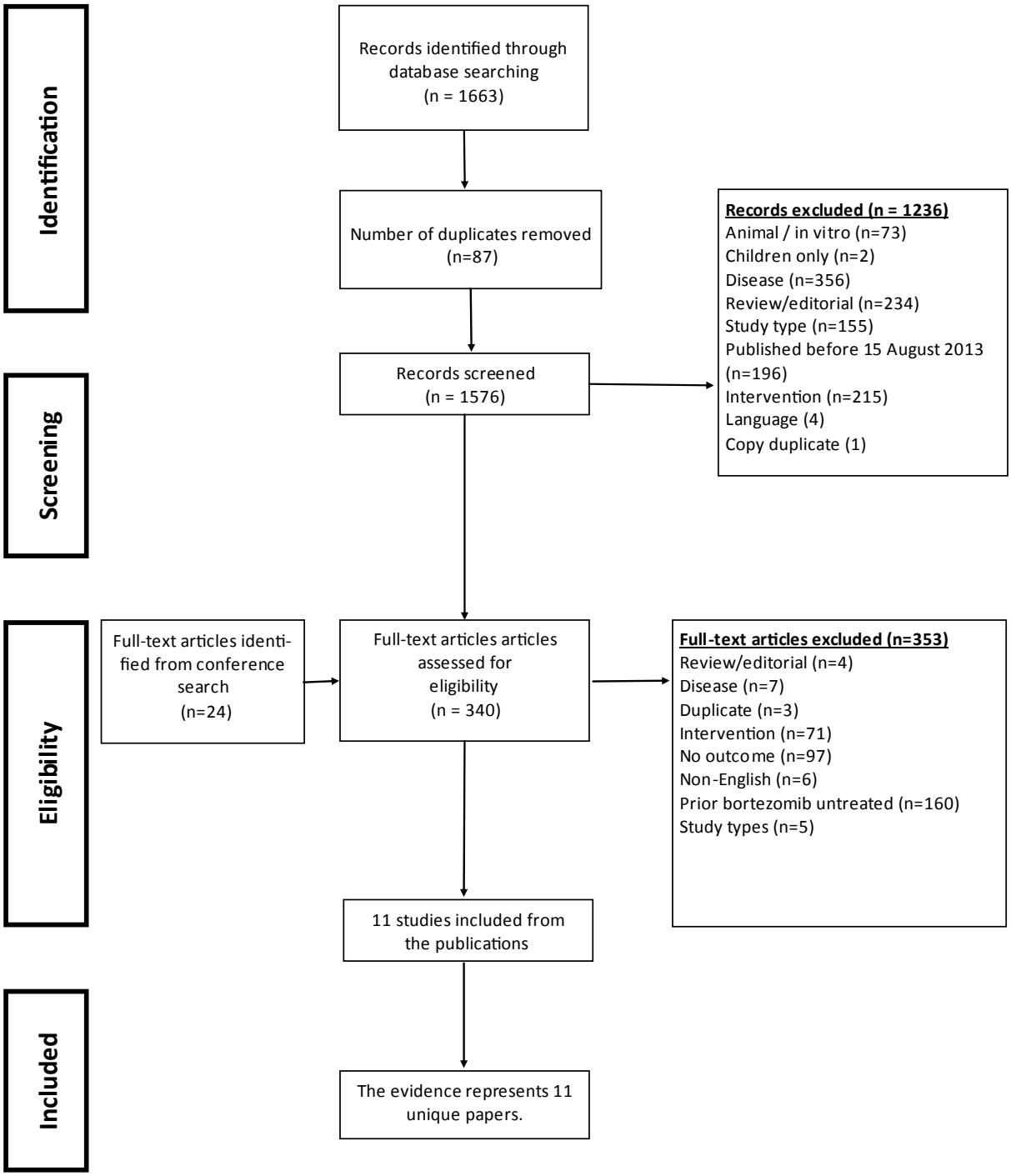
The lack of detail in Celgene's account of study selection introduces a level of uncertainty regarding the exact procedure followed. We propose that Celgene's methodological approach most likely deviates from standard procedures for systematic reviews where two researchers independently review the titles / abstracts and the full-texts of studies, and that discrepancies between investigators are resolved by involving a third investigator and coming to a consensus.[14] This is supported by our observation that no formal measure of agreement between two authors making simple inclusion / exclusion decisions (e.g. calculation of a kappa statistic) is reported. A lack of adherence to standard procedures, particularly the apparent lack of independent screening by two reviewers with a third person available to resolve discrepancies, could have introduced bias and increased the possibility that relevant studies were discarded.

From the 1,663 citations Celgene identified from their searches, 1,236 citations were excluded and 340 were taken to full-text screening at the abstract screening stage. The company identified 24 further full-text articles from conference search. From the full text screening, 353 citations were excluded and brief reasons for exclusion provided in the PRISMA diagram reported in Celgene's report (Figure 1, p.20).

3.1.3 Data extraction

Eleven papers involving relevant comparator therapies of potential interest were identified (Table 3, p.21)

Figure 1. Preferred Reporting Items for Systematic Reviews and Meta-Analyses diagram



Source: Celgene submission, Section 4.2, Figure 2, p.14

Table 3. Data extraction

Publication details	San-Miguel 2015[2]	Hulin 2013[3]	Ahn 2014[4]	Berenson 2014[5]	Kim 2015[6]	Mao 2014[7]	Orio 2014[8a]	Cerchione 2014[9]	Reece 2014[10]	Jagannath 2015[11]	Cerchione 2013[12]
Study type	RCT	POS	ROS	ROS	ROS	ROS	ROS	ROS	ROS	ROS	ROS
Treatment regimen of interest	BOR	BOR	BOR	BOR	BOR	CTD	BOR	BOR	BEND	CyBorP/D	BOR
Geographical region	Worldwide	Worldwide	Korea	NR	Korea	China	EU	NR	NR	NR	NR
N (on arm of interest)	51[b]	96	30	58	6	20	35	24	98	69	16
% previous use of DOX	NR	NR	NR	NR	NR	100	11	NR	NR	NR	NR
Lines of previous therapy (median or mean no. / % with 1 prior therapy)	NR	NR	2-57%, 3-20%, 4-13%, 5-10%	2(2-10)	75% (1 prior)	4 (2-11)	100% (1 prior)	6.3 (4-8)	2 (1-6)	2	5.7 (4-8)
ISS stage	NR	11-19	I -15/25 II -5/25, III – 6/25	NR	NR	NR	NR	ISS was equally distributed	NR	LII – 46.9%	ISS was equally distributed
Beta 2M	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
% worsening extramedullary disease	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
Age (years)	NR	62 (34-80)	67 (51-81)	64 (31-84)	75.1 (66.7 - 78.8)	63 (39 -72)	63 (34 - 84)	66 (48 - 83)	64 (36 -88)	68	62.6 (39 - 82)
Time from diagnosis (months)	NR	NR	43.6 (16.9 - 249.6)	NR	NR	NR	0.8 (0 - 3)	NR	NR	NR	NR
% receiving concomitant steroids	NR	41	76.7	NR	100	100	51	8%	100	NR	NR
% receiving other relevant concomitant therapies (chemotherapy)	NR	NR	100	NR	0	100	81	100	NR	NR	100
Previous SCT (%)	NR	NR	50	NR	0	5	15	58	75	46.8	69
Cycles of therapy received	NR	4	6 (2 - 12)	NR	15 (4 - 23)	2 (2 - 4)	4 (3 - 13)	4.3 (2 - 9)	5 (-47)	NR	4.7 (2 - 6)

Publication details	San-Miguel 2015[2]	Hulin 2013[3]	Ahn 2014[4]	Berenson 2014[5]	Kim 2015[6]	Mao 2014[7]	Orio 2014[8a]	Cerchione 2014[9]	Reece 2014[10]	Jagannath 2015[11]	Cerchione 2013[12]
Response rate (%)	NR	NR	73.3	NR	16.7	50	NR	62.5	68	NR	68
Median OS (95% CI)	19.5 (14.1 - 32.5)	17.6 (14.4 - 23.5)	13.4 (6.1 - 20.7)	NR	Not reached	NR	NR	6.7 (2 - 19)	24.2 (20.2 - 28.1)	NR	3.6 (2 - 6)
OS KM included (y/n)	Y	N	Y	NR	N	N	N	N	N	N	N
Median TTP (95% CI)	NR	NR	5.8 (2.6 - 9)	NR	NR	NR	NR	NR	NR	NR	NR
TTP KM included (y/n)	N	N	Y	N	N	N	N	N	N	N	N
Median PFS (95% CI)	N	6.9 (4.6 - 8.2)	5.8 (4.2 - 6.8)	NR	Not reached	20 (6 - 32)	15 (1-40)	NR	14.9 (12 - 17.9)	7.3 (4.2-10.8)	NR
TTP / PFS KM included (y/n)	N	N	Y	NR	N	N	Y	N	N	N	N
Median (or mean) time on treatment (95% CI or SD)	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
Time on treatment KM included (y/n)	N	N	N	N	N	N	N	N	N	N	N

Key: BOR, bortezomib; BEND, bendamustine; CI confidence interval; CTD, cyclophosphomide, thalidomide, dexamethasone; CyBorP, cyclophosphamide, bortezomib, prednisone; D, dexamethasone; DOX, bortezomib-dexamethasone-doxorubicin; EU, Europe; ISS, International Staging System; KM, Kaplan-Meier; N, no; NR, not reported; OS overall survival; PFS, progression free survival; POS, prospective observational study; RCT, randomised control trial; ROS, retrospective observational study; SCT, stem-cell transplant; SD, standard deviation; TTP, time to progression; Y, yes.

Notes: a Orio 2014 is reported in Celgene's data extraction table but is referenced as Oriol et al. (2015)[8] in the submission's reference section. We have unsuccessfully attempted to locate Orio 2014 and so have concluded that this is a typographical error; b Sub-population who had been previously treated with bortezomib and an immunomodulatory drug (IMiD); outcomes were not reported separately for participants only previously treated with bortezomib.

Source: Celgene submission, Section 4.2, Table 1, p.15

The report explains the process of data extraction used as follows:

"These papers were reviewed to determine whether the information contained within the papers would add to the weight of evidence available for both comparators based upon the following criteria:

- *Provision of Kaplan-Meier data for overall survival*
- *Provision of Kaplan-Meier data for progression-free survival*
- *Sample size"* (Celgene submission, Section 4.2, p.12)

This data extraction process undertaken by Celgene is not as rigorous as the standard review process which usually involves two investigators independently extracting data on study characteristics, interventions, patient characteristics at baseline, and outcomes for the study populations of interest. Any discrepancies found between the data extracted by the two data extractors is usually resolved by involving a third reviewer and coming to a consensus. In the current submission, Celgene do not state if the data extraction has been carried out by more than one author and no formal measure of agreement between two authors making simple inclusion / exclusion decisions (e.g. calculation of a kappa statistic) has been reported, suggesting that this process was not followed. Without evidence to the contrary, we conclude that it is likely that the data has been extracted by a single author.

We note that in relation to the 11 unique papers, as well as typical data (such as study type, geographical region and age) being extracted from the paper, additional data including Kaplan-Meier data for overall survival (OS) and Kaplan-Meier data for progression-free survival (PFS) were also extracted. We found at least one error in the placement of extracted data reported in the data extraction table of their submission (Celgene submission, Section 4.2, Table 1, p. 15). We observe that the median OS reported in the data extraction table for Hulin et al. (2013) is 17.6 (14.4 - 23.5) and assume that an error was made in reading the table and the wrong value reported.[3] Although such mistakes may indicate a lack of diligence on the part of the company, we are satisfied that these reporting errors are typographical mistakes and that the correct values have been used for any data included in the economic analysis.

3.1.4 Quality assessment

No quality assessment was reported.

3.1.5 Evidence synthesis

One study (Ahn et al., 2014)[4] was identified as having relevant data for the economic analysis and so synthesis of the evidence was not required. Two papers were identified providing overall survival Kaplan-Meier data.[2, 4] The Ahn et al. (2014) paper additionally provided Kaplan-Meier data for PFS and was therefore selected for inclusion in the economic model. We agree with Celgene that the paper by San Miguel et al. should not be included in the economic model "*due to a lack of information on patient characteristics in the relevant patient subgroup (those who received prior bortezomib)*" (Celgene submission, Section 4.2, p.12). Additionally all participants had received prior lenalidomide.

3.2 Critique of trials of the technology of interest, their analysis and interpretation (and any standard meta-analyses of these)

A single study of interest for updating the economic model was identified in the submission (main publication by Ahn et al., 2014).[4]

3.2.1 Study design and treatment

Ahn et al. (2014) was a retrospective observational study investigating the clinical efficacy and safety of bortezomib retreatment in patients with relapsed or refractory multiple myeloma (MM).[4] A total of 30 Korean participants who relapsed or progressed after ≥6 months since the last dose of their previous bortezomib therapy were included in this study and received retreatment with bortezomib in combination with dexamethasone and cyclophosphamide and thalidomide for a median of 6 (2 - 12) cycles. 20 participants received this retreatment, while 10 participants received the same regimen with the exclusion of thalidomide. All participants had received prior bortezomib-based salvage therapy, with 57%, 20%, 13% and 10% receiving 2, 3, 4 and 5 prior lines of therapy respectively. In addition to this, 50% of participants had also received stem cell transplantation (SCT). In the original paper, the median age of participants was reported as 67 years (range: 51 - 58). [4] We note a discrepancy in the submission with Celgene reporting the median age of participants as 62 years (range 51 - 81) (Celgene submission, Section 4.2, p.13), However, we note that the correct value (67 years) was reported by Celgene in the submission data extraction table (Table 3). The median time from diagnosis was 43.6 (16.9 – 249.6) months. Median time to progression (TTP) was 5.8 (2.6 - 9) months. Median PFS was 5.8 (4.2 - 6.8) months, and median OS was 13.4 (6.1 - 20.7) months. Ahn et al. note that there was no significant difference in PFS or OS according to the combination therapy regimen with bortezomib retreatment (i.e. either with or without thalidomide).[4] We note a discrepancy in the submission with Celgene reporting the median OS as 17.6 (14.4 - 23.5) months (Celgene submission, Section 4.2, p.13), although we also note that the correct value (13.4 months) was reported by Celgene in the submission data extraction table (Table 3).

The study was conducted in Korea. In the submission Celgene state that Ahn et al. observe that though the data represented a Korean cohort, OS and PFS estimates were consistent with Caucasian data (Celgene submission, section 4.2, p. 13). We note that this is a misrepresentation by Celgene, as in their original publication, Ahn et al. observed that TTP and OS estimates were consistent with Caucasian data.[4]

3.3 Critique of other evidence sources

Celgene provided ten further sources (eight retrospective observational studies, one prospective observational study and one RCT) of evidence within their current submission.

3.3.1 Retrospective observational evidence

The eight retrospective observational studies relevant to the decision problem identified by Celgene were by Berenson et al.,[5] Kim et al.,[6] Mao et al.,[7] Oriol et al.,[8] Cerchione et al.,[9, 12] Reece et al.[10] and Jagannath et al.[11] Celgene extracted data from all eight studies but did not use any of the data in the economic model, stating that none of the papers provided anything other than median survival to allow comparison to bortezomib. We have independently checked all eight studies and agree with Celgene that none of the papers

provided Kaplan-Meier data either for overall survival or progression-free survival suitable for use in the economic model.

3.3.2 Prospective observational evidence

The single prospective observational study relevant to the decision problem identified by Celgene was by Hulin et al. (2013).[3] Celgene extracted data from this paper but did not use any of this data in the economic model. We have independently checked this study and agree with Celgene that it does not provide Kaplan-Meier (K-M) data for overall survival (or progression-free survival) suitable for use in the economic model.

3.3.3 RCT evidence

The single RCT study relevant to the decision problem identified by Celgene was by San-Miguel et al.[2] The abstract by San-Miguel et al. (2015)[2] included OS K-M data for the prior treatment with bortezomib subgroup, but did not provide PFS K-M data and so was not used by Celgene to update the economic model for bortezomib retreatment. We have independently checked this and are satisfied that the abstract included OS K-M data but did not provide PFS K-M data. We independently identified a related and recently published paper (San-Miguel et al. 2014)[15] which provided PFS HRs for the previous use of bortezomib but no PFS K-M data for this subpopulation. We assume that this paper was excluded by Celgene at full text screening stage, as summarised in the PRISMA diagram (Figure 1), although this cannot be verified as no further details of excluded studies are provided in the current submission.

3.4 Conclusions of the clinical effectiveness section

We are satisfied that the company presented a reasonable update of the systematic review of the literature, although we noted a lack of rigor in reporting of both systematic review methodology and the data extraction process which may have increased the possibility of bias. Their searches were thought to be satisfactory and their inclusion criteria, while not fitting the Scope exactly, were considered appropriate. We have concluded that, despite ten studies being excluded on the basis of being non-English language, the company is unlikely to have missed any evidence.

The updated search identified 11 unique papers (both journal articles and conference papers) which included participants who had received bortezomib as their initial treatment. The data extraction table contained typographical errors, both in data values and references, but we are satisfied that this did not affect the use of any relevant data in the economic model.

Two papers were identified providing overall survival Kaplan-Meier data (Ahn 2014; San-Miguel 2015). The Ahn et al. (2014) paper, a retrospective observational study investigating the clinical efficacy and safety of bortezomib retreatment in participants with relapsed or refractory multiple myeloma, additionally provided Kaplan-Meier data for progression-free survival and was therefore selected for inclusion within the economic model. Again, we noted a number of typographical errors in the reporting of the Ahn et al. paper, but were satisfied that this did not affect the use of the correct data in the economic model.

We independently checked the remaining nine papers and were satisfied that none contained Kaplan-Meier data for overall survival and hence did not contain usable data for the economic model.

We note that the data from Ahn et al. used to update the scenario analysis includes mainly participants suited to thalidomide. [4] This is in line with the use of data from MM009[16] and MM010[17] and Taverna et al. (2012),[18] where intention to treat data (including participants suited to thalidomide) was used. Although the proportion of participants with prior treatment of thalidomide changes with study population, we do not believe this to be a significant concern. Previously in this appraisal, the NICE Appraisal Committee recognised that the MM-009 / MM-010 trial populations did not match the population set out in the decision problem and concluded that study populations that include participants with prior treatment of thalidomide are generalisable to the population of interest.[19]

4 Cost-effectiveness

4.1 Celgene's review of cost-effectiveness evidence

The cost effectiveness searches were not updated for the submission. Additional documentation supplied by Celgene, following our request to see the full search strategies, states that: "*only the clinical effectiveness search was updated, as this was where the comparative effectiveness estimates were where the key uncertainty lay*" (Celgene 2016, data on file).

4.2 Summary and critique of the economic evaluation submitted by Celgene

4.2.1 Summary of the 2014 model

Prior to the current model submitted by Celgene (the "2016 Celgene model"), we believe that the most recent previous version of the model is the Evidence Review Group (ERG)-corrected version submitted in June 2014, which followed responses to the original appraisal consultation document (ACD). In April 2014 Celgene produced a new analysis, in which the ICERs changed significantly from those reported in their 2013 submission. In this 2014 submission, the lenalidomide (LEN) arm dominated bortezomib (BOR) and had an ICER of £55,000 per QALY versus melphalan plus prednisolone (MP). The ERG response from June 2014 highlighted several concerns with this model. Under the preferred analysis, where no subsequent lines of treatment were included and the corrections suggested by the ERG response were made, the ICERs came to £11,000 per QALY for LEN versus BOR and £44,000 for lenalidomide versus MP. This analysis maintained Celgene's preferred adjustment for covariates (mean of covariates), used to calculate the comparative efficacy of the treatment arms. We believe this is the last version of the model that the Committee has seen. This will henceforth be referred to as the '2014 model'. The main assumptions of the 2014 model are reported in Table 4.

Of note, subsequent (3rd-line) therapies are not included in the 2014 model, as recommended in the ERG 2014 document. Prior to the 2014 model, the original Celgene response to the ACD included subsequent therapies where lenalidomide was included after discontinuation of the comparator and as part of best supportive care (BSC) in 4th line treatment for both arms. Dexamethasone (DEX) was included as part of a mix of treatments used in third and fourth line for both arms. Celgene state that these treatments were identified from 2007-2009 Haematological Malignancy Research Network (HMRN) data collected by York (Celgene executable model, 2016).

In the base case for the 2014 model, LEN arm, OS was fitted with a Gompertz distribution, PFS with a Gamma distribution and TTF with a Gamma distribution independent of PFS.

The PFS and OS hazard ratios (HRs) for LEN vs. BOR and MP were estimated using the median PFS and OS for LEN and the median PFS and OS for the comparators. Median PFS and OS were estimated from Taverna et al (2012)[18] for the BOR arm and Petrucci et al. (1989)[20] for the chemotherapy arm (MP).

As noted in the ERG response to Celgene's comments following the ACD in 2014: "*PFS [progression free survival], TTF [time to failure] and OS [overall survival] curves derived from*

MM-010 data were adjusted using the mean of covariates method, by which average values of covariates (like for example the beta-2 microglobulin count and presence of bone lesions for the baseline MM-010 population) were entered into a proportional hazards regression equation. The choice of relevant predictors of PFS, TTF and was not very transparent in the submission. The Committee agreed that several factors predicting survival were not adjusted for, and that it was not clear why the 3 factors used had been chosen. The Committee therefore concluded that the data had not been appropriately adjusted (ACD, section 4.8)"[21]

For the April 2014 iteration of the model Celgene had updated the mean of covariates (MOC) method to use the pooled data for the LEN arm from the MM-010 and MM-009 trials. The ERG response in June 2014 also provided an alternative method for adjusting the covariates (corrected group prognosis [CGP]), but MOC is the method originally employed by the company and we believe is the method the committee is most familiar with. It is used in the 2014 model.

Though not reported, costs appear to be estimated for 2012-13.

Other important assumptions in the 2014 model include:

- The number of BOR cycles was not capped.
- The patient access scheme (PAS) for lenalidomide was not applied in 2nd-line treatment, but was applied in the comparator treatment arms in 3rd-line.

Table 4. Comparison of 2014 model and 2016 Celgene model.

Model input	2014 base case	2016 changes
Time horizon	25 years	-
Cycle length	4 weeks	-
Subsequent therapies	No	Yes
Comparators	BOR, MP	-
Method for covariate adjustment in estimation of PFS and OS	Mean of covariates	Multi-state modelling
OS distribution	Gompertz	Multi-state modelling
PFS	Gamma	Multi-state modelling
TTF	Gamma (separate to PFS)	Multi-state modelling
Comparator efficacy	HR obtained using median PFS, OS for LEN, BOR and MP	As 2014 model, but with adjustment for OS for BOR and MP
Comparator efficacy source	BOR: Taverna et al. (2012) MP: Petrucci et al. (1989)	-
Maximum number of BOR+DEX treatment cycles	Unlimited	8
Lenalidomide PAS	3rd-line, comparator arm only	LEN in any arm or treatment line
Transport costs	Yes	No
Cost year	2012-2013	-

Key: -, unchanged from 2014; BOR, bortezomib; BSC, best supportive care; DEX, dexamethasone; HR, hazard ratio; LEN, lenalidomide; MP, melphalan plus prednisolone; OS, overall survival; PAS, patient access scheme; PFS, progression free survival; QALYs, quality adjusted life years; TTF, time to failure

4.2.2 Changes from 2014 model to Celgene 2016 model

Table 4 gives the main similarities and differences between the 2014 and 2016 models (i.e. Celgene's latest model version). These changes are discussed in the subsections below.

In summary, the changes that have had a substantial impact on the cost-effectiveness of LEN compared to bortezomib or chemotherapy are given in Figure 2, Figure 3 and Table 5, and are:

a) Incorporation of the PAS for 2nd-line LEN.

b) The duration of bortezomib (BOR) treatment is now restricted to 8, 3-week treatment cycles.

c) Estimation of TTP PFS and OS for LEN is now based on the multi-state modelling (MSM) method. OS for LEN is now substantially longer tailed than in the 2014 model.

d) The estimation of the HR for PFS and OS for LEN vs. BOR and LEN vs. MP is now adjusted using the MSM method, whereas in the 2014 model, the mean of covariates method was used.

e) After application of the relevant HR, OS for BOR and MP has then been altered (columns M and N of worksheet "PF.Comparator", Celgene 2016 model). Celgene have not reported this important change to their model. We discuss this change in detail in Section 4.2.2.4, pp.34-36.

f) In the comparator arms, LEN as assumed to be taken 3rd-line by all patients that survive to progression.

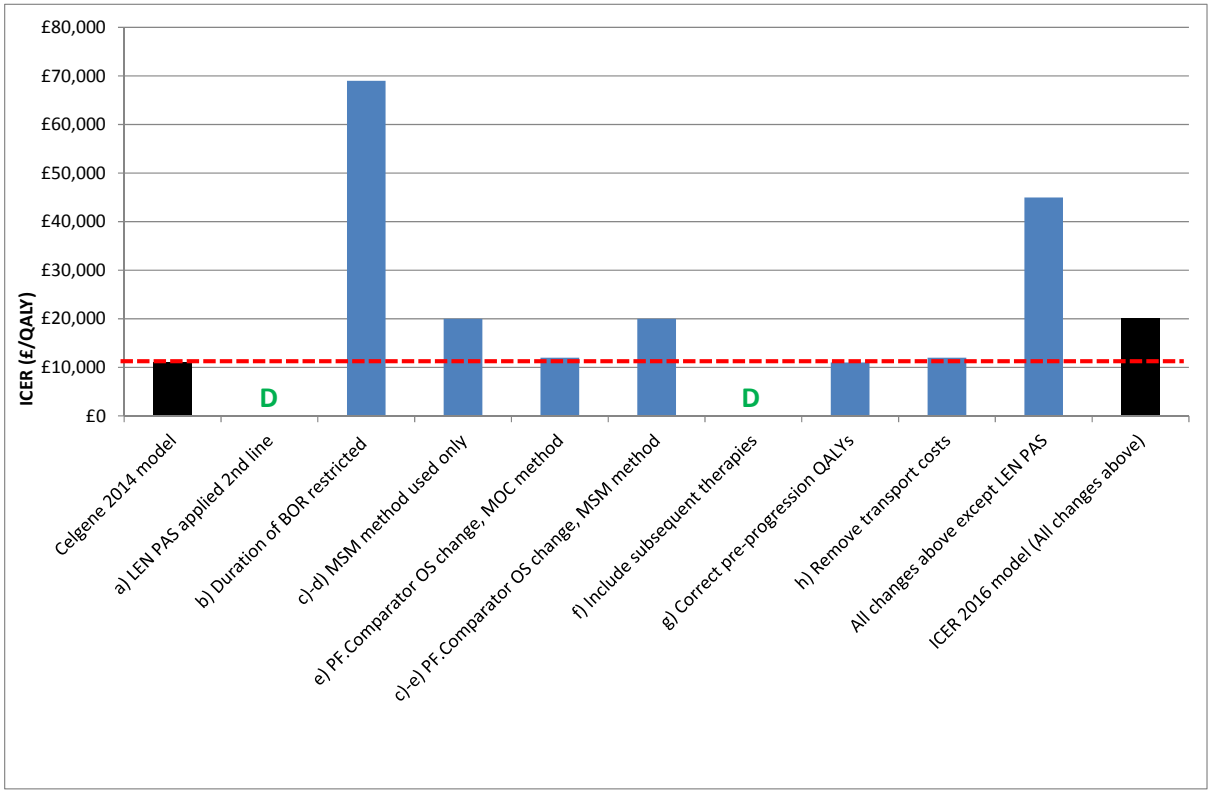
Changes which have a minimal effect on the cost-effectiveness of LEN are:

g) Correction of pre-progression QALYs to all rows of comparator arm of the Markov model (the original correction was only applied to the first row).

h) Removal of transport costs.

We have considered both the appropriateness of the changes and the implementation of these changes within the model.

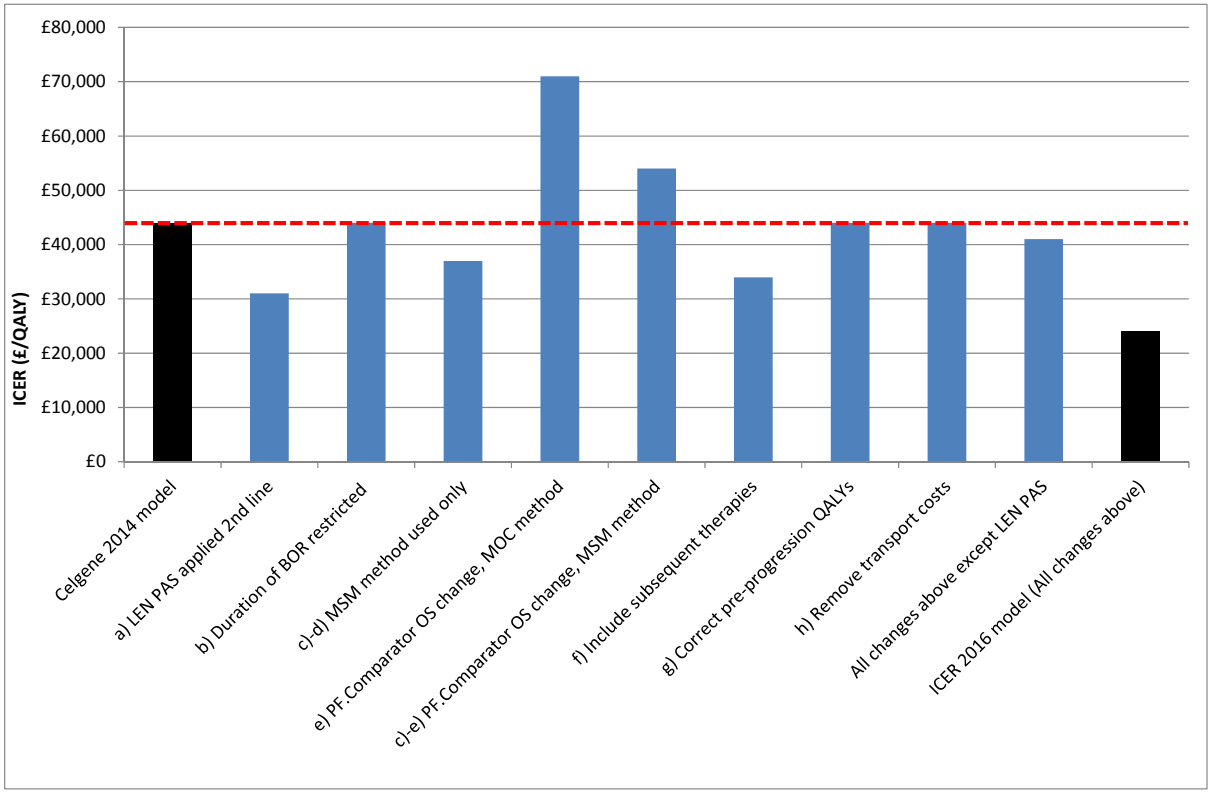
Figure 2. Impact of changes to 2014 model on ICERs for LEN vs. BOR



Key: BOR, bortezomib; D, lenalidomide dominates; ICER, incremental cost-effectiveness ratio; LEN, lenalidomide; MOC, mean of covariates; MSM, multi-state modelling; OS, overall survival; PAS, patient access scheme; QALY, quality adjusted life year

Notes: ICERs are rounded to the nearest £1,000. Negative ICERs are not given on this figure. Instead, **D** represents results where lenalidomide dominated the comparator (i.e. had reduced costs and increased QALY gains). The red dotted line represents the 2014 ICER, LEN vs. BOR.

Figure 3. Impact of changes to 2014 model on ICERs for LEN vs. MP



Key: BOR, bortezomib; ICER, incremental cost-effectiveness ratio; LEN, lenalidomide; MOC, mean of covariates; MP, melphalan plus prednisolone; MSM, multi-state modelling; OS, overall survival; PAS, patient access scheme; QALY, quality adjusted life year

Notes: ICERs are rounded to the nearest £1,000. The red dotted line represents the 2014 ICER, LEN vs MP.

Table 5. Impact on ICERs of Celgene changes to the 2014 model

	LEN vs. BOR	LEN vs. MP
Celgene submission post-ACD	LEN dominates	£55,000
2014 model	**£11,000**	**£44,000**
a) LEN PAS applied second line	LEN dominates	£31,000
b) Duration of BOR restricted	£69,000	£44,000
c) and d) MSM method used only (previously MOC method)	£20,000	£37,000
e) Change to PF.Comparator OS modelling keeping MOC method	£12,000	£71,000
c)-e) Change to PF.Comparator OS modelling and MSM method	£20,000	£54,000
f) Include subsequent therapies (LEN taken 3rd-line in comparator arm)	LEN dominates	£34,000
g) Correct pre-progression QALYs	£11,000	£44,000
h) Remove transport costs	£12,000	£44,000
All changes above except LEN PAS	£45,000	£41,000
Celgene 2016 model (All changes above)	**£20,000**	**£24,000**

Key: BOR, bortezomib; ICER, incremental cost-effectiveness ratio; LEN, lenalidomide; MOC, mean of covariates; MP, melphalan plus prednisolone; MSM, multi-state modelling; OS, overall survival; PAS, patient access scheme; QALY, quality adjusted life year

Notes: LEN dominates refers to the instances where lenalidomide arm had both reduced costs and increased QALYs compared to the comparator arm.

4.2.2.1 LEN PAS

In the 2014, model, the LEN PAS was applied for 3rd-line treatment only in the comparator arms. The PAS is as follows: for any patient who stays on lenalidomide for more than treatment 26 cycles, the treatment costs post cycle 26 are met by Celgene.

In the 2016 model, the LEN PAS is included in all treatment lines, including 2nd-line. This is appropriate.

4.2.2.2 Bortezomib treatment duration

In their response to the ACD in 2014, Celgene advocated treatment to progression with bortezomib. However the ERG response from June 2014 questioned this, citing current clinical practice and the Taverna et al. study where >90% patients received 1-6 treatment cycles of BOR and only 2% received >10 cycles.[18, 21] Therefore in Celgene's current submission, a cap of 8 cycles has been applied to treatment with BOR.

Celgene explain that BOR is given in 3-weekly cycles, and so 8 treatment cycles equates to 6 model cycles. We agree, and this has been modelled correctly.

However, in Taverna et al. (2012), the study which Celgene use for the clinical effectiveness of BOR, and as acknowledged in the ACD (p57),[19] there was a median of 3 cycles of BOR re-treatment. Further, we estimate a mean of 3.8 treatment cycles (in our amended version of Celgene's 2016, cell CW8, model worksheet "PF.Comparator").[18] This is substantially shorter than the modelled mean of 6.6 treatment cycles. Given that PFS and OS for BOR are estimated from Taverna et al. (2012), we believe that it is important to also estimate the

mean number of BOR cycles from Taverna et al. (2012). For the model, we consider re-treatment with BOR at 2nd-line. It could therefore be argued that the estimated mean of 3.8 cycles from Taverna et al. (2012) is too short, as this study considers patients with a median of 2 prior treatments, and we might expect treatment duration to be shorter at later lines of treatments. However, there is little evidence for this, as initial BOR treatment in Taverna et al. (2012) was for a median of 4 cycles,[18] and we estimate a mean of 4.5 cycles (in our amended version of Celgene's 2016, cell CY8, model worksheet "PF.Comparator") which is similar to the number of cycles for BOR re-treatment.

Also, it could be argued that we ought to adjust the BOR treatment duration from Taverna et al. (2012) to estimate BOR treatment duration if it had been used in the MM RCTs, by adjusting for the differences in patient characteristics between the Taverna et al. and MM RCTs. We notice that the median PFS of LEN of 48.1 weeks in the MM RCT increases to 56 weeks under the MSM method, when assuming patient characteristics in Taverna et al. (2012). Therefore, arguably, we should assume a mean duration of BOR, assuming patient characteristics in the MM RCTs, of 48.1 / 56 x 3.8 = 3.3 cycles. However, given that we cannot be certain that the adjustment to treatment duration is directly analogous to that for PFS, we leave unchanged the estimated mean duration of BOR at 3.8 cycles.

In the PenTAG base case, the total per patient cost of acquisition and cost of administration of BOR is multiplied by a factor of 58% (in our amended version of Celgene's 2016, cell CW10, model worksheet "PF.Comparator"), equal to the ratio of the estimated mean cycles from Taverna et al. (2012), 3.8, to the mean modelled 6.6 cycles. In this case, Celgene's 2016 base case ICER for LEN vs. BOR increases substantially, from £20,000 to £29,000 per QALY.

4.2.2.3 Modelled TTF, PFS and OS LEN

In both the 2014 and 2016 models, the modelled TTF, PFS and OS for LEN were estimated from the 2nd-line pooled data from the MM RCTs.

TTF, PFS and OS in the 2014 model were estimated based on the mean of covariates (MOC) method, and by the multi-state modelling (MSM) method in the 2016 model.

TTF for LEN is virtually unchanged from the 2014 model, whereas PFS and OS for LEN have become longer tailed (Figure 4). Celgene have not reported any validation of this change to the tails. Mean PFS has increased slightly from 2.6 to 3.0 years and mean OS substantially from 4.7 to 5.9 years.

As required, the PFS and OS curves do not cross. Also, OS is clearly below that of the general population of England and Wales, which was a concern in a previous version of Celgene's model (ACD p.28).[19]

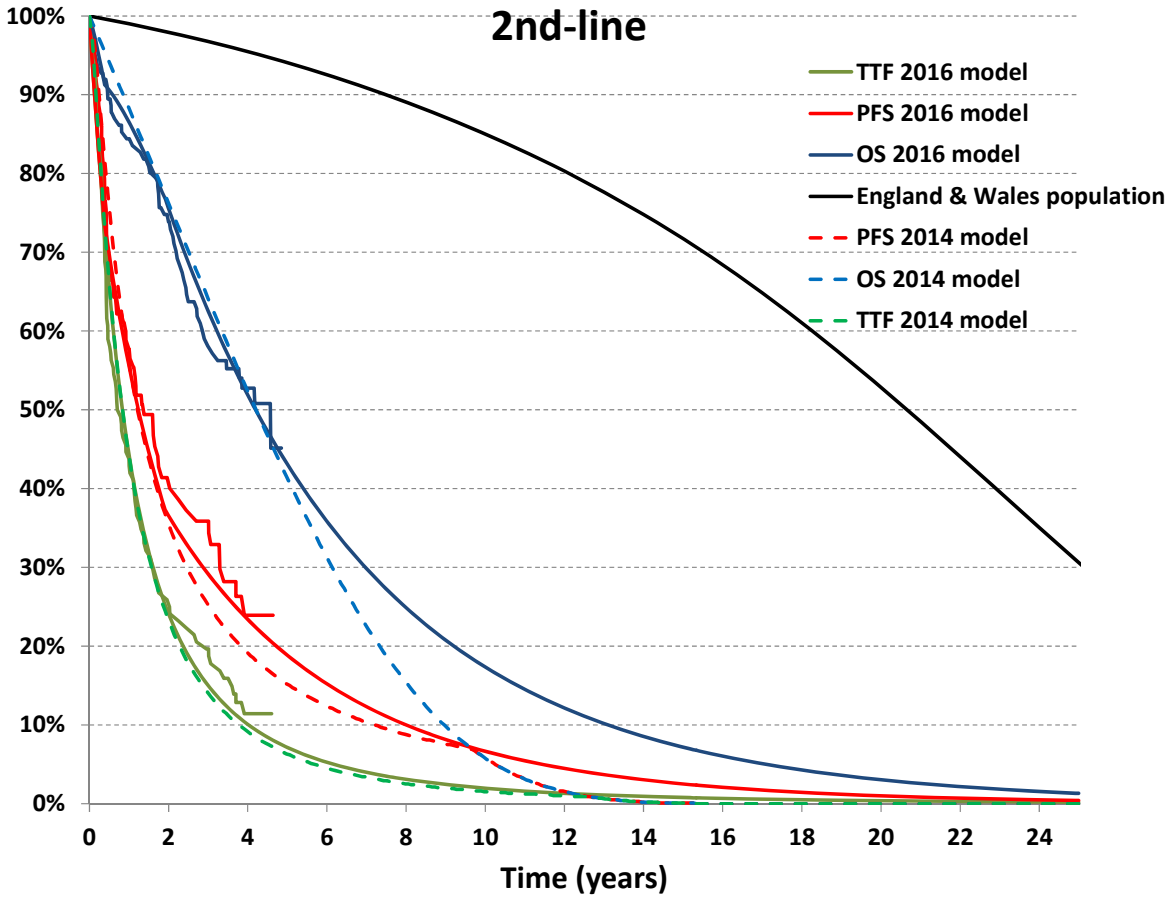
The changes to TTF, PFS and OS for LEN alone substantially improve the cost-effectiveness of LEN. Based on the 2016 model, but assuming TTF, PFS and OS for LEN in 2014, whilst leaving the treatment effects of the comparators unchanged, the ICER for LEN vs. BOR increases from £20,000 to £30,000 per QALY, and for LEN vs MP from £24,000 to £34,000 per QALY.

In the next section, we explain that Celgene have included a correction to the estimation of OS for BOR and MP. We argue below that we disagree with this correction. If we remove the correction, and, as before, if we assume TTF, PFS and OS for LEN in 2014, whilst leaving the treatment effects of the comparators unchanged, the ICER for LEN vs. BOR

increases from £20,000 to £24,000 per QALY, and decreases for LEN vs MP from £24,000 to £22,000 per QALY.

The updated modelled TTF, PFS and OS visually appear to fit well to the empirical data for 2nd-line LEN+DEX from the pooled MM09 and MM10 RCTs (Figure 4). However, PFS and particularly OS are immature. This means that there is great uncertainty in the extrapolated portions of these curves.

Figure 4. TTF, PFS and OS for LEN: 2014 vs. 2016 models



Key: OS, overall survival; PFS, progression-free survival; TTF, time to treatment failure.

4.2.2.4 Modelled PFS and OS for BOR and MP

The clinical data underlying the estimate of PFS and OS for BOR is taken from Taverna et al. (2012).[18] In summary, this is very poor quality evidence, because clinical effectiveness is not randomised between LEN and BOR, and Taverna et al. (2012) was a retrospective survey of just 42 patients in Switzerland.

In its original submission, Celgene assumed that the efficacy of MP was the same as bendamustine (p35 ACD and Table p16 ACD).[19]

At the ACD stage and now, OS and PFS for MP is taken from Petrucci et al.(1989).[20] This represents extremely poor quality evidence for the following reasons:

- Again, clinical effectiveness is not randomised between LEN and MP.
- Petrucci et al. (1989) was published 27 years ago and is unlikely to represent current clinical practice or outcomes.
- It considered a small cohort of 34 patients.
- As we, the ERG report from 2014, and our clinical advisor previously noted (p43 ACD),[19] it is likely that OS for MP from Petrucci (1989) underestimates the effect relevant for this STA, as all patients had received prior chemotherapy, and it is likely that this prior treatment reduces the effectiveness of chemotherapy.

In the 2014 model, the HRs for PFS and OS for LEN vs. BOR and LEN vs. MP were estimated using the mean of covariates (MOC) method. In the 2016 model, these HRs are instead estimated using the MSM method (Table 4, Table 6).

To summarise, the HR of 1.34 for PFS for LEN vs. BOR under the MSM method in the 2016 model is estimated as follows.

1. Take from Taverna et al. (2012) the median PFS for BOR of 45.7 weeks.
2. Take from the pooled MM RCTs the median PFS for LEN for all patients (2nd-line and beyond) of 48.1 weeks.

A naïve estimate of the PFS HR would then be 48.1 / 45.7 = 1.1 favouring LEN.

However, the method continues:

3. The MSM method is then applied to adjust median PFS for LEN from 48.1 to 56.0 weeks for differences in patient baseline characteristics between Taverna et al. (2012) and the MM RCTs.
4. A correction factor is applied to convert between TTP and PFS, as Taverna et al. (2012) reports only TTP, whereas we require PFS (which is reported from the MM RCTs).
5. The PFS HR is estimated from just the adjusted median PFS for LEN and the unadjusted median PFS for BOR, whilst assuming constant hazards in both treatment arms.

A similar procedure is followed to estimate the OS HR of 1.71 for LEN vs. BOR.

1. Take from Taverna et al. (2012) the median OS for BOR of 88.7 weeks.
2. Take from the pooled MM RCTs the median OS for LEN for all patients (2nd line and beyond) of 164.3 weeks.
3. Under the MSM method, the median OS for LEN is reduced from 164.3 to 152 weeks in an attempt to adjust for differences in patient baseline characteristics between Taverna et al. (2012) and the MM RCTs.
4. The HR is estimated simply as the ratio of these medians 152 / 88.7 = 1.71, again assuming constant hazards in both treatment arms.

Table 6. Celgene estimated HRs for PFS and OS: 2014 vs. 2016 model

	2014 model (MOC)	2016 model (MSM)
PFS: LEN vs. BOR	1.19	1.34
OS: LEN vs. BOR	1.89	1.71
PFS: LEN vs. MP	4.66	4.37
OS: LEN vs. MP	4.66	4.37

Key: BOR, bortezomib; HR, hazard ratio; LEN, lenalidomide; MOC, mean of covariates; MP, melphalan plus prednisolone; MSM, multi-state modelling; OS, overall survival; PFS, progression free survival

PFS and OS for BOR and MP are then calculated by applying the HRs as follows:

$$S(t)_{BOR/MP} = S(t)_{LEN} \text{ ^ } HR_{BOR/MP}, \text{ where S(t) represents PFS or OS.}$$

We agree that the amended HRs have been applied correctly in the 2016 model.

We note that the cost-effectiveness of LEN is not sensitive to differences in patient baseline characteristics between Taverna et al. (2012) and the MM RCTs as follows. If we did not adjust the HRs for difference in such characteristics, i.e. the MSM method is not applied for the HRs, the ICER LEN vs. BOR increases only slightly, from £20,000 to £24,000 per QALY, and the ICER for LEN vs. MP is unchanged at £24,000 per QALY.

We further note that the cost-effectiveness of LEN is insensitive to the change in PFS and OS HRs between the 2014 and 2016 models, as follows. When we use the 2016 model with the HRs from the 2014 model, the ICER LEN vs. BOR increases only slightly, from £20,000 to £22,000 per QALY, and the ICER for LEN vs. MP is unchanged at £24,000 per QALY.

We find all the above to be reasonable. However, in their 2016 model, Celgene now adjust OS for BOR and MP further. We do not see the need for this adjustment, as explained in Section 4.2.4, p.38.

4.2.2.5 3rd-line and 4th-line treatments

As stated in the ERG report from 2014 (p.59),[22] only the costs of subsequent treatment are modelled. Any effects of subsequent treatment on OS are ignored.

We believe that it is essential that if the costs of subsequent treatments are modelled, then the clinical effectiveness of all treatment arms should reflect use of such subsequent treatments. However, we do not know which subsequent treatments were used in the MM RCTs that inform the effectiveness of LEN, in the Taverna et al. (2012) study that informs the effectiveness of BOR or in the Petrucci et al. (1989) study that informs the effectiveness of MP.

In the 2014 model, no subsequent (3rd-line) treatments were assumed (Table 4, p.28). In the 2016 model, Celgene assume that, in both comparator arms, LEN is taken 3rd-line by all patients that survive to progression. This adds a very substantial cost to the comparator arms (£20,000 per patient in the BOR arm, and £23,000 in the MP arm). They also assume very low cost (£400 per patient) 3rd-line treatment (a mixture of dexamethasone, prednisone, prednisolone, cisplatin, cyclophosphamide, doxorubicin, etoposide, melphalan, vincristine, bortezomib and thalidomide) in the LEN arm, and 4th-line treatment (a mixture of cyclophosphamide, thalidomide, melphalan and lenalidomide) of similar cost for all treatment arms (£6,100 per patient in the LEN arm, £5,300 in the BOR arm and £5,900 in the MP arm).

Given that we do not know the nature of subsequent treatments in the underlying clinical data, we believe that this change is inappropriate and that the 2014 model was preferable in this respect.

Indeed, in our base case, we assume no costs for subsequent treatments. This is consistent with the NICE committee's view at the ACD stage (ACD p.55): "*The Committee ….. noted that the manufacturer had modelled the costs and effects differently in the lenalidomide and the comparator arms in all 4 base cases presented, and agreed that this was not appropriate and instead costs and effectiveness should be modelled consistently in the lenalidomide and comparator arms.*"[19]

When we remove all costs of subsequent treatment, the cost-effectiveness of LEN worsens substantially: the ICER for LEN vs. BOR increases from £20,000 to £36,000 per QALY for LEN vs. MP from £24,000 to £37,000 per QALY.

We note the following concerns of the ERG from 2014 (ACD p.30): "*The ERG commented that the manufacturer's defined clinical treatment pathway differed from the third- and fourth-line treatments included in the model, and that several modelled treatments were no longer routinely used in clinical practice in the UK. The ERG questioned whether clinicians in the UK would offer third- or fourth-line treatments in clinical practice. Taking this and the other issues into account (see sections 3.42 and 3.43), the ERG therefore questioned whether these treatment lines should be included in the model*." [19]

Next, we recommend that Celgene give more clarity on how they cost subsequent treatments. Indeed, it is noted in the ERG critique of Celgene's model from 2014, that the calculation of third line costs (with or without the PAS) was unclear; and they described the PAS calculations as "*confusing*".[22] From our examination of the model, Celgene appear to have estimated the average number of LEN cycles per patient (and the average number of cycles that are costed before the PAS) based on 3rd-line TTF estimates from the MM-010 and MM-009 trials. The comparator sheet of the Markov model then uses this to estimate the number of new people entering 3rd-line lenalidomide treatment per cycle and the number who have remained on lenalidomide treatment from the previous cycle. However the model appears set up to only remember the last 12 model cycles, suggesting patients are being followed for only a year.

We believe that it is unlikely that 3rd-line LEN is taken by all patients in the comparator arms that survive to progression in clinical practice. It is also unlikely to have been the case in Taverna et al. (2012) and almost certainly not in Petrucci et al. (1989) given the age of this study.

Treatments for multiple myeloma are continually changing. For example, panobinostat in combination with bortezomib and dexamethasone has recently been approved by NICE for adults with relapsed and/or refractory multiple myeloma who have received at least 2 prior regimens including bortezomib and an immunomodulatory agent.[1]

We also note that the mix of treatments assumed 4th-line, is taken from HMRN data from 2007-2009. This is unlikely to reflect the current distribution of treatments given that this is now several years old.

4.2.2.6 Modelling of pre-progression QALYs

In 2014, the ERG report identified a wiring error in the modelling of pre-progression QALYs in the comparator arm (2014 model worksheet "PF.Comparator"). Whilst the formula was

corrected in the first row of calculations in 2014, this had not been copied it down into all rows of the model. Celgene have now implemented it correctly in all rows of the comparator arm. As noted in our comparisons between models, this has minimal impact on the ICER (Table 5, p.32). We discuss this no further.

4.2.2.7 Cost of transport

The ERG response to Celgene's comments post-ACD states: "*for the purpose of calculating transportation costs Celgene assumed that 50% of patients require transportation for their treatment administration and also that if more than one treatment occurs during one week the patient will be kept in the hospital for up to one week. Clinical opinion sought by the ERG did not believe this to be a reasonable assumption. Firstly the percentage of MM patients requiring transportation to the hospital was considered to be significantly lower than 50%. Secondly, the assumption that patients in need of more than one treatment per week would stay in the hospital was believed to be unrealistic. These assumptions are likely overestimating the cost of Bort [bortezomib] in the economic analysis as only one administration visit is considered for Len [lenalidomide].*" (ERG response June 2014)[21]

Therefore, in the current submission, Celgene have attempted to remove transport costs from the model. These have minimal impact on cost-effectiveness results, but when implementing the changes, we note that the transport cost of doxorubicin (~£17 per cycle) is still included. However, this is a very small cost with negligible impact on the ICER. Therefore, we discuss this issue no further.

4.2.3 Checking the Celgene 2016 model

We checked the 2016 model in two ways.

First, as explained in the previous section, we were able to reconcile the 2016 model with the 2014 model, subject to the changes described above.

Second, we built a simplified version of the 2016 model, which was based on mean times in state, as opposed to use of discrete model cycles. Approximations were made for discounting of costs and QALYs. We were able to recreate the results of the full 2016 model to a reasonable degree of accuracy.

4.2.4 Outstanding and further criticisms of the Celgene 2016 model

4.2.4.1 Modelled TTF, PFS and OS for LEN

The maximum follow-up of TTF, PFS and OS data underlying the analysis is 4.6 years (Figure 4). PFS and OS are immature, which thus necessitates substantial extrapolation in the model. The empirical data is given in Dimopoulos et al. (2009), which gives data up to July 2008. The maximum follow-up now, March 2016, is approximately 4.6 + 7.5 = 12.1 years. Therefore, we encourage Celgene to use outcomes data that is far more mature than presented in their latest submission.

If much more mature data were used, this would substantially reduce the uncertainty in cost-effectiveness of LEN as we would no longer need to extrapolate PFS and OS over a long period of time.

4.2.4.2 Modelled OS for BOR and MP

As explained previously, in the 2014 model, OS for BOR and MP was calculated by raising the OS for LEN to the power of the appropriate OS HR. As mentioned in Section 4.2.2.4, p.34, we are satisfied with this.

In the 2016 model, the resulting OS for BOR and MP is now adjusted.

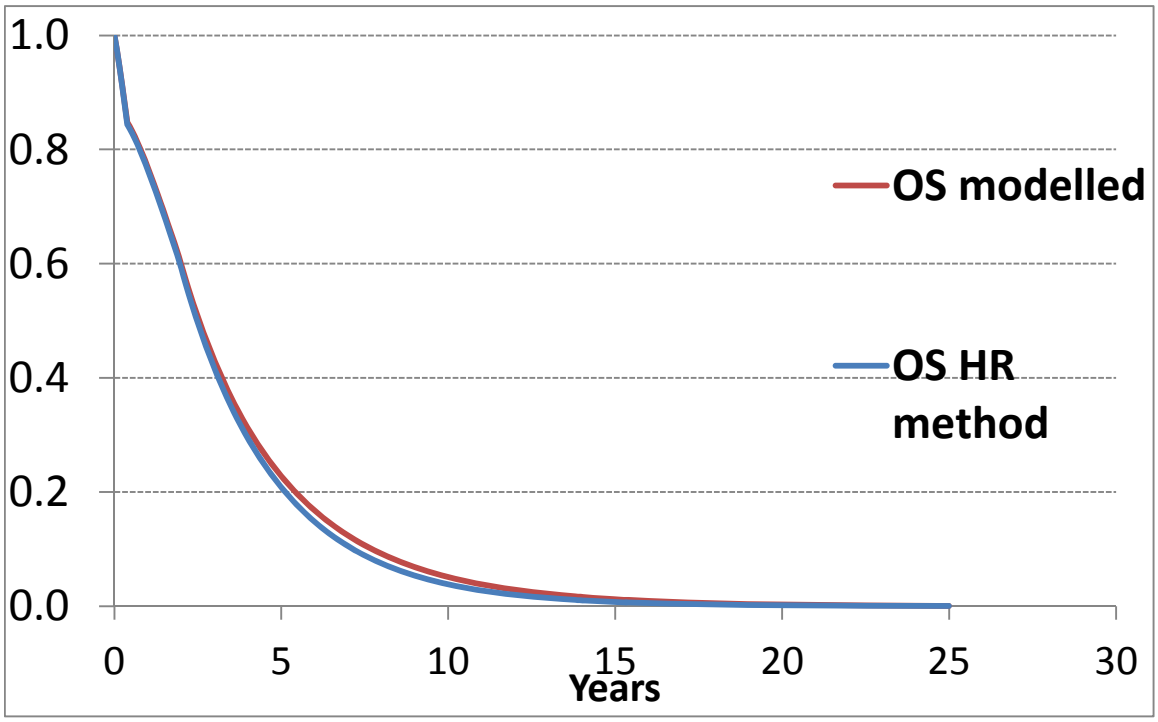
This change extends OS for BOR only incrementally, but it extends OS substantially for MP. The differences in OS between that predicted by the HR alone and that actually modelled (with the adjustment) are given in Figure 5 and Figure 6.

Celgene do not discuss this important change in their recent document discussing the PAS. Furthermore, we can see no reason for this adjustment.

Therefore, in our base case analysis, we remove the adjustment, and instead estimate OS for BOR and MP based on the HR method.

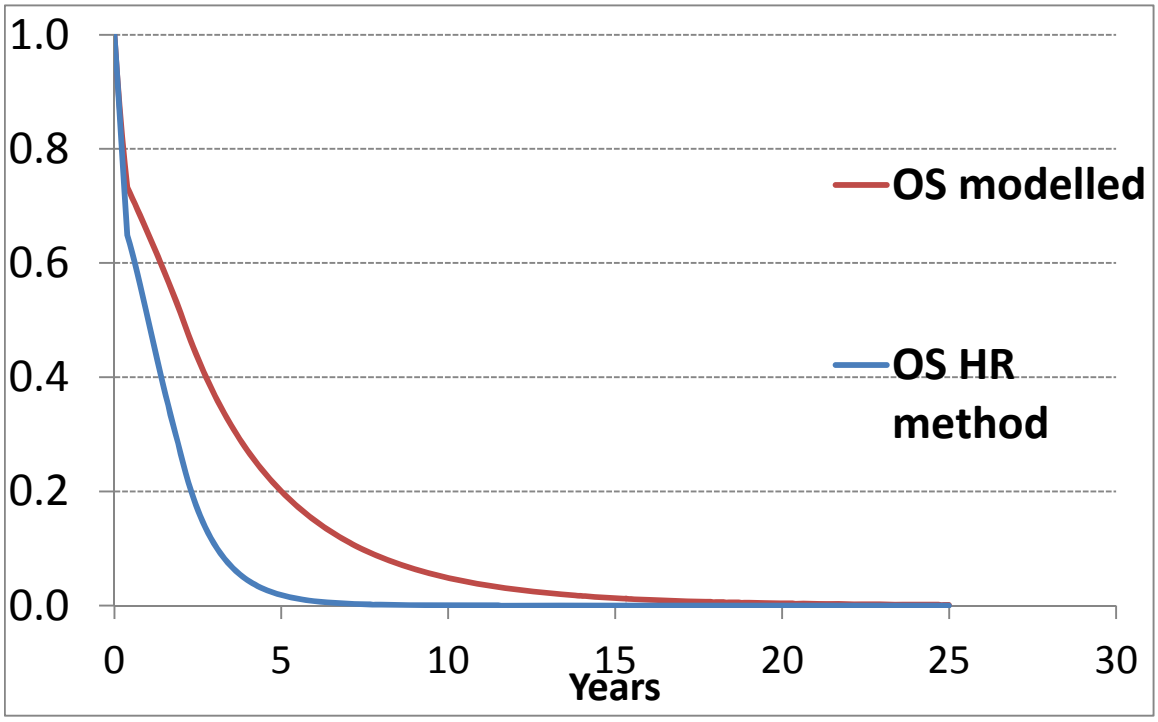
When we remove the adjustment, the cost-effectiveness of LEN is virtually unchanged for LEN vs BOR, but improves substantially vs MP: the ICER for LEN vs. MP decreases from £24,000 to £19,000 per QALY.

Figure 5. OS for BOR: Celgene 2016 modelled (with adjustment) vs. estimated by HR alone



Key: BOR, bortezomib; HR, hazard ratio; OS, overall survival

Figure 6. OS for MP: Celgene 2016 modelled (with adjustment) vs. estimated by HR alone



Key: HR, hazard ratio; MP, melphalan plus prednisolone; OS, overall survival

4.2.4.3 Utilities

We understand that quality of life data was not collected in the MM RCTs.

Celgene continue to estimate utilities based on Agthoven et al. (2004) from a 2002 PhD thesis which, to the Committee's knowledge, has never been published in a peer-reviewed journal (ACD p.58).[19] Therefore, the utilities are highly uncertain. This then acts to increase the uncertainty in the ICERs.

At the ACD stage, the Committee noted that the utility values were derived from a younger population, and were higher than the average population of the same age (ACD p.58).[19] We find that the utilities in PFS (0.81 up to 2 years and 0.77 thereafter) are similar to, but not higher than, those for members of the general public of England and Wales.

4.2.4.4 Cost of administration of bortezomib

Celgene currently assume that all patients take BOR intravenously. However, as mentioned in the ACD (p58),[19] a proportion will likely take BOR subcutaneously. The Committee agreed that, if some patients receive BOR subcutaneously, the costs of administration and transportation, and the disutility associated with intravenous therapy would fall, and the cost effectiveness of LEN would likely worsen vs. BOR (ACD p.58).[19]

4.2.4.5 Comparator treatments

It is clearly important to select the most relevant comparators to assess the cost-effectiveness of lenalidomide. We are concerned that treatments for multiple myeloma may have changed since the start of the current appraisal. We contacted our clinical expert to confirm what treatments are currently in place on the NHS and also considered the current

NICE guidance for our relevant patient population. A summary of our findings are presented in Table 7.

BOR monotherapy is currently the only treatment with guidelines second line. However, our clinical expert suggests that retreatment with BOR is unlikely to be used in current practice.

Our clinical expert suggests that bendamustine would be the most likely comparator, though we note that this application is not included in the license.[23]

Table 7. Second line comparator treatments

2nd-line intervention/comparator included in scope	Included in Celgene 2016 model?	Recommended 2nd line by NICE guidelines	Use in clinical practice	View from NICE ACD 2014[19]
Lenalidomide + dexamethasone	Yes	None	~0% (previously on CDF)	N/A
Bortezomib (+dexamethasone)	Yes	TA129 recommends BOR monotherapy for patients who have received one prior treatment, under a rebate scheme[1]	~0% unless post thalidomide. Monotherapy rarely used.	The Committee heard from the clinical specialist that this population may receive bortezomib re-treatment at first relapse after initial treatment with bortezomib. However, the clinical specialist stated that this may not be appropriate for more than half of this population either because their condition does not respond to bortezomib or because of adverse reactions. The clinical specialist explained that for these people, in the absence of lenalidomide, current treatment options would be limited to standard chemotherapy and bendamustine.(ACD p.49).
Chemo incl. regimens based on melphalan, vincristine, cyclophosphamide and doxorubicin	Yes	None	Rarely used	Committee agreed this is a comparator (ACD p.49). The ERG sought the opinion of a clinical specialist in 2014 who stated that chemotherapy was likely to be less effective than lenalidomide, and that only about 5% of people with MM receive chemotherapy as a 2nd-line treatment (ACD p.43).
Bendamustine	No	None	For patients contraindicated to thalidomide this is most likely treatment	Based on the opinion of the clinical specialist, the Committee agreed that it was appropriate not to consider bendamustine as a second-line treatment (ACD p.49). The clinical specialist and patient experts stated that bendamustine was rarely offered to patients because it is not licensed for second-line treatment. They noted that it is licensed for first-line treatment but is not used in this way, that it is available through the Cancer Drugs Fund for relapsed multiple myeloma where other treatments are not appropriate, and that it is used in clinical practice as fourth- and fifth-line treatment. (ACD p.49).

Notes: Clinical expert opinion was provided by Dr Claudius Rudin (Royal Devon and Exeter Hospital); and suggests that <5% patients would be contraindicated to thalidomide, both thalidomide and bortezomib are used interchangeably as first and second line options, and that bortezomib retreatment is not funded by commissioners.

4.3 Cost effectiveness results of Celgene 2016 model

4.3.1 Base case

Celgene present results both with and without the lenalidomide second line PAS. Reporting of results in the submission generally matched those in the model, subject to a couple of minor rounding errors.

As this was an update of their previous analysis and the model has been through many iterations, we would have appreciated a summary comparison of the current results to the 2014 results.

We also stress that the limited clinical evidence for the comparators means these results remain extremely uncertain. Indeed, the NICE committee understood the extreme uncertainty at the ACD stage (ACD p.59).[19]

Table 8. Celgene submission 2016 base case results, no PAS 2nd line

	Lenalidomide	Bortezomib retreatment	Melphalan plus Prednisone
Intervention cost (£)	90,945	19,036	**926**
Other costs (£)	20,031	39,672	**41,684**
Total costs (£)	110,976	58,708	**42,610**
Difference in total costs Lenalidomide vs comparator (£)	-	52,269	**68,366**
LYG	5.867	3.685	**3.153**
LYG difference Lenalidomide vs comparator	-	2.182	**2.714**
QALYs	3.546	2.374	**1.880**
QALY difference Lenalidomide vs comparator	-	1.172	**1.666**
ICER Lenalidomide vs comparator (£) per QALY	-	44,605	**41,030**

Key: LYG: life-year gained; QALY: quality-adjusted life-year; ICER: incremental cost-effectiveness ratio
Notes: We noted in the Celgene submission that the CIC information presented here could be back-calculated. We have therefore marked up additional cells as CIC
Source: Celgene PAS submission, Table 5, p.30

Though not reported in Celgene's submission, we note in the model that the life years spent in pre-progression and post-progression are nearly equal in the LEN arm (just under 3 years). In the BOR arm, most life years are gained in the pre-progression state, but the difference between pre- and post-progression states is not stark (~2 years vs. ~1.7 years). In comparison, for the MP arm, very few life years are accrued pre-progression (<0.5 years) and over 2.7 years are accrued in the post-progression health state. Despite the OS HRs suggesting that BOR arm should have a greater survival gain than MP, the total life years accrued are similar for BOR and MP compared with the 5.87 life years accrued in the lenalidomide arm (3.69 for bortezomib, 3.15 for melphalan plus prednisolone). This is due to the adjustment made to the OS of the comparators (Section 4.2.4.2, p.39).

The largest cost in both comparator arms is for LEN given third line. The cost of acquisition of bortezomib is also large.

Table 9. Celgene submission 2016 base case results, PAS 2nd line

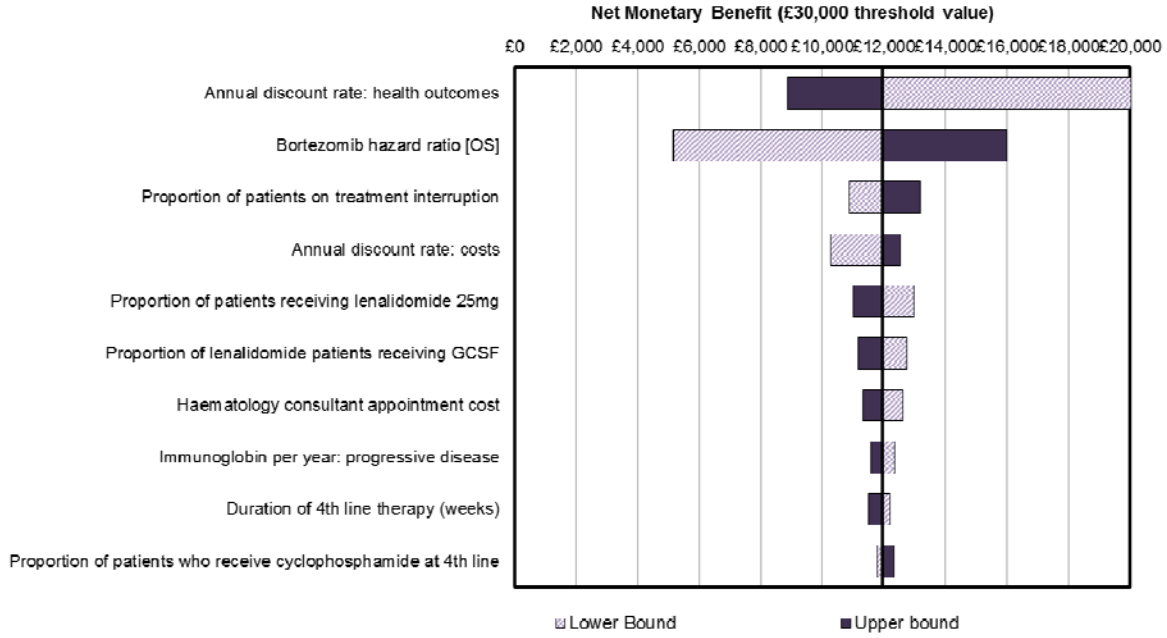
	Lenalidomide	Bortezomib retreatment	Melphalan plus Prednisone
Intervention cost (£)	61,856	19,036	**926**
Other costs (£)	20,031	39,672	**41,684**
Total costs (£)	81,887	58,708	**42,610**
Difference in total costs Lenalidomide vs comparator (£)	-	23,179	**39,277**
LYG	5.867	3.685	**3.153**
LYG difference Lenalidomide vs comparator	-	2.182	**2.714**
QALYs	3.546	2.374	**1.880**
QALY difference Lenalidomide vs comparator	-	1.172	**1.666**
ICER Lenalidomide vs comparator (£) per QALY	**-**	**19,781**	**23,572**

Source: Celgene PAS submission, Table 6, p.30
Key: LYG: life-year gained; QALY: quality-adjusted life-year; ICER: incremental cost-effectiveness ratio
Notes: We noted in their submission that the CIC information presented here could be back-calculated using the differences. We have therefore marked up additional cells as CIC

4.3.2 Sensitivity analyses

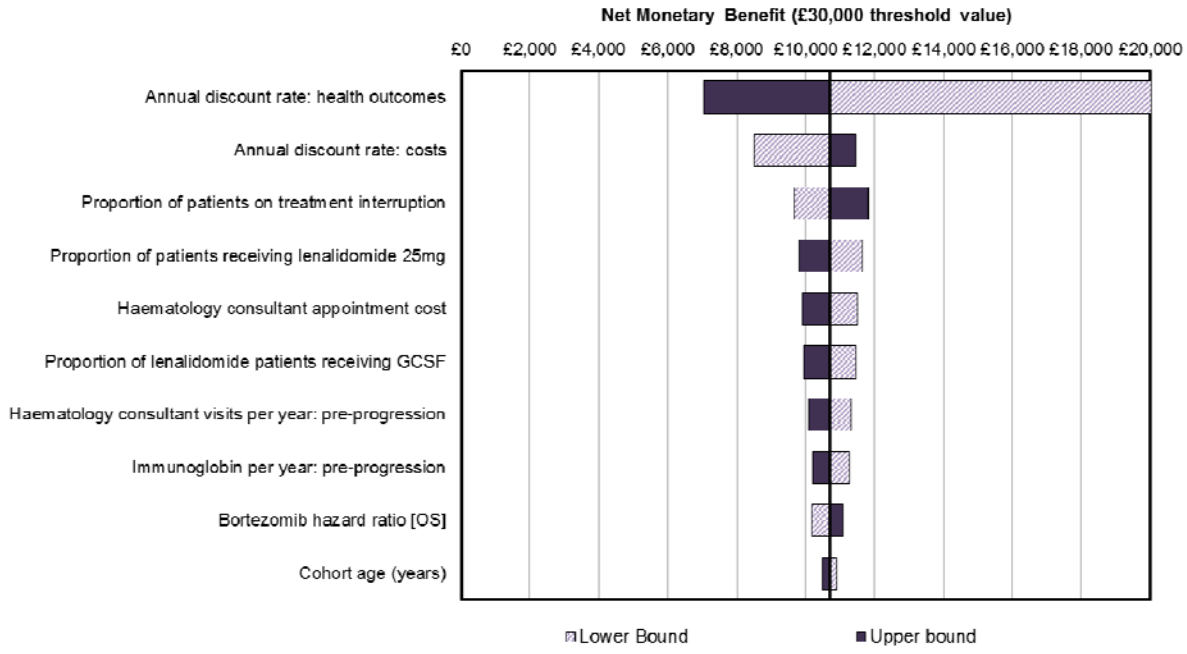
Celgene presents one way sensitivity analyses for the PAS results. Notably no sensitivity analysis causes the net monetary benefit (NMB) to fall below £0 at the £30,000 per QALY willingness to pay (WTP) threshold. Rerunning with a £20,000 per QALY WTP threshold results in negative NMB for nearly all sensitivity analyses when MP is the comparator (with exception of the health outcomes annual discount rate) and all the sensitivity analyses cross the £0 NMB line for the BOR arm. Of note, the magnitude of the difference in NMB has reduced from the 2014 analyses, and the discount rates for costs and QALYs are now some of the most influential parameters.

Figure 7. Tornado diagram LEN (with PAS) vs. BOR



Key: BOR, bortezomib; GCSF, Granulocyte-colony stimulating factor; LEN, lenalidomide; mg, milligrams; OS, overall survival; PAS, patient access scheme
Source: Celgene submission 2016, page 33, Figure 6

Figure 8. Tornado diagram LEN (with PAS) vs. MP

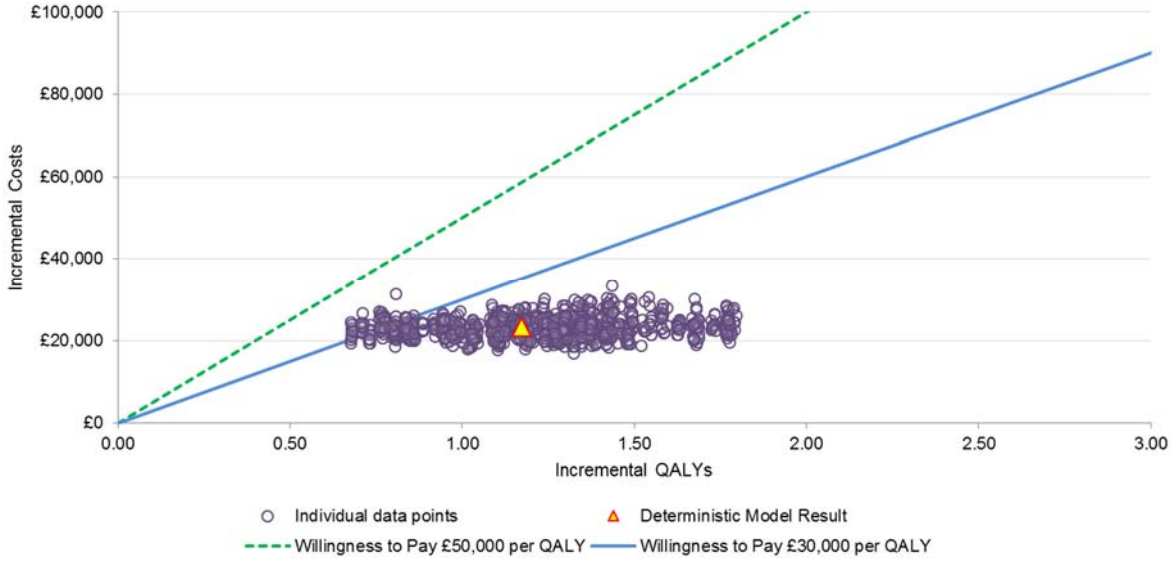


Key: GCSF, Granulocyte-colony stimulating factor; LEN, lenalidomide; mg, milligrams; MP, melphalan plus prednisolone; OS, overall survival; PAS, patient access scheme
Source: Celgene submission 2016, page 34, Figure 7

Celgene also present the probabilistic sensitivity analyses (PSA) under the PAS. Most uncertainty lies in the QALY gain when compared to both comparators and this uncertainty drives whether lenalidomide appears cost-effective compared to BOR or MP at a willingness to pay threshold of £30,000 per QALY. The PSA scatterplots fall within the upper right quadrant of the cost-effectiveness plane, such that lenalidomide always results in increased costs and increased QALYs. The cost-effectiveness acceptability curves (CEACs) suggest that for both comparators, LEN is most likely to be cost-effective at a willingness to pay threshold of £30,000 per QALY. We have not critiqued the parameter ranges used to produce the PSA results, nor their implementation, so cannot comment on their appropriateness.

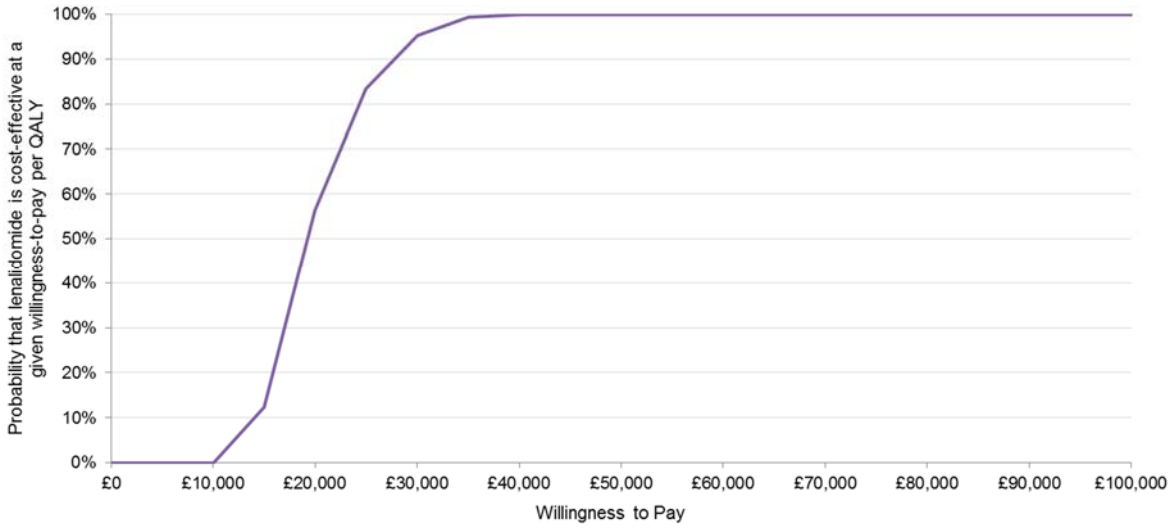
We note that though the PSAs may identify some of the parameter uncertainty, they cannot represent all of the uncertainty in the model, particularly structural uncertainty. Given that there is substantial structural uncertainty in Celgene's model, we consider the PSA results to be of little value.

Figure 9. PSA scatterplot LEN (with PAS) vs. BOR



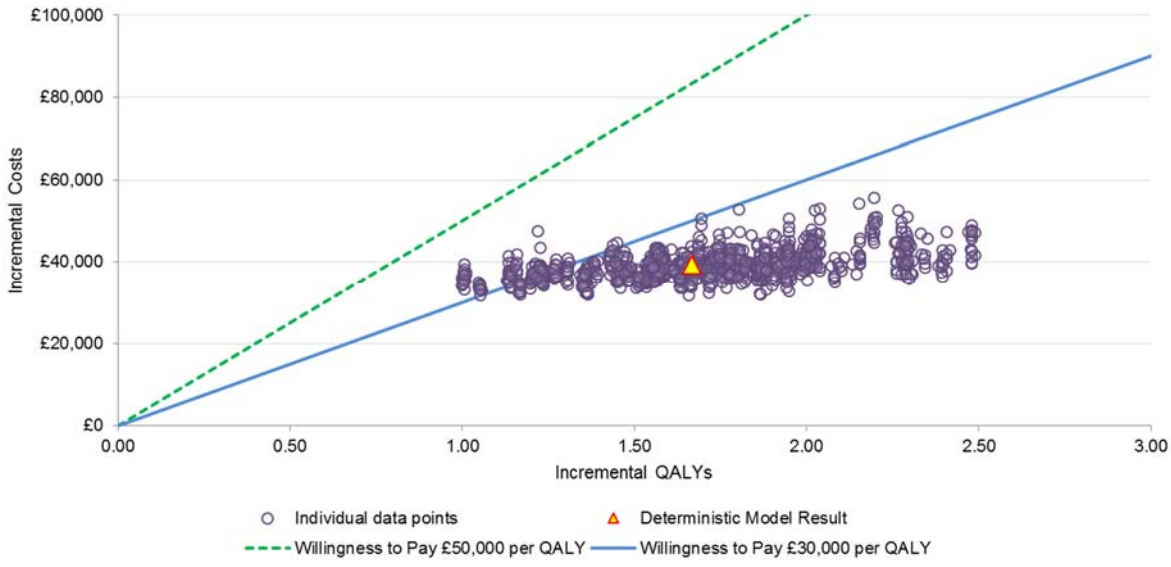
Key: BOR, bortezomib; CEAC, cost-effectiveness acceptability curve; LEN, lenalidomide; PAS, patient access scheme; PSA, probabilistic sensitivity analysis, QALY, quality adjusted life year

Figure 10. CEAC LEN (with PAS) vs BOR



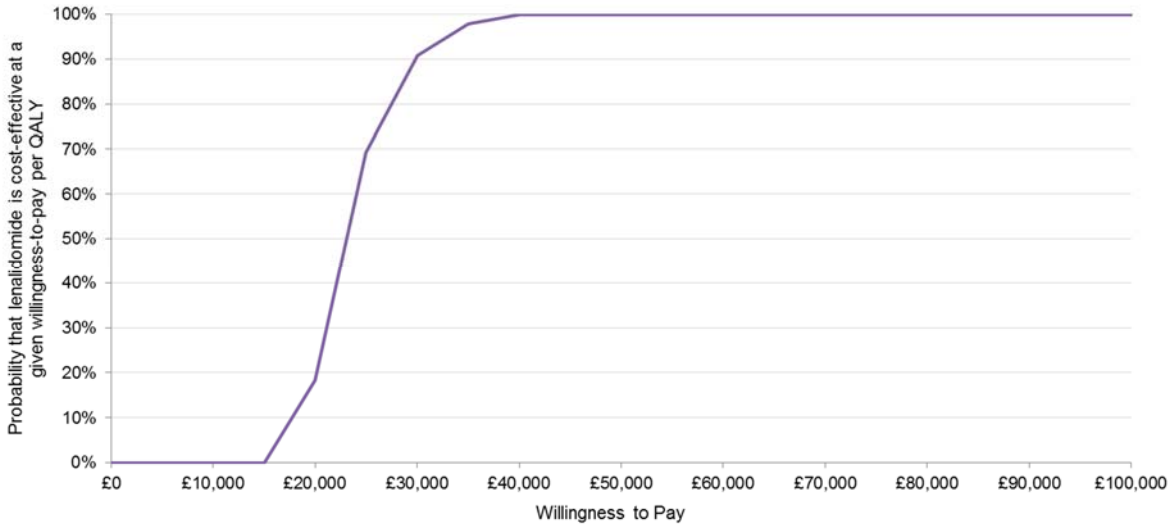
Key: BOR, bortezomib; CEAC, cost-effectiveness acceptability curve; LEN, lenalidomide; PAS, patient access scheme; PSA, probabilistic sensitivity analysis, QALY, quality adjusted life year

Figure 11. PSA scatterplot LEN (with PAS) vs. MP



Key: LEN, lenalidomide; MP, melphalan plus prednisolone; PAS, patient access scheme; PSA, probabilistic sensitivity analysis, QALY, quality adjusted life year

Figure 12. CEAC LEN (with PAS) vs. MP



Key: CEAC, cost-effectiveness acceptability curve; LEN, lenalidomide; MP, melphalan plus prednisolone; PAS, patient access scheme; PSA, probabilistic sensitivity analysis, QALY, quality adjusted life year

4.3.3 Scenario analyses

Celgene provide several new scenario analyses. The reported results of these match those given in the model. A summary of the results is given below, with further details given in the Celgene submission pp.39-46.

Table 10. Celgene scenario analyses results

Scenario	ICER LEN vs. BOR	ICER LEN vs. MP
Celgene 2016 base case	£20,000	£24,000
MOC method	£26,000	£29,000
CGP method	£28,000	£31,000
Uncapped BOR	LEN dominates	£24,000
MSM using adjust parametric survival curves	£17,000	£24,000
MSM using TTF based on HR to PFS curve	£19,000	£22,000
Weighted average of subsequent treatment in both arms[a]	£27,000	£24,000

Notes: a Celgene refer to this scenario as "MSM with no 3rd line lenalidomide".
Source: Celgene submission, Tables 11-17, pp. 39-46

We do note that Celgene's second scenario is supposed to serve a similar function to the results we provided in Section 4.2.2, p.29, as it is supposed to demonstrate "*the impact of the PAS in isolation of the methodological improvements*" (Celgene submission, pp.39-40). However, by altering the model to use the MOC method (as opposed to the MSM method), this really only demonstrates the impact of the change in methodology, and not in PAS. We believe our analysis (applying the PAS to the 2014 model) demonstrates the impact of the PAS (separate to the other changes), better than this scenario.

4.4 Exploratory and sensitivity analyses undertaken by PenTAG

The PenTAG base case comprises three changes to the Celgene 2016 model. These changes have all been discussed above:

- No subsequent (3rd- and 4th-line) treatments, as assumed in the 2014 model (Section 4.2.2.5, p.36).
- Reduce the mean duration of BOR from 6.6 to 3.8 treatment cycles (Section 4.2.2.2, p.32)
- OS for BOR and MP based on HR alone, as assumed in the 2014 model (Section 4.2.2.4, p.34).

At the ACD stage, the NICE committee considered that the End of Life criteria do not apply in this assessment (ACD, p.60).[21] We believe this remains the case. Under Celgene's current base case, life expectancy is clearly greater than the 2 year threshold for BOR and MP. Under our base case, life expectancy for BOR remains clearly greater than 2 years; therefore, the criteria do not apply for the comparison with BOR. Conversely, life expectancy for MP is now below the threshold, at 1.5 years. However, we consider that the criteria do not apply for the comparison with MP, as the clinical data is not robust.

4.5 Impact on the ICER of additional clinical and economic analyses undertaken by PenTAG

Table 11 shows the impact of our additional analyses on the ICER. It is very important to appreciate that all ICERs in Table 11 are highly uncertain because:

1. The underlying clinical data is not randomised.
2. The quality of the clinical data used to inform PFS and OS for BOR and MP is extremely low.
3. Given that modelled TTF, PFS and OS for LEN are based on immature data, substantial extrapolation is required. .
4. The utilities are highly uncertain.
5. The nature of subsequent treatments is uncertain.

Celgene currently assume that all patients take BOR intravenously. However, as mentioned above, a proportion will likely take BOR subcutaneously. If we were to model a proportion taking BOR subcutaneously, this would increase the ICER of LEN vs. BOR.

Table 11. Impact on the ICER of additional analyses undertaken by PenTAG

	LEN vs. BOR	LEN vs. MP
Celgene 2016 model	**£20,000**	**£24,000**
1: No 3rd- or 4th-line treatments	£36,000	£37,000
2: Reduce mean duration of BOR from 6.6 to 3.8 treatment cycles	£29,000	£24,000
3: OS for BOR and MP based on HR alone	£20,000	£19,000
1 & 2	£45,000	£37,000
1 & 3	£35,000	£26,000
2 & 3	£28,000	£19,000
PenTAG base case (1 & 2 & 3)	**£44,000**	**£26,000**

Key: BOR, bortezomib; HR, hazard ratio; ICER, incremental cost-effectiveness ration; LEN, lenalidomide; MP, melphalan plus prednisolone; OS, overall survival

5 Overall conclusions

In Celgene's current (2016) base case, the ICER for LEN vs. BOR is £20,000 per QALY and for LEN vs. MP, £24,000 per QALY.

We made three key changes to Celgene's current model to derive our base case:

- Remove all subsequent treatments.
- Reduce the mean duration of BOR from 6.6 to 3.8 treatment cycles.
- Remove the Celgene's 2016 adjustment to OS for BOR and MP.

Under our base case, the ICER for LEN vs. BOR is £44,000 per QALY and for LEN vs. MP, £26,000 per QALY.

5.1 Strengths and limitations of the clinical update

The company present a reasonable update of the systematic review of the literature, although we note a lack of rigor in reporting of both systematic review methodology and the data extraction process which may have increased the possibility of bias. Celgene's searches were thought to be satisfactory and their inclusion criteria, while not fitting the Scope exactly, were considered appropriate. We have concluded that, despite ten studies being excluded on the basis of being non-English language, the company is unlikely to have missed any evidence.

The updated search identified 11 unique papers (both journal articles and conference papers) which included participants who had received bortezomib as their initial treatment. We independently checked all papers and were satisfied that only one (Ahn et al. 2014; San-Miguel et al. 2015) provided Kaplan-Meier data for both overall survival and progression-free survival and was therefore suitable for inclusion within the economic model.

No studies which included both lenalidomide and a relevant comparator arm were identified.

5.2 Strengths and limitations of the cost-effectiveness update

Celgene present an update of their model from 2014, where they have addressed some of the important concerns of the ERG and Committee raised in the 2014 documents. They have implemented most of their changes correctly and no new major wiring errors have been identified in the model.

Celgene present a scenario analysis using the Ahn et al. (2014) study, for the bortezomib (BOR) arm, but give the results using the parametric survival curves, which they have not done for the base case study (Taverna et al., 2012). They have also not explained their reasons for preferring Taverna et al. (2012) as their base case.

However our main concerns are with the remaining uncertainties, detailed below.

5.3 Remaining uncertainties

We believe the important remaining uncertainties are:

- Choice of comparator. This project began in 2013, and we feel it would be prudent to assess whether the comparators defined in 2013 are still relevant to current clinical practice. The opinion of our clinical expert suggests this is not the case.
- Overall survival for lenalidomide. We believe that much more mature data for OS for LEN should be available. If such data were available, this would substantially reduce the uncertainty in the extrapolation of lenalidomide survival.
- Quality of evidence for comparator clinical effectiveness. The MM-009 and MM-010 provide high quality evidence for the effectiveness of LEN vs. DEX. However, the evidence for the comparator effectiveness of LEN vs. BOR or LEN vs. MP is of very low quality. This is because the evidence is not randomised, and because the studies of BOR and MP are very small.

5.4 Implications for research

Here we highlight research priorities:

- Choice of treatment comparators. As this STA began three years ago, in 2013, we feel it would be appropriate to assess whether the comparators identified then are still applicable.
- Consider the possibility of conducting a RCT for patients relevant to the current STA of lenalidomide vs. the most relevant comparators.
- Overall survival for lenalidomide. We believe that much more mature data for OS for LEN should be available. If such data were available, this would substantially reduce the uncertainty about in the extrapolation of the lenalidomide survival estimates.

References

1. National Institute for Health and Care Excellence. NICE pathways: Multiple myeloma 2016 [20th February 2016]. Available from: http://pathways.nice.org.uk/pathways/blood-and-bone-marrow-cancers.

2. San-Miguel JF, Vania TM, Hungria YS-S, et al. Final analysis of overall survival from the Phase 3 Panorama 1 trial of panobinostat plus bortezomib and dexamethasone versus placebo plus bortezomib and dexamethasone in patients with relapsed or relapsed and refractory multiple myeloma. American Society of Hematology; Orlando, FL, USA. 5-8 December 2015. 2015.

3. Hulin C, Rubia J, Dimopoulos M, et al. Bortezomin (BTZ) retreatment for relapsed/refractory (REL/REF) multiple myeloma (MM) in real-world medical practice: final data from the international, non-interventional, observational EVOBS study. Haematologica; June 13-16; Stokholm, Sweden. 2013.

4. Ahn JS, Jung SH, Lee SS, Ahn SY, Yang DH, Kim YK, et al. Clinical outcome of bortezomib retreatment in patients with relapsed or refractory multiple myeloma. Biomed Res Int. 2014;2014.

5. Berenson J, Bitran J, Tabbara I, et al. A Retrospective Study to Evaluate the Incidence and Severity of Peripheral Neuropathy for Relapsed/Refractory Multiple Myeloma Patients upon Retreatment with Bortezomib. American Society of Hematology; Dec 6-9; san Francisco, CA, United States. 2014.

6. Kim HJ, Kwon J, Jung JY, et al. Metronomic chemotherapy of thalidomide, cyclophosphamide, and dexamethasone for relapsed/refractory multiple myeloma patients. Clinical Lymphoma Myeloma and Leukemia. 2015; 15:e311-e2.

7. Mao J, Cheng F, Chen H, Wang J, Zhou X, Jiang Y, et al. Short-course bortezomib-based retreatment for patients with multiple myeloma who had received bortezomib-thalidomide-dexamethasone (VTD) as an initial therapy: A single-center case series. Exp Ther Med. 2014; 7(4):977-81.

8. Oriol A, Giraldo P, Kotsianidis I, Couturier C, Olie R, Angermund R, et al. Efficacy and safety of bortezomib-based retreatment at the first relapse in multiple myeloma patients: A retrospective study. Hematology. 2015; 20(7):405-9.

9. Cerchione C, Madonna E, Pareto A, et al. Management of relapsed and refractory multiple myeloma with bendamustine. Haematologica; June 12-15; Milan, Italy. 2014.

10. Reece D, Trieu Y, Masih-Khan E, et al. Treatment of relapsed/refractory multiple myeloma with weekly cyclophosphamide plus bortezomib plus prednisone or dexamethasone (CyBorP/D): Updated experience at Princess Margaret Cancer Centre (PMCC). American Society of Clinical Oncology; May 30 - June 3; Chicago, IL, United States. 2014. p. e19568.

11. Jagannath S, Roy A, Kish J, et al. Progression-Free Survival in Third-Line Treatment of Relapsed Refractory Multiple Myeloma Among Patients Who Have Received Prior Bortezomib and Immunomodulatory Drugs. American Society of Hematology; December 5-8; Orlando, FL, United States. 2015.

12. Cerchione C, Madonna E, Marano L, et al. A new option for relapsed/refractory multiple myeloma. Haematologica; June 13-16; Stokholm, Sweden. 2013.

13. National Institute for Health and Care Excellence. NICE technology appraisal guidance [TA171]. Lenalidomide for the treatment of multiple myeloma in people who have received at least one prior therapy. National Institute for Health and Care Excellence, TA171.

14. National Institute for Health and Care Excellence. Guide to the methods of technology appraisal. . 2013.

15. San-Miguel JF, Hungria VT, Yoon SS, Beksac M, Dimopoulos MA, Elghandour A, et al. Panobinostat plus bortezomib and dexamethasone versus placebo plus bortezomib and dexamethasone in patients with relapsed or relapsed and refractory multiple myeloma: a multicentre, randomised, double-blind phase 3 trial. Lancet Oncol. 2014;15(11):1195-206.

16. Weber DM, Chen C, Niesvizky R, Wang M, Belch A, Stadtmauer EA, et al. Lenalidomide plus dexamethasone for relapsed multiple myeloma in North America. N Engl J Med. 2007;357(21):2133-42.

17. Dimopoulos MA, Chen C, Spencer A, Niesvizky R, Attal M, Stadtmauer EA, et al. Long-term follow-up on overall survival from the MM-009 and MM-010 phase III trials of lenalidomide plus dexamethasone in patients with relapsed or refractory multiple myeloma. Leukemia. 2009;23(11):2147-52.

18. Taverna C, Voegeli J, Trojan A, Olie RA, von Rohr A. Effective response with bortezomib retreatment in relapsed multiple myeloma--a multicentre retrospective survey in Switzerland. Swiss Med Wkly. 2012;142:w13562.

19. National Institute for Health and Care Excellence. Lenalidomide for treating multiple myeloma after 1 prior treatment with bortezomib (part-review of TA171): appraisal consultation document.

20. Petrucci MT, Avvisati G, Tribalto M, Cantonetti M, Giovangrossi P, Mandelli F. Intermediate-dose (25 mg/m2) intravenous melphalan for patients with multiple myeloma in relapse or refractory to standard treatment. Eur J Haematol. 1989;42(3):233-7.

21. Bacelar M, Durand A, Hyde C, Latimer N. The clinical and cost-effectiveness of lenalidomide for people who have received at least one prior therapy with bortezomib (partial review of TA171). Celgene's response to Technology Appraisal Committee's Critique of the Economic Case. A Critique of the Response Submitted by Celgene. Matrix and Peninsula Technology Assessment Group, 2014.

22. Bacelar M, Durand A, Cooper C, Hyde C, Latimer N, Murray D. The clinical and cost-effectiveness of lenalidomide for people who have received at least one prior therapy with bortezomib (partial review of TA171). Single Technology Appraisal. Matrix and Peninsula Technology Assessment Group, 2014.

23. electronic Medicines Compendium. Summary of Product Characteristics: Bendamustine hydrochloride Accord 2.5 mg/ ml Powder for concentrate for solution for infusion. 2016 [cited 2016 1/3/16]. Available from: http://www.medicines.org.uk/emc/medicine/31286.

The clinical and cost-effectiveness of lenalidomide for people who have received at least one prior therapy with bortezomib (partial review of TA171)

Addendum- additional PenTAG analyses

1st April 2016

Contents

List of tables

Summary

The most recent ERG report for this STA was submitted by PenTAG to NICE on 7th March 2016.

In that submission, the PenTAG base case differed from Celgene's base case in three ways:

- Third and fourth line treatments were removed.
- The mean number of bortezomib cycles was reduced from 6.6 to 3.8.
- Overall survival for the comparators was based on the overall survival hazard ratios only.

This addendum presents additional analyses requested by NICE, following the submission from 7th March. Table 1 presents the summary ICERs for the new analyses and the sections where more detailed results can be found.

In Section 1 we provide additional clarification on the multi-state Markov model produced by Celgene in their 2016 submission.

In Section 2 we produce a scenario analysis for the PenTAG base case, where the patient access scheme (PAS) for second line bortezomib is assumed to equal an absolute cost discount of 15% (Celgene previously estimated this to be 8.3%).

In Section 3 we explore the impact of assuming that the effectiveness of bortezomib is equal to lenalidomide (i.e. the PFS and OS for bortezomib equal the PFS and OS for lenalidomide). This is applied to the following scenarios:

- Celgene's base case (Section 3.1)
- PenTAG's base case (Section 3.2)
- PenTAG's base case with the bortezomib PAS equal to an absolute cost discount of 15% for second line bortezomib (Section 3.3)

Table 1. Summary analyses

	ICER LEN vs. BOR (£/QALY)	Section in this addendum
Celgene 2016 model	**£20,000**	
PenTAG base case	**£44,000**	
PenTAG base case, BOR PAS = 15% discount	£44,000	Section 2, p.6
LEN and BOR PFS and OS assumed equal		
Celgene 2016 model	£183,000	Section 3.1, p.7
PenTAG base case	£2,120,000	Section 3.2, p.9
PenTAG base case, BOR PAS = 15% discount	£2,155,000	Section 3.3, p.10

Notes: ICERs rounded to nearest 1,000
Key: BOR, bortezomib; HR, hazard ratio; ICER, incremental cost-effectiveness ration; LEN, lenalidomide; OS, overall survival; PAS, patient access scheme; PFS, progression free survival; QALY, quality adjusted life year

1 Multi-state model clarification

In order to help clarify the multi-state modelling (MSM) technique that Celgene have presented, we provide the following information.

MSM is a way of describing a process in which an individual moves through a series of states in continuous time. In our case, progression free survival (PFS), progressive disease (PD), dead. The changes of state in a MSM usually occur at unknown times.

The probabilities of moving from PFS to death, PFS to PD and PD to death are estimated by maximum likelihood. It is possible to specify either that these probabilities are constant over time or vary over time. We understand that this is a valid and well established technique. Indeed, there is a package ("msm") in the R statistics program that does MSM. Celgene have used this package. To run this routine, you need the time of the observation, the observed health state, and the patient ID number.

The alternative to the MSM method is to fit PFS by maximum likelihood to the underlying individual patient data (IPD) and then independently fit overall survival (OS) by maximum likelihood, which is the method commonly used by companies.

We understand that there are the following advantages of the MSM method over the normal method:

- It is impossible for the PFS and OS curves to cross. Crossing was a problem earlier in this STA.
- The correlation between PFS and OS can be modelled correctly for the probabilistic sensitivity analysis (PSA).

However, we are not convinced that these represent a major advantage with the normal method.

Regardless of whether the MSM method or the normal independent fitting is performed, we still need to make assumptions for the extrapolation period (from the time of maximum follow up in the MM RCTs and until all patients are dead). By using the MSM method for the extrapolation period as well as during the trial follow up period, Celgene are assuming that the trends in PFS and OS observed from the MM RCTs apply for the extrapolation period. They provide no evidence to test this important assumption, e.g. they do not compare extrapolated PFS or OS versus observational mature data.

2 Additional analyses: bortezomib patient access scheme

The patient access scheme (PAS) for bortezomib provides a refund for patients who do not achieve "*at least a Minimum Response (a 25% or greater reduction in serum M-protein) within the first 4 cycles of treatment*".[1]

Celgene implemented this complex PAS into their model in a simplified manner. In the summary report for the bortezomib PAS, Janssen-Cilag provided an estimate of the expected absolute rebate as "*at least 15% of the total cost of Velcade used in the NHS in England and Wales*".[1] Celgene therefore translated this into a 15% cost reduction for all patients receiving bortezomib retreatment second line in their model.

A 2009 report into the uptake of PAS (Williams 2009) stated that 55% of respondents could confirm that they had received all refunds for bortezomib.[2] Celgene used the value of 55% in their model to adjust the 15% absolute discount to a value of 8.3% to allow for patients for whom the PAS was not received.

To explore the scenario where all bortezomib PAS refunds are received, we readjust the absolute cost discount for all patients receiving second line bortezomib to 15% and apply this to the PenTAG base case. This can be achieved by changing Cell D129 on the "*Controls*" worksheet from 55% to 100% (Excel workbook "*LEN MM 2016 model PenTAG base case SUBMIT TO NICE 7 March 2016.xlsb"*)

The results for this are presented in Table 2.

With the total absolute discount for bortezomib second line assumed to be 15%, the discounted costs in the bortezomib arm reduce by ~£800. QALY gains remain unchanged. This increases the ICER of the lenalidomide arm versus the bortezomib arm from £43,654 per QALY gained to £44,322 per QALY gained.

Table 2. Impact of bortezomib PAS on the PenTAG base case

	PenTAG 2016 base case (BOR PAS discount 8.3%)			**PenTAG 2016 base case, BOR PAS discount increased to 15%**		
Arm	**BOR**	**LEN**	**LEN vs. BOR**	**BOR**	**LEN**	**LEN vs. BOR**
Total discounted cost second line treatment (£)	11,027	61,856	50,830	10,216	61,856	51,640
Total discounted costs (£)	22,427	75,388	52,961	21,617	75,388	53,771
Undiscounted LYs	3.59	5.87	2.28	3.59	5.87	2.28
Discounted QALYs	2.33	3.55	1.21	2.33	3.55	1.21
ICER (£/QALY)			43,654			44,322

Key: BOR, bortezomib; HR, hazard ratio; ICER, incremental cost-effectiveness ration; LEN, lenalidomide; LYs, life years; OS, overall survival; PAS, patient access scheme; QALY, quality adjusted life year

3 Additional analyses: effectiveness of bortezomib and lenalidomide assumed equal

The evidence which informs the effectiveness estimates of bortezomib versus lenalidomide is of poor quality. As such, it is difficult to ascertain whether or not lenalidomide is truly more effective than bortezomib. In their base case, Celgene assume that the progression free survival hazard ratio (PFS HR) for lenalidomide versus bortezomib is 1.34, resulting in an additional 0.98 undiscounted progression free years. The overall survival hazard ratio (OS HR) is estimated as 1.71, resulting in an additional 2.18 undiscounted life years for the lenalidomide arm compared with the bortezomib arm.

We present scenarios where the effectiveness of bortezomib and lenalidomide are assumed equal; i.e., PFS and OS for bortezomib is the same as PFS and OS for lenalidomide.

We note that for all scenarios, the QALY gains differ across arms, despite the estimated life years being equal. The difference in QALYs between the lenalidomide and bortezomib arms arises from the difference in utility decrements associated with adverse events. In the lenalidomide arm the impact of adverse events are represented with a utility decrement of 0.008, applied in every model cycle for patients on second line treatment. In the bortezomib arm, this decrement is 0.033.

3.1 Applying equal effectiveness for bortezomib and lenalidomide to the Celgene base case

Table 3 demonstrates the impact of assuming bortezomib and lenalidomide have equal effectiveness upon the Celgene 2016 base case. We have provided both the scenario where the PFS and OS HRs for lenalidomide versus bortezomib should equal 1 (i.e. cells G23:24 in the "*Selected Progression & Survival*" Excel worksheet are set to 1), and the scenario where the life years are equal in the lenalidomide and bortezomib arms.

We expect these two scenarios to produce identical results, but Table 3 shows that this is not the case. Setting the PFS and OS HRs to equal 1 in the Celgene model does not result in equal effectiveness estimates for lenalidomide and bortezomib. This scenario is therefore presented only to highlight the impact of Celgene's additional adjustment to OS in the comparator arm of their model, further details of which can be found in our submission from March 7th 2016.

When equal effectiveness of the treatments is correctly implemented (i.e., the undiscounted life years for lenalidomide and bortezomib are equal), the ICER for lenalidomide versus bortezomib increases greatly from ~£20,000 per QALY gained to ~£183,000 per QALY gained.

In this scenario, the driver of the difference in QALYs between arms is the impact of adverse events upon the quality of life in each arm. The main driver of the costs in the bortezomib arm is the cost of subsequent treatments. The discounted cost of third line treatment in the bortezomib arm is £31,736, compared to £385 in the lenalidomide arm. The next largest cost in the bortezomib arm, the acquisition cost for second line bortezomib, totals £19,814 (discounted). The largest cost in the lenalidomide arm is the acquisition cost for second line lenalidomide (£61,856, discounted).

Table 3. Impact of equal effectiveness for BOR and LEN applied to the Celgene base case

	Celgene 2016 base case			Celgene 2016 base case, LEN vs. BOR PFS HR=1, OS HR=1[a]			Celgene 2016 base case, LEN vs. BOR PFS HR=1, OS HR=1, LYs set equal		
Arm	**BOR**	**LEN**	**LEN vs. BOR**	**BOR**	**LEN**	**LEN vs. BOR**	**BOR**	**LEN**	**LEN vs. BOR**
Total discounted costs (£)	58,708	81,887	23,180	61,730	81,887	20,157	77,642	81,887	4,246
Undiscounted PFS	1.97	2.96	0.98	2.94	2.96	0.01	2.96	2.96	0.00
Undiscounted total LYs	3.69	5.87	2.18	4.37	5.87	1.50	5.87	5.87	0.00
Discounted QALYs	2.37	3.55	1.17	2.83	3.55	0.72	3.52	3.55	0.02[b]
ICER (£/QALY)			19,781			28,126			182,773

Notes: a Due to the additional adjustment in the Markov model for the comparator arm in the Celgene base case, setting the OSHR=1 does not give equal LYs in both arms of the model. b When the life years are equal in both arms, QALYs still differ due to differences in adverse event utility decrement between arms. If utility decrements are assumed equal, bortezomib dominates lenalidomide.

Key: BOR, bortezomib; HR, hazard ratio; ICER, incremental cost-effectiveness ration; LEN, lenalidomide; LYs, life years; OS, overall survival; PAS, patient access scheme; PFS, progression free survival; QALY, quality adjusted life year

3.2 Applying equal effectiveness for bortezomib and lenalidomide to the PenTAG base case

Table 4 demonstrates the impact of assuming equal effectiveness for bortezomib and lenalidomide upon the PenTAG base case. Again the QALYs differ between arms due to the different adverse event utility decrements for lenalidomide compared with bortezomib.

Table 4. Impact of equal effectiveness for BOR and LEN on the PenTAG base case

	PenTAG 2016 base case			PenTAG 2016 base case, LEN vs. BOR PFS HR=1, OS HR=1		
Arm	**BOR**	**LEN**	**LEN vs. BOR**	**BOR**	**LEN**	**LEN vs. BOR**
Total discounted costs (£)	22,427	75,388	52,961	26,148	75,388	49,240
Undiscounted PFS	1.97	2.96	0.98	2.96	2.96	0.00
Undiscounted total LYs	3.59	5.87	2.28	5.87	5.87	0.00
Discounted QALYs	2.33	3.55	1.21	3.52	3.55	0.02
ICER (£/QALY)			43,654			2,119,810

Notes: QALY difference is driven by differing adverse event utility decrements between arms. If these decrements are also set as equal, bortezomib dominates lenalidomide.

Key: BOR, bortezomib; HR, hazard ratio; ICER, incremental cost-effectiveness ration; LEN, lenalidomide; LYs, life years; OS, overall survival; PAS, patient access scheme; PFS, progression free survival; QALY, quality adjusted life year

In contrast to Celgene's base case, the OS for bortezomib in the PenTAG base case is calculated using only the hazard ratio compared to lenalidomide. Therefore, setting the PFS and OS HRs equal to 1 gives equal numbers of undiscounted life years in the two arms, as expected.

In this scenario, the benefit of the lenalidomide arm versus bortezomib is greatly reduced (0.02 QALYs gained, compared to 1.21 QALYs gained in the PenTAG base case), and the cost of the bortezomib arm is increased by less than £4,000. The ICER for lenalidomide versus bortezomib increases significantly to over £2,000,000 per QALY gained.

3.3 Applying equal effectiveness for bortezomib and lenalidomide to the PenTAG base case, bortezomib PAS equal to 15% cost discount

This scenario combines the assumption of equal effectiveness of bortezomib and lenalidomide with the results of the scenario from Section 2, pp.6-7. Similar ICERs were produced whether the bortezomib discount was 8.3% or 15%, as reported in Section 2. Therefore, the results for this scenario, presented in Table 5, are similar to those reported in Table 4. Again the ICER for lenalidomide versus bortezomib is increased to over £2,000,000 per QALY gained.

Table 5. Impact of equal effectiveness for bortezomib and lenalidomide on the PenTAG base case, with BOR PAS equal to 15% discount in cost

	PenTAG 2016 base case, BOR PAS discount 15%			**PenTAG 2016 base case, BOR PAS discount 15%, LEN vs. BOR PFS HR=1, OS HR=1**		
Arm	**BOR**	**LEN**	**LEN vs. BOR**	**BOR**	**LEN**	**LEN vs. BOR**
Total discounted costs (£)	21,617	75,388	53,771	25,338	75,388	50,050
Undiscounted PFS	1.97	2.96	0.98	2.96	2.96	0.00
Undiscounted LYs	3.59	5.87	2.28	5.87	5.87	0.00
Discounted QALYs	2.33	3.55	1.21	3.52	3.55	0.02
ICER (£/QALY)			44,322			2,154,696

Notes: QALY difference is driven by differing adverse event utility decrements between arms. If these decrements are also set as equal, bortezomib dominates lenalidomide.

Key: BOR, bortezomib; HR, hazard ratio; ICER, incremental cost-effectiveness ration; LEN, lenalidomide; LYs, life years; OS, overall survival; PAS, patient access scheme; PFS, progression free survival; QALY, quality adjusted life year

References

1. Department of Health. Summary of VELCADE® Response Scheme. https://www.nice.org.uk/guidance/ta129/documents/department-of-health-summary-of-responder-scheme2 2007.

2. Williamson S. A Report into the Uptake of Patient Access Schemes in the NHS. http://www.bopawebsite.org/contentimages/publications/Report_into_Uptake_of_Patient_Access_Schemes_8_11_09.pdf Cancer Network Pharmacist Forum, 2009.

Lenalidomide for treating multiple myeloma after 1 prior treatment with bortezomib (part-review of TA171)

Request for additional evidence, April 2016

Dear [redacted]

Thank you for your contribution to the appraisal committee discussion of lenalidomide on the 6th April. We are writing to inform you that this appraisal has been suspended because the committee needs to see additional evidence before it can reach a decision. Accordingly, NICE requests Celgene to submit the evidence described below by the end of 17June 2016. Please upload your response to NICE Docs.

1. Validation of model

The committee noted that Celgene's model predicts a mean survival benefit of 2.2 years for lenalidomide versus bortezomib, whereas the MM-009 and MM-010 trials show a median survival benefit of 6.4 months for lenalidomide versus high-dose dexamethasone. This raised concerns about the external validity of Celgene's model.

Please provide clinical outcomes (QALYs and LYG) when using the model to compare lenalidomide with the comparator in the randomised trials (that is, high-dose dexamethasone alone), using the pooled data from the MM-009 and MM-010 trials. Please **do not adjust** for crossover from the comparator group to lenalidomide (adjustment is not needed because the use of third-line lenalidomide in the comparator arm of the trial reflects current NHS practice). The comparator arm should be modelled in the same way as the lenalidomide arm using multi-state modelling. Please present the results in a table showing the discounted and undiscounted estimates of QALYs and LYG for: A) the comparator arm of the trials; B) Celgene's base-case predictions for chemotherapy; C) Celgene's base-case predictions for bortezomib retreatment; D) Celgene's base-case predictions for lenalidomide.

The committee views this analysis as an important aspect of model validation. It accepts that placebo plus dexamethasone is not an appropriate comparator from a clinical or patient point of view. The committee seeks reassurance that the predicted outcomes for comparator treatments, based on the current model using indirect comparisons, are not worse than those using the MM trial comparator arms (**not adjusted** for crossover).

2. Subsequent treatments

Please provide full details of how the comparator arms of the model are adjusted to reflect the **costs and benefits** of third- and fourth-line treatments. Please provide full details of all calculations. In addition, kindly include a table or figure showing which third- and fourth-line treatments are included in the model and the assumptions about use of each (i.e. an updated version of figure 22 from the main company submission).

The rationale behind this question is that the ERG finds the impact of the adjustment difficult to understand. For instance, in Celgene's base case, adjusting for third line lenalidomide in the comparator arm modestly increases overall survival for the bortezomib arm (3.59 versus 3.69 LYs), but the costs increase substantially from £33,497 to £60,555. If PFS and OS second-line hazard ratios are set to 1, the third line lenalidomide adjustment reduces the overall survival with bortezomib (4.37 LYs compared to 5.87 without adjustment). Furthermore, when PFS and OS second-line hazard ratios are set to 1, there appear to be fewer life years gained in progressed disease for the bortezomib arm compared with the base case (1.43 compared to 1.71 LYs in the base case), but higher costs (£32,549 compared to £31,918 in the base case).

3. New clinical evidence

Professor Kwee Yong made the committee aware of new evidence on bortezomib retreatment from UCLH. Although the new study shares the shortcomings of the existing evidence insofar as it was a small record review and comes from a single centre, it does include recent NHS patients. If these data are available, please submit this evidence to NICE.

4. Explanation of the 2016 model

Please provide a clear explanation of the multistate model in language suitable for non-specialists. Please specify the methods for:

- predicting outcomes with lenalidomide, including how these were extrapolated and adjusted to reflect second-line patients

- calculating hazard ratios, including adjustments for differences in trial populations.

In addition, please comment on whether the model satisfies the underlying statistical assumptions for using ratios of medians to calculate hazard ratios.

5. Scenario assuming no survival benefit after stopping treatment

The committee requests a scenario analysis using a hazard ratio of 1 after stopping treatment, because there is no evidence of an ongoing survival benefit beyond that time. Please present this analysis for both comparators (bortezomib retreatment and melphalan).

If Celgene's response contains confidential information, please:

- Underline and highlight all confidential information: 'commercial in confidence' in turquoise and 'academic in confidence' in yellow.
- Provide 2 versions, one with confidential information highlighted and one with confidential information redacted ([redacted]).
- Provide a completed checklist of confidential information (attached).

Please also provide an updated version of your economic model along with a list of the changes made to the version submitted in February 2016.

The appraisal has been suspended to allow time to prepare the new evidence. We will be in touch with more information about timelines. I look forward to receiving your new evidence by 17 June 2016.

Best wishes,

Dr Melinda Goodall
Associate Director – Committee B

Celgene Ltd
1 Longwalk Road
Stockley Park, Uxbridge
UB11 1DB, UK
Telephone +44 (0)208 831 8620
Facsimile +44 (0)208 831 8301

Dr Melinda Goodall,
Associate Director – Committee B,
National Institute for Health and Care Excellence
10 Spring Gardens | London SW1A 2BU | United Kingdom
Tel: 44 (0)20 7045 2248 | Fax: 44 (0)20 7061 9830

Dear Dr Goodall,

RE: **Multiple myeloma - lenalidomide (post bortezomib) (part review TA171) [ID667]**

Please see below the additional information requested by the committee.

Celgene would like to thank the Committee for these requests and especially for question 1 which has given us a much greater understanding of the validity of the results produced by the model. This allows us to provide a meaningful and valid comparison with which a decision can be made for this part review.

We would also like to highlight the comments of the patient and clinical experts at the last committee meeting who stated that as bortezomib re-treatment is no longer funded via the CDF and NHS England have informed hospitals that they will not be funding re-treatment via TA129, the comparison with MP is the most-appropriate for decision making.

Yours Sincerely,

[redacted]

Lenalidomide for treating multiple myeloma after 1 prior treatment with bortezomib (part-review of TA171)

Request for additional evidence, April 2016

1. Validation of model

The committee noted that Celgene's model predicts a mean survival benefit of 2.2 years for lenalidomide versus bortezomib, whereas the MM-009 and MM-010 trials show a median survival benefit of 6.4 months for lenalidomide versus high-dose dexamethasone. This raised concerns about the external validity of Celgene's model.

Please provide clinical outcomes (QALYs and LYG) when using the model to compare lenalidomide with the comparator in the randomised trials (that is, high-dose dexamethasone alone), using the pooled data from the MM-009 and MM-010 trials. Please **do not adjust** for crossover from the comparator group to lenalidomide (adjustment is not needed because the use of third-line lenalidomide in the comparator arm of the trial reflects current NHS practice). The comparator arm should be modelled in the same way as the lenalidomide arm using multi-state modelling. Please present the results in a table showing the discounted and undiscounted estimates of QALYs and LYG for: A) the comparator arm of the trials; B) Celgene's base-case predictions for chemotherapy; C) Celgene's base-case predictions for bortezomib retreatment; D) Celgene's base-case predictions for lenalidomide.

The committee views this analysis as an important aspect of model validation. It accepts that placebo plus dexamethasone is not an appropriate comparator from a clinical or patient point of view. The committee seeks reassurance that the predicted outcomes for comparator treatments, based on the current model using indirect comparisons, are not worse than those using the MM trial comparator arms (**not adjusted** for crossover).

Clinical outcome comparison of previous comparative effectiveness estimates compared to DEX outcomes from the clinical trial

Methods

As requested, DEX outcomes were projected using the same methodology as the LEN arm, please see response to question 4 for further clarification. The cut-offs selected for DEX were 168 & 616 (days). Covariates included within both analyses were kept the same.

Results

Table 1 provides a comparison of clinical outcomes for all treatments. For transparency a breakdown of Life Years (LYs) both pre and post progression has been provided.

Table 1: Comparison of clinical outcomes for all treatments

Treatment	Total LYs	Total PFS LYs	Total PPS LYs	Total QALYs
Dexamethasone (DEX)	3.97	0.99	2.98	2.39
Melphalan + prednisolone (MP)	1.14	0.43	0.72	0.78
Bortezomib retreatment (BORT)	2.70	1.87	0.83	1.86
Lenalidomide + dexamethasone (LEN+DEX)	5.87	2.96	2.91	3.55

Conclusion

When looking at LYs, there are evidently issues with the validity of the comparative effectiveness compared to MP and BORT retreatment. These issues stem from the necessity of comparing to poor quality non-randomised evidence which in some cases included only reported medians and from the inability to properly adjust the evidence available to account for the benefit that might be expected from third-line use of LEN (see answer to question 2), again due to lack of reported information for the comparators.

However, when LYs are broken down into pre-progression and post-progression, it can be seen that that the differences lie in the post-progression gain which suggests that the efforts made to include subsequent LEN at third-line for the comparator arms have not accurately reflected the true benefits as seen in the trial. When considering Progression-free LYs, the model returns the ordering of efficacy which would be expected based on our clinical knowledge of the treatments; LEN+DEX, BORT retreatment, DEX, MP. Based upon this information, we believe that the comparison to MP at least should be considered further and that the high-quality RCT data from MM-009 and MM-010 for DEX can be used to estimate a valid result for MP and

allow a comparison to LEN+DEX. We provide evidence below of the equivalence of DEX and MP.

We then present ICERs when the post-progression survival (PPS) for MP is set equal to that of DEX and also when the PFS and PPS for MP is set equal to DEX; so that the committee has the full information available on which to make a decision and to explore the uncertainty around the use of the OS HR for PFS (which was used as no PFS was reported) from the Petrucci 1989[1] paper for MP.

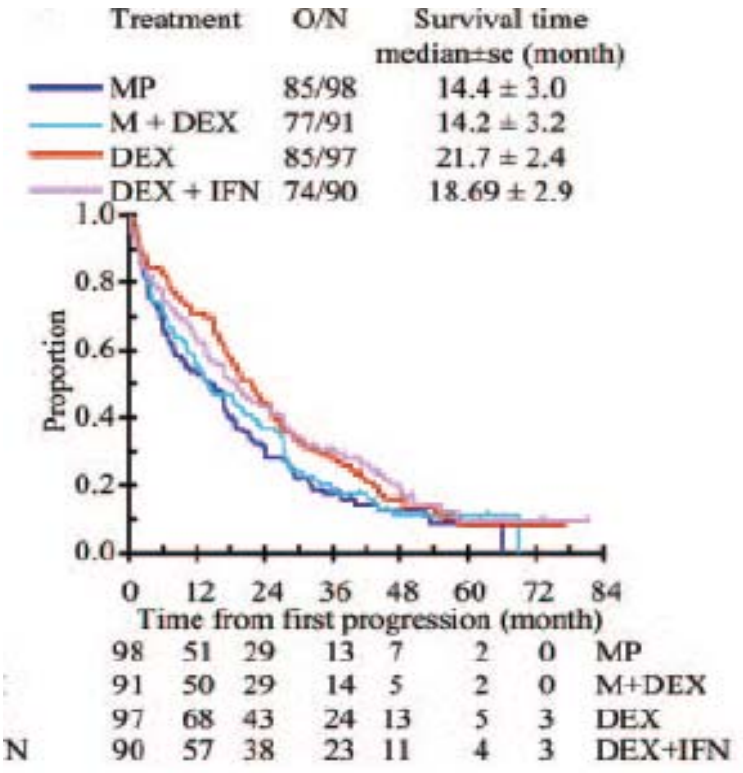
Evidence for relative effectiveness of MP and dexamethasone

MP has been in use since the late 1960s and as has been demonstrated within this review there is little high quality evidence demonstrating the effectiveness of MP for the treatment of multiple myeloma at second line.

Within the MM-009 and MM-010 clinical trials DEX was adopted as the control arm as it represented a standard anti-myeloma therapy for the treatment of subjects with relapsed or refractory disease at the time the trials were initiated[2,3].

In 2006 the IFM group published data on DEX vs. MP in transplant ineligible patients[4] which demonstrated no significant difference in OS between the 4 regimens studied: melphalan-prednisone, dexamethasone alone, melphalan dexamethasone, and dexamethasone–interferon alpha. Whilst this study is in first-line patients this represents the only study available comparing outcomes in patients receiving MP vs DEX.

Figure 1: Facon 2006 OS for MP vs DEX



It should also be noted that within the original TA171 review, outcomes from the MRC database for patients treated with conventional chemotherapy were accepted as a suitable proxy for outcomes for patients receiving DEX in TA171.

The available evidence indicates that DEX is equivalent to MP. The advantages of using the evidence from the DEX arm in the MM-009 and MM-010 trials are numerous:

- Maintains randomization allowing for a statistically robust comparison to be made.
- Allows the impact of receipt of LEN at 3rd line on OS to be accurately measured as patients were allowed to receive LEN post progression.
- Removes uncertainty surrounding comparator effectiveness as outcomes for PFS and OS are available from a mature dataset with patient level data available, meaning that comparison using medians alone is no longer required.

Cost-effectiveness of LEN vs MP assuming equivalence to DEX outcomes

Methods

In order to produce a more statistically robust cost-effectiveness analysis as detailed above, the LY and QALY projections generated for DEX using the MSM method were used to produce a full cost-effectiveness analysis for MP vs LEN+DEX.

The number of patients who crossed over within the DEX arm of the clinical trials to receive subsequent LEN was 167 out of the original 351. The cost of LEN is therefore included within the model for exactly this proportion of patients and as stated in the question, adjustment is not needed because the use of third-line lenalidomide in the comparator arm of the trial reflects current NHS practice.

Two scenarios have been considered for MP in order to explore model sensitivity around assumptions for PFS given that the information previously used for MP PFS was drawn from a crude hazard ratio produced using a small non RCT dataset containing only OS data:

- MP is assumed for have the same PFS as submitted previously; DEX information is used only to inform PPS.
- MP is assumed to have an identical PFS and PPS to DEX.

We consider the second scenario to be the more valid of the two given that this retains full statistical validity (entirely within trial comparison). This is also the more conservative of the two scenarios as DEX outcomes within MM-009 and MM-010 are more favourable than those originally estimated for MP.

Results

As can be seen in the tables below in both scenarios, LEN+DEX is cost-effective vs MP based upon a willingness to pay threshold of £30,000 per QALY.

Table 2: MP with pre-progression profile MP, post-progression profile of DEX

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER
MP	£58,573	4.32	2.44	-	-	-	-
LEN + DEX	£84,568	5.87	3.55	£25,995	1.55	1.11	£23,462

Table 3: MP with pre and post progression profile of DEX

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER
MP	£65,848	4.86	2.76	-	-	-	-
LEN + DEX	£84,568	5.87	3.55	£18,720	1.01	0.79	£23,810

ICERs within both scenarios remain very similar, this is due mainly to the increase in LYs (12%) for MP, which effects both the costs and the QALYs by the same magnitude.

The previously presented results using the digitized curves from the Petrucci 1989[1] paper instead of the medians, produced similar results (MP LYs, 3.15; QALYs, 1.88; ICER, £23,618) to those seen here when using the DEX data which suggests these results are reliable and interpretable.

2. Subsequent treatments

Please provide full details of how the comparator arms of the model are adjusted to reflect the **costs and benefits** of third- and fourth-line treatments. Please provide full details of all calculations

Information is provided below separately for third and fourth line as these are modelled differently (third line accounts for effectiveness and cost; fourth line only accounts for cost).

Third line

Calculation of third-line subsequent therapy costs and effectiveness is split into two parts:

- Patients receiving subsequent LEN
- Patients not receiving subsequent LEN

It is assumed that the impact on effectiveness of all 3rd line therapies other than LEN has already been captured within the OS curves in the various source publications.

Detailed calculation information is provided below using the comparator patient flow sheet as an example.

Adjustment of OS to account for third-line treatments

OS transitions were split into two parts when subsequent treatment with LEN was included within the model calculations:

- Initial OS transition (in Column P, used for second-line survival)
- OS from third line onwards (in Column R) was set equal to Column P when the treatment received was not LEN; when receiving LEN, the OS transition was taken from survival for patients receiving third-line LEN (OS 3L sheet in Column AL)
    - In the latest version of the model, these third-line transitions use the MSM model structure, and are derived in the same way as the second-line LEN transitions
    - This method is an imperfect proxy for reality as it is not possible to derive pre-progression survival information from the sources available for comparator effectiveness –OS is used as a proxy for the transitions

for death prior to third-line treatment. The assumption is that pre- and post-progression survival hazards are equivalent within the comparator dataset

- The final OS transition in Column N is calculated using the second- and third-line transitions and the proportion of patients at second line versus third line in the previous cycle. This transition is used to create a composite OS curve in Column M.

Calculation of costs

1. The weighted average cost of third-line treatment for all treatments other than LEN is calculated in Row 57 of the 'Post Progression' sheet based upon data from HMRN, and the costs for each treatment detailed in Table 55 of the original submission. Dosing regimens used can be found in Rows 106 to 117 in the costs sheet. Both drug and administration costs are included within the calculation.
2. The proportion of patients requiring third-line therapy is calculated within the patient flow sheet:
    a. Column AL tracks the proportion of patients expected to enter third-line therapy each cycle assuming that death happens equally across all health states post completion of second-line therapy
        i. This is calculated using the half cycle correction OS transition for third-line treatment stored in Column S
    b. This figure is taken into Column DD
    c. In Column DG, we then use the information on the duration of third-line therapy for patients not receiving LEN (taken from HMRN data as 4 cycles) to calculate how many cycles to apply cost for once patients start treatment – the OS transition is used to account for the proportion of patients who die each cycle having started treatment, assuming that death happens equally across all health states post completion of second-line therapy
        i. This is calculated using the proportion of patients entering treatment each cycle from Column DD and the half cycle correction OS transition for third-line treatment stored in Column S;

ii. Patients from the previous cycles are included for up to 4 cycles of treatment

d. In Column DE, we calculate the proportion of patients who would receive the costs of third-line LEN if this treatment is selected to be received assuming that death happens equally across all health states post completion of second-line therapy
    i. This is calculated using the proportion of patients entering treatment each cycle from Column DD and the half cycle correction OS transition for third-line treatment stored in Column S
    ii. Patients from the previous cycles are included for up to 12 cycles of treatment
        1. This figure is taken from Cell V15 in the TTF 3L sheet, which sums the average number of cycles of treatment received according to the fitted curves for TTF 3L
e. In Column DF, we calculate the proportion of patients who would receive the costs of third-line LEN if the PAS is applied again assuming that death happens equally across all health states post completion of second-line therapy
    i. This is calculated using the proportion of patients entering treatment each cycle from Column DD and the half cycle correction OS transition for third-line treatment stored in Column S
    ii. Patients from the previous cycles are included up to [redacted] of treatment
        1. This figure is taken from Cell V16 in the TTF 3L sheet, which sums the average number of cycles of treatment received according to the fitted curves for TTF 3L before the PAS would kick in at 26 cycles of treatment

3. For patients not receiving LEN, cost is calculated within Columns DJ and DK by multiplying the proportion of patients receiving third-line therapy by the weighted average cost of therapy (drug and administration costs)
4. For patients receiving LEN, cost is calculated within Columns DH, DI and DV based upon:

a. The proportion of patients receiving LEN at third line
b. The per cycle cost of LEN at third line (drug cost, GCSF and AEs), which is assumed the same as at second line

Fourth line

Subsequent treatment is included only as a cost impact. No impact on effectiveness is modelled.

An example of a calculation within the patient flow sheet for LEN is described below:

1. The weighted average cost of fourth-line treatment is calculated in Cell F78 of the 'Post Progression' sheet based upon data from HMRN, and the costs for each treatment detailed in Table 55 of the original submission. Dosing regimens used can be found in Rows 106 to 117 in the costs sheet
2. The proportion of patients requiring fourth-line therapy is calculated within the patient flow sheet
    a. Column AL tracks the proportion of patients expected to enter third-line therapy each cycle, assuming that death happens equally across all health states post completion of third-line therapy
        i. This is calculated as sum of (pre-progression third-line and post-progression health states this cycle minus sum of pre-progression third-line and post-progression health states in the previous cycle) * (1 minus proportion transitioning to death in that cycle)
    b. Column CV denotes when these patients would be expected to start fourth-line treatment – in this case, 4 cycles following the start of third line based upon HMRN data for the duration of third-line treatment
    c. Column CY simply references Column AL (to allow later lookup)
    d. In Column DG, we use this information to calculate the number of patients newly finishing third-line treatment each model cycle using a lookup based upon Columns CV and CY
    e. In Column DH, we then use the information on the duration of fourth-line therapy (again taken from HMRN data) to calculate how many cycles to apply cost for once patients start treatment – again, the OS transition is used to account for the proportion of patients who die each cycle having started treatment

3. Cost is calculated within Column DI by multiplying the proportion of patients receiving fourth-line therapy by the weighted average cost of therapy

Additional information on how the above fits with previous ERG scenario analysis

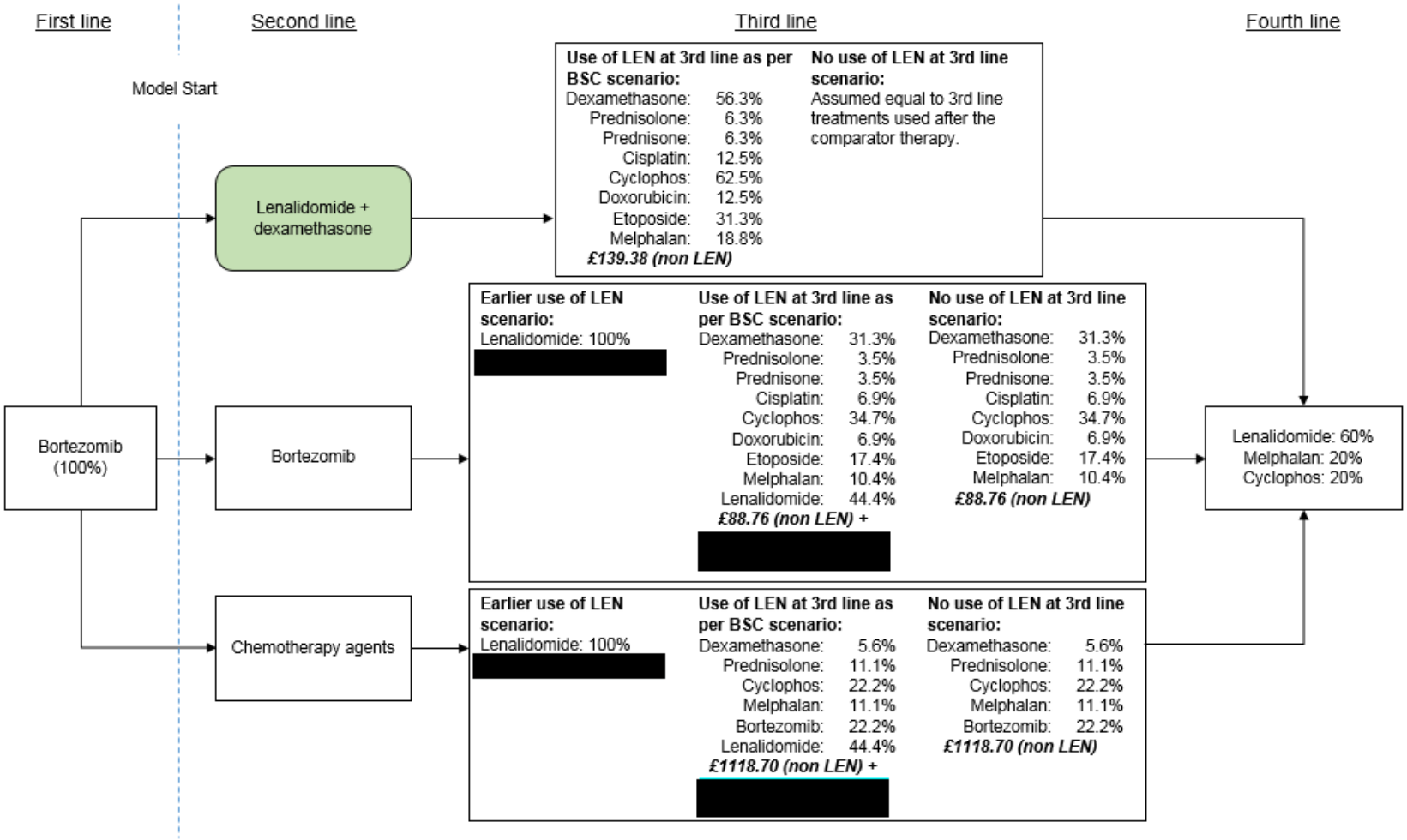
- Within ERG scenario analysis 3 (Comparator OS estimated from hazard ratio only), the adjustment to the OS transitions detailed for third-line LEN is removed from the calculation
- Within ERG scenario analysis 1, the cost impact of all third- and fourth-line treatments is removed from the calculations (but not the adjustment to the OS transition)

ERG scenario analyses 1 and 3 combined is equivalent to:

- Removing subsequent treatment costs and effects of subsequent LEN OS. This is a technically accurate calculation, but not in line with the current treatment pathway; and
- Ignoring the costs of subsequent therapies intrinsic to the OS curves used to estimate effectiveness for both the comparator and LEN+DEX within the calculations. This is not technically correct, particularly as a major benefit of including LEN within the treatment pathway earlier is to reduce the requirement for use of subsequent treatment.

In addition, kindly include a table or figure showing which third- and fourth-line treatments are included in the model and the assumptions about use of each (i.e. an updated version of figure 22 from the main company submission).

Figure 2 – Replication of figure 22 from main submission



The rationale behind this question is that the ERG finds the impact of the adjustment difficult to understand. For instance, in Celgene's base case, adjusting for third line lenalidomide in the comparator arm modestly increases overall survival for the bortezomib arm (3.59 versus 3.69 LYs), but the costs increase substantially from £33,497 to £60,555. If PFS and OS second-line hazard ratios are set to 1, the third line lenalidomide adjustment reduces the overall survival with bortezomib (4.37 LYs compared to 5.87 without adjustment). Furthermore, when PFS and OS second-line hazard ratios are set to 1, there appear to be fewer life years gained in progressed disease for the bortezomib arm compared with the base case (1.43 compared to 1.71 LYs in the base case), but higher costs (£32,549 compared to £31,918 in the base case).

Having thoroughly reviewed this, the issue in implementing the impact of third-line LEN into the economic model, which is causing spurious results such as that seen above, is the requirement to make assumptions around the pre- and post-progression mortality within the OS data available for the comparator. The model currently assumes that pre- and post-progression mortality are equal (which is necessary as only OS transition data are available from the publications). This means that it is possible to produce results where the third-line LEN transition is higher (i.e. assuming a greater mortality) than the OS transition from the comparator publication (which is used for both pre- and post-progression survival when third-line LEN is not included in the calculations).

This information should be considered within the following light:

- Lack of information for comparators does not allow for a more accurate modelling structure.
- This problem does not exist for the comparison versus MP, where the third-line LEN outcomes are consistently better than those with MP (for which a published KM was actually available for OS).

3. New clinical evidence

Professor Kwee Yong made the committee aware of new evidence on bortezomib retreatment from UCLH. Although the new study shares the shortcomings of the existing evidence insofar as it was a small record review and comes from a single centre, it does include recent NHS patients. If these data are available, please submit this evidence to NICE.

These data have now been published[5]. The baseline characteristics and clinical results are presented below. However, given that funding has been removed for BORT-retreatment during the course of this appraisal, we do not believe the BORT-retreatment comparison should be considered further.

Table 4: Patient demographics

Characteristic	
Median age, years (range)	63.6 (38.7-97.7)
Gender, n (%)	
Female	11 (48)
Male	12 (52)
Median follow-up, years (range)	3.8 (1.0 -7.3)
Myeloma type, n (%)	
IgG	10 (43)
IgA	8 (35)
Light chain only	4 (17)
Plasma cell leukaemia	1(4)
ISS, n (%)	
1	11 (48)
2	6 (26)
3	6 (26)
Cytogenetic risk, n (%)	
Standard	9 (35)
Adverse	6 (22)
Not known	8 (43)
First treatment regimen, n (%)	
PAD	10 (43)
VCD	5 (22)
Bortezomib Dexamethasone	5 (22)
MPV	3 (13)
Second treatment regimen, n (%)	
VTD-based	13 (57)
VCD	4 (17)
Bortezomib, bendamustine, dexamethasone	2 (9)
Bortezomib, dexamethasone	4 (17)

Adverse risk cytogenetics: t(4;14), t(14;16), p53 del (>50%)
Standard risk cytogenetics: no high risk lesion
PAD, bortezomib, doxorubicin, dexamethasone. VCD, bortezomib, cyclophosphamide, dexamethasone. MPV, melphalan, prednisolone, bortezomib. VTD, bortezomib, thalidomide, dexamethasone.

Table 5: Response to bortezomib

	First-line bortezomib	Bortezomib re-treatment
Number of cycles, median (range)	5 (4–8)	5 (1–8)
Response, *n* (%)		
Overall response	23 (100)	20 (87)
Complete response	6 (26)	3 (13)
Very good partial response	14 (61)	8 (35)
Partial response	3 (13)	9 (39)
<partial response	0	3 (13)
Time to best response, months; median (range)	3·5 (0·7–9·0)	4·1 (0·7–15·0)
Duration of response, months; median (range)	14·9 (4·7–44·5)	11·5 (1·0–18·5)
Autologous stem cell transplantation, *n* (%)		
Yes	10 (43)	11 (48)
No	13 (57)	12 (52)

Note: only 14 patients had disease progression at time of analysis.

4. Explanation of the 2016 model

Please provide a clear explanation of the multistate model in language suitable for non-specialists. Please specify the methods for:

- predicting outcomes with lenalidomide, including how these were extrapolated and adjusted to reflect second-line patients
- calculating hazard ratios, including adjustments for differences in trial populations.

In addition, please comment on whether the model satisfies the underlying statistical assumptions for using ratios of medians to calculate hazard ratios.

A more technical explanation is presented in Appendix 1 below for the ERG and technical members of the committee. For non-specialists, please see the explanation here:

Model Structure

A multi-state model describes how an individual moves between a series of states over time. Suppose an individual is in state 1 at one point in time. The next state, 2, to which the individual moves, and the time of the change, are governed by a set of transition intensities. Intensities, like hazards, represent the instantaneous risk of moving from state 1 to state 2.

Of note is that this model allows for direct movement between the pre-progression and death states without observing progression (i.e., patients can die before they progress). The previously adopted method did not explicitly model this transition; the death time would have been used as the date of event for both PFS and OS.

Model Output

The primary output of the MSM model is a transition probability matrix. This gives the probability of moving between states over a fixed time interval; the interval length was chosen as 28 days for this analysis.

For a **time-homogeneous** Markov model this matrix is independent of time e.g. the instantaneous risk of moving from state 1 (Pre-Progression) to state 2 (Progression) is constant across the full life horizon of the study. For a **time inhomogeneous**

model, transition intensities are allowed to change at a series of times common to each individual.

For a **time-inhomogeneous** Markov model, more than one transition probability matrix is produced; the number of matrices equals the number of time-points for which the transition intensities are allowed to change plus one.

Similarly, **covariate adjusted models** allow this transition intensity matrix to change according to levels of covariates. The number of matrices depends on the number of unique levels of covariate combinations used in the fitted model.

Model Fitting

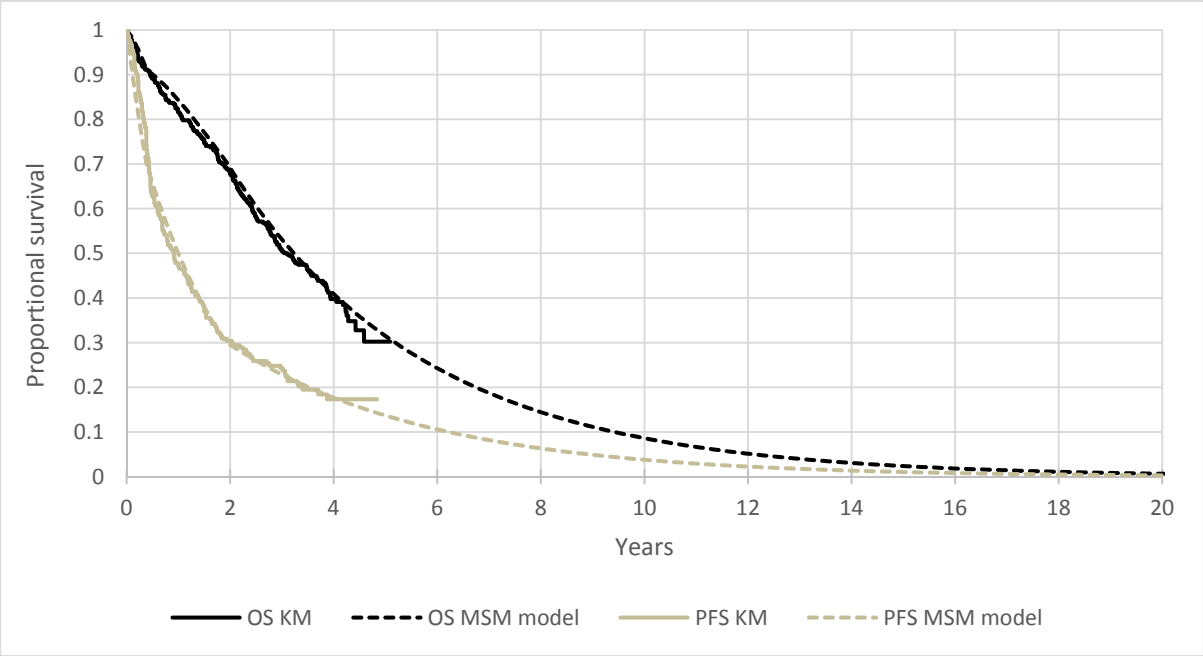
For the combined MM-009 and MM-010 data, it was found that a time inhomogeneous model significantly improved the model fit to the data. Hazard plots for OS and PFS were used to determine that there was a change in hazard at 168 and 728 days. There were therefore three transition matrices produced, one for each time interval (1-167, 168-727 and 728+ days). Further details on the model selection fitting process were provided in the original response.

Predicting Outcomes with Output from Multi-State Model

Survival curves for PFS and OS were produced by calculating the estimated percentage of patients in state 1 and 2 (OS) and the estimated percentage of patients in state 1 (PFS) at each time point.

This procedure was extended beyond the end of the trial assuming that the matrix used for the last time point (728+ days) can be used to estimate survival beyond the observed trial period as shown in Figure **3**.

Figure 3: Survival probabilities estimated from time-inhomogeneous* multi-state Markov models



Key: KM, Kaplan–Meier; OS, overall survival; PFS, progression-free survival.
Notes: *The hazard of moving between states is free to vary at 168 and 728 days in the statistical model

Covariate Adjustment

As described above, Multi-state Markov models allow for covariate adjustment by permitting the transition intensities to change according to levels of a covariate i.e. allowing the risk of moving between states to change depending on covariate values. Four covariates were found to be statistically significant in the model and were included in the final model. These were:

- prior treatments [1 vs 2 or 3],
- baseline beta-2 microglobulin [>2.5 or ≤2.5 mg/L],
- prior doxorubicin [Yes or No],
- worsening extramedullary plasmacytoma disease [Yes or no]

As such, 16 (one for each covariate combination; 2^4) transition probability matrices were outputted from the Multi-state Markov model for each time interval (1-167, 168-727 and 728+ days). The group prognosis method suggested by the ERG[6] was utilised to create an overall survival and progression-free survival curve for all patients. This involves outputting one matrix for each time point for each combination of covariates (i.e. 16 groups of patients in total).

Using the frequency of people in each combination of covariates (see supplementary table 1) survival curves are produced for each group of patients using the procedure described in the step above, a weighted survival curve is then produced simply by multiplying each survival curve by the proportion of patients within the relevant group and summing.

This procedure allows the survival curves to be calculated for OS and PFS for the group of patients of interest, e.g. second line patients. When survival curves have been estimated for second line patients only, survival curves have been weighted according to the % of patients in each of the 8 groups of second line patients.

How comparative effectiveness is estimated through calculation and comparison of adjusted survival curves

Comparative effectiveness was estimated by firstly calculating the projected OS and PFS for LEN+DEX in the group of patients most reflective of the patient characteristics in the clinical trial. This was done using the group prognosis method (as detailed above) by estimating the proportion of patients within each of the 16 groups based upon the information published within the clinical trial. Where no information was published the proportion of patients was assumed the same as within MM-009 and MM-010. This process produced the estimated survival for LEN+DEX if it was given to the same patients as received the comparator treatment within the study publication.

Following this the median OS and PFS was derived from the parametric survival curve simply by looking up the time-point at which 50% of patients had either died, or in the case of PFS either progressed or died. The median for LEN+DEX was then compared to the median for the comparator and the hazard ratio calculated as the ratio of the two (this assumes underlying constant hazards for both arms – see the comment on this below).

Applying hazard ratios calculated from medians

As described above, for the Multi-State Markov model utilised, the amount of time in each state is exponentially-distributed. As such, there are parallels between the

output of an MSM and a piecewise exponential survival curve fitted to the data in a standard way. Technically applying a hazard ratio calculated from a median is only correct when this is applied to an exponential curve (as the assumption is that we can reflect comparative mean survival using only the point estimate for survival at the 50% mark i.e. that hazards are constant across time). This means that applying a hazard ratio calculated from a median is only technically correct for a time homogenous MSM model.

A balance must therefore be struck between:

- Modelling survival for LEN+DEX accurately – given that time homogenous models (and the exponential curve previously) do not fit the data well
- The technical incorrectness of a applying a HR calculated from a median

Analysis has been provided to the Committee as part of the response to the previous documentation demonstrating the impact of applying a HR calculated from a median compared to projecting use outcomes for curves fitted independently to survival curves for the comparator therapy using Petrucci 1989[1] (MP) and Ahn 2014[7] (BORT retreatment) datasets.

This comparison (presented again below) demonstrated that survival curves produced using hazard ratios produced from medians are generally similar / more optimistic in their projections of comparator outcomes than the survival curves produced using curves fit to Kaplan Meier data.

We therefore don't consider that any bias implicit in the requirement to use medians for comparison to some data sources is in favour of LEN+DEX (in fact the opposite is likely to be the case).

Figure 4: Comparison of survival curves fit to digitised data versus hazard ratios produced by median estimates using Petrucci 1989

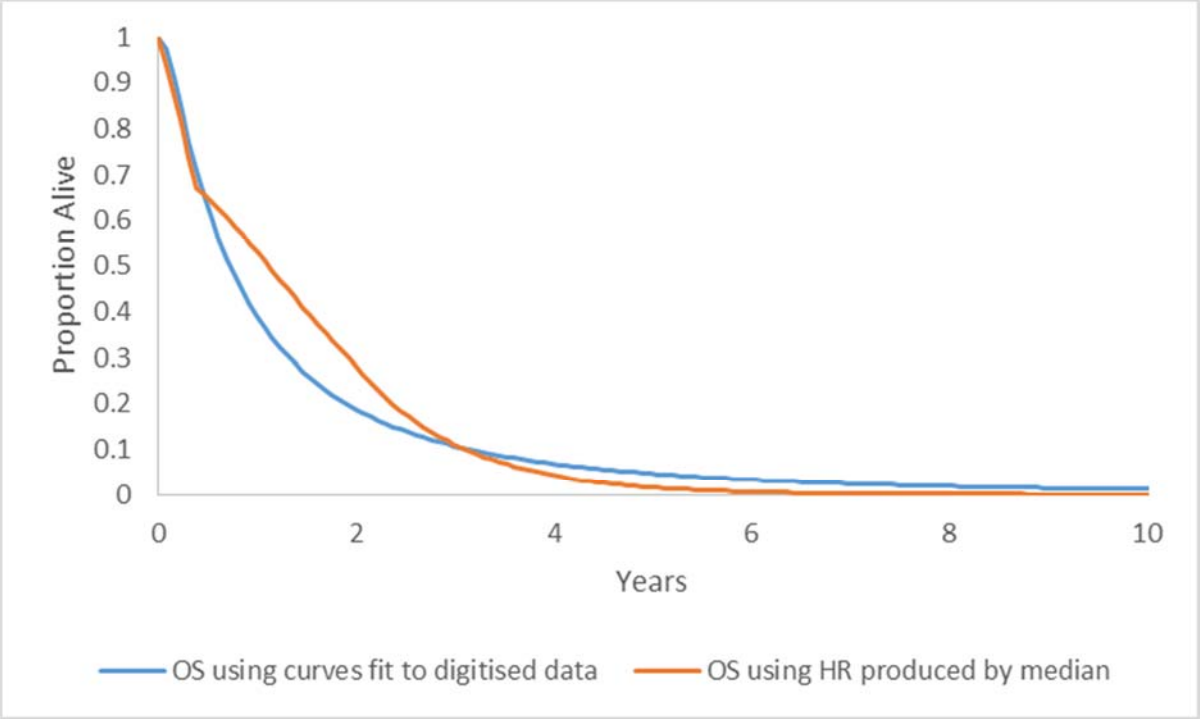
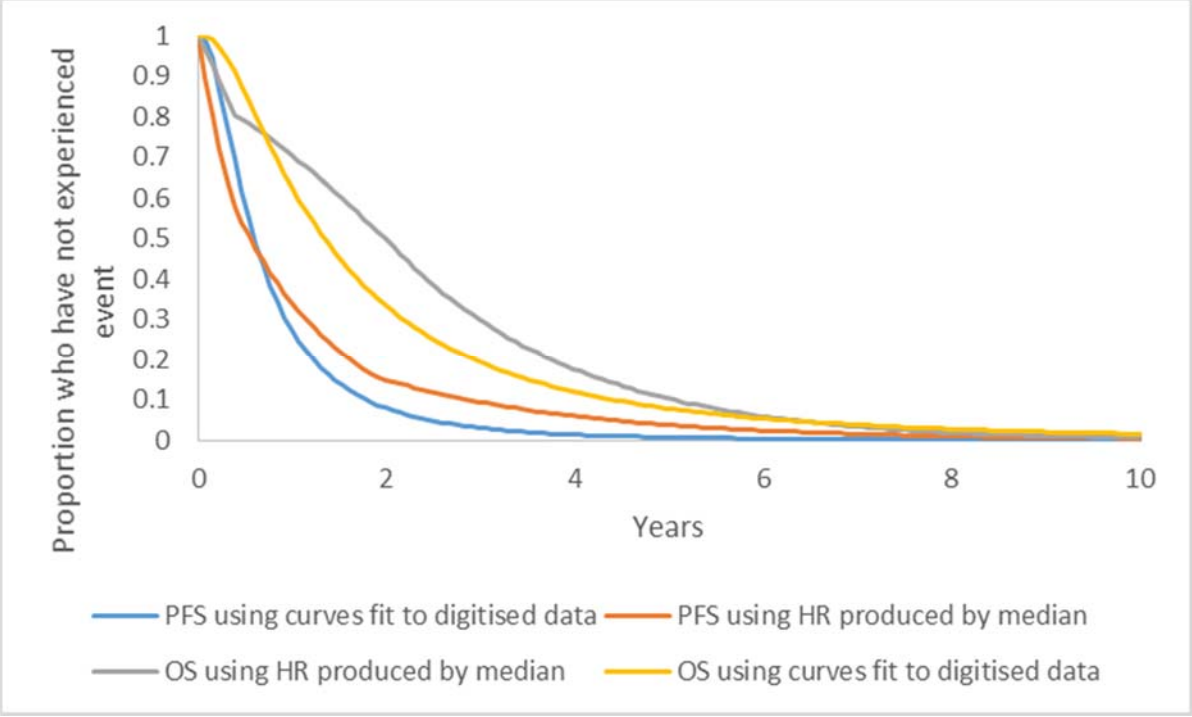


Figure 5: Comparison of survival curves fit to digitised data versus hazard ratios produced by median estimates using Ahn 2014



5. Scenario assuming no survival benefit after stopping treatment

The committee requests a scenario analysis using a hazard ratio of 1 after stopping treatment, because there is no evidence of an ongoing survival benefit beyond that time. Please present this analysis for both comparators (bortezomib retreatment and melphalan).

Celgene do not believe this is a valid comparison as the MM-009 and MM-010 data shows a PPS benefit for LEN+DEX over DEX, when the DEX arm is adjusted for cross-over to LEN.

In addition, in responders, the lack of cumulative toxicity and immunomodulatory activity of LEN resulting in enhanced antitumour effects mediated by T and natural killer (NK) cells[8] , may prolong remissions. This beneficial immunomodulatory effect is less likely with immunosuppressive regimens containing bortezomib[9] or conventional chemotherapy[10] which are limited to a fixed treatment durations due to cumulative toxicities.

Methods

Given the structure of the MSM model employed, this exact request could not be executed within the model; instead equal PPS across all arms was explored.

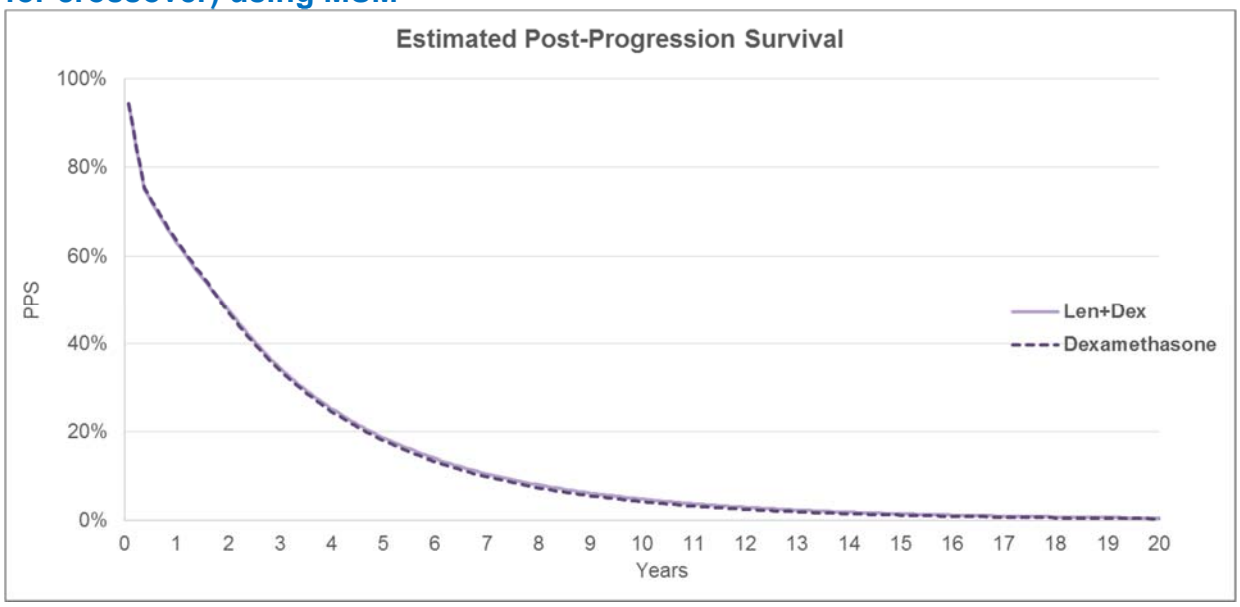
In order to explore this concept the MSM modelling used to estimate PPS for LEN was applied to the comparator arms as follows:

- comparator transitions were calculated via MSM methodology rather than using the original parametric survival curves.
- transitions for BORT/MP for progression to death and remaining in progression were then set equal to LEN.
- the probability of remaining progression free was calculated based upon the original hazard ratios relative to the LEN curve probability of remaining progression free.
- transitions for BORT /MP for progression free to progression and death were set proportionally equal to LEN after excluding those patients that remain progression free (defined by BORT /MP)

Results

Figure 6 provides a comparison of estimated PPS for LEN+DEX and DEX based upon the MSM models fitted to the LEN+DEX and DEX arms of MM-009 & MM-010. As requested in Q1 DEX outcomes have not been adjusted for crossover.

Figure 6: MM009&MM010 estimated PPS for LEN+DEX and DEX (not adjusting for crossover) using MSM



Conclusion

As can be seen, PPS for DEX is almost identical to that of the LEN+DEX arm, even though more than half of the patients on the DEX arm are receiving active treatment with LEN post progression and the same is clearly not the case on the LEN+DEX arm. This demonstrates that there is a sustained survival benefit for LEN even after the treatment is stopped, and that patients are more able to benefit from LEN if this is administered earlier, echoing the conclusions from the Stadtmauer paper[11]. This is not unexpected given the mechanism of action of LEN as LEN is known to have an immunomodulatory effect. This evidence shows that it is not appropriate to perform a scenario assuming no benefit post progression.

Nevertheless please find the results requested below. Again it should be noted that comparison to BORT retreatment is no longer relevant to clinical practice following its delisting from the CDF.

Table 6: BORT with pre-progression profile BORT, post-progression profile of LEN

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER
BORT retreatment	£64,598	4.93	2.99	-	-	-	-
LEN + DEX	£81,887	5.87	3.55	£17,289	0.94	0.56	£31,048

Table 7: MP with pre-progression profile MP, post-progression profile of LEN

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER
MP	£44,741	3.13	1.87	-	-	-	-
LEN + DEX	£81,887	5.87	3.55	£37,146	2.73	1.68	£22,172

When considering the impact of the validation exercise undertaken for question 1, Table 8 presents a scenario where pre-progression is set equal for MP to DEX and post-progression for MP is set equal to LEN+DEX.

Table 8: MP with pre-progression profile DEX, post-progression profile of LEN

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER
MP	£50,949	2.21	3.68	-	-	-	-
LEN + DEX	£81,887	5.87	3.55	£30,938	2.18	1.34	£23,152

For completeness, if the subsequent LEN costs are removed from the model and the assumption is made that there is equal efficacy post-progression of the comparators to LEN+DEX, the ICERs are as follows:

Table 9: BORT retreatment with pre-progression profile BORT retreatment, post-progression profile of LEN+DEX and no subsequent LEN costs

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER
BORT retreatment	£41,962	4.93	2.99	-	-	-	-
LEN + DEX	£81,748	5.87	3.55	£39,786	0.94	0.56	£71,449

Table 10: MP with pre-progression profile MP, post-progression profile of LEN and no subsequent LEN costs

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER
MP	£24,450	3.13	1.87	-	-	-	-
LEN + DEX	£84,568	5.87	3.55	£60,028	2.73	1.68	£35,830

Table 11: MP with pre-progression profile DEX, post-progression profile of LEN and no subsequent LEN costs

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER
MP	£33,433	2.21	3.68	-	-	-	-
LEN + DEX	£84,568	5.87	3.55	£51,135	2.18	1.34	£38,267

Overall Conclusions

The validation exercise in question 1 demonstrated that there are concerns over the estimates of efficacy for MP and BORT retreatment from the model and especially in the post-progression state where it appears that the efforts to include subsequent LEN have not fully reflected the true benefit.

During the course of this appraisal BORT re-treatment has lost funding on the NHS. When this is combined with the uncertainty identified by the analysis in question 1, the MP comparison should be considered most informative.

Data has been presented to show the equivalence of MP and DEX and a number of scenarios presented to explore the potential cost-effectiveness of LEN+DEX against MP. In all scenarios except where post-progression is assumed (without any evidence) to be equal to LEN+DEX and the cost of subsequent LEN is excluded, LEN+DEX is shown to be cost-effective against MP.

There is currently a gap in the treatment algorithm at 2nd line in the UK and we believe LEN+DEX should be funded post first-line BORT; as every possible effort has now been made to prove efficacy and cost-effectiveness in this population where the comparator data is poor.

Appendix 1 – Technical MSM Description

Introduction

Following the release of the original Appraisal Consultation Document (ACD) 2 report, the Evidence Review Group (ERG) noted that the modelled parametric overall survival (OS) curves fell below the progression-free survival (PFS) curves in the extreme ends of the extrapolated tails of the fitted curves. This implies that patients would progress after they had died and would therefore not give a realistic prediction of long-term survival. Both the committee and the ERG expressed a methodological preference that these curves should be modelled in a combined way such that the logical relationship between progression and death is retained; patients progress prior to death.

Research was conducted to identify a suitable method to produce a combined model of OS and PFS. A multi-state Markov modelling approach[12] was selected for this analysis as it allowed both:

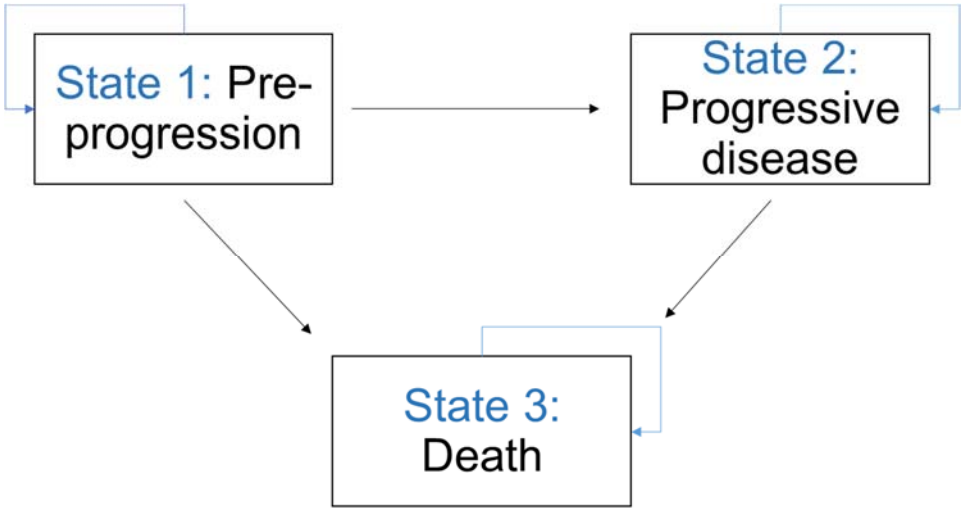
- Simultaneous modelling of the transitions between progression and death which prevents the survival curves for PFS and OS crossing.
- Covariate adjustment of survival curves which are required in order to be able to use all the data from the MM-009 and MM-010 trials to estimate outcomes in a second line patient population and to be able to produce estimates of comparative effectiveness

This methodology was implemented in the R package 'msm'[13].

Model Structure

The model structure is presented in Figure 7. Patients were defined as being in one of three states; pre-progression (state 1), progressive disease (state 2) and death (state 3). All patients start in state 1 at time of randomisation. Patients move into state 2 at time of progression (determined by central review assessment), if their progression date is censored, they will skip state 2 and move straight to state 3 at time of death. Death is a final, absorbing state that patients can move to from any other health state. Patients who were known to be alive at the end of the study were considered to be in a censored state.

Figure 7: Model structure



Note: Progressive disease is defined by the variable time to progression (central review assessment).

Multi-State Markov Models

Model Structure

A multi-state model describes how an individual moves between a series of states in continuous time. Suppose an individual is in state S(r) at time t. The next state, s, to which the individual moves, and the time of the change, are governed by a set of transition intensities; q_{rs}. Intensities, like hazards, represent the instantaneous risk of moving from state r to state s. e.g. q_{12} is the instantaneous risk of moving from state 1 (Pre-Progression) to state 2 (Progression). For the model structure defined in Figure 7, the transition intensity matrix is defined as;

$$Q(t) = \begin{bmatrix} q_{11} & q_{12} & q_{13} \\ 0 & q_{22} & q_{23} \\ 0 & 0 & 0 \end{bmatrix}$$

Of note is that this model allows for direct movement between the pre-progression and death (q_{13}) states without observing progression. The previously adopted method did not explicitly model this transition; the death time would have been used as the date of event for both PFS and OS. The sojourn time in each state r is exponentially-distributed with mean -1/qrr. The probability that an individual in state r moves to the next to state is -qrs/qrr.

For a **time-homogeneous** Markov model this matrix is independent of time e.g. the instantaneous risk of moving from state 1 (Pre-Progression) to state 2 (Progression)

is constant across the full life horizon of the study. For a **time inhomogeneous** model, transition intensities are allowed to change at a series of times common to each individual. Therefore, the transition intensity matrix changes at each of these time-points. Similarly, **covariate adjusted models** allow this transition intensity matrix to change according to levels of covariates.

Following specification of the transition intensity matrix, the likelihood (L(Q)) is maximised in terms of log(q_{rs}) using standard optimisation algorithms. The choice between model structures (time-homogeneous/inhomogeneous) and covariate selection can be based upon standard techniques; visual inspection of fitted curves, goodness of fit statistics (AIC/BIC) and p-value selection for nested models.

Model Output

The primary output of the MSM model is a transition probability matrix; P(t). The p_{rs} element of which denotes the probability of moving from state r to s over a fixed time interval; the interval length was chosen as 28 days for this analysis. For the model structure defined in Figure 1, the transition probability matrix is defined as;

$$P(t) = \begin{bmatrix} p_{11} & p_{12} & p_{13} \\ 0 & p_{22} & p_{23} \\ 0 & 0 & 1 \end{bmatrix}$$

For example, p_{12} is probability of moving from state 1 (Pre-Progression) to state 2 (Progression). As for the transition intensity matrices defined above; the transition probability matrix for a **time-homogeneous** Markov model is independent of time. As such, a time homogeneous model would only produce a single transition probability matrix. For a **time-inhomogeneous** Markov model, more than one matrix would be produced; the number of matrices being equal to the number of time-points for which the transition intensities are allowed to change plus one. Similarly, for a **covariate adjusted model**, the number of matrices would depend on the number of unique levels of covariates combinations used in the fitted model e.g. for a time homogenous model including covariates of Gender and Age [<65,>=65], four transition probability matrices would be produced.

Model Fitting

For the combined MM-009 and MM-010 data, it was found that a time inhomogeneous model significantly improved the model fit to the data compared to a

time homogeneous model. Hazard plots for OS and PFS were used to determine that there was a change in hazard at 168 and 728 days. There were therefore three probability transition matrices produced, one for each time interval (1-167, 168-727 and 728+ days). Further details on the model selection fitting process were provided in the original response.

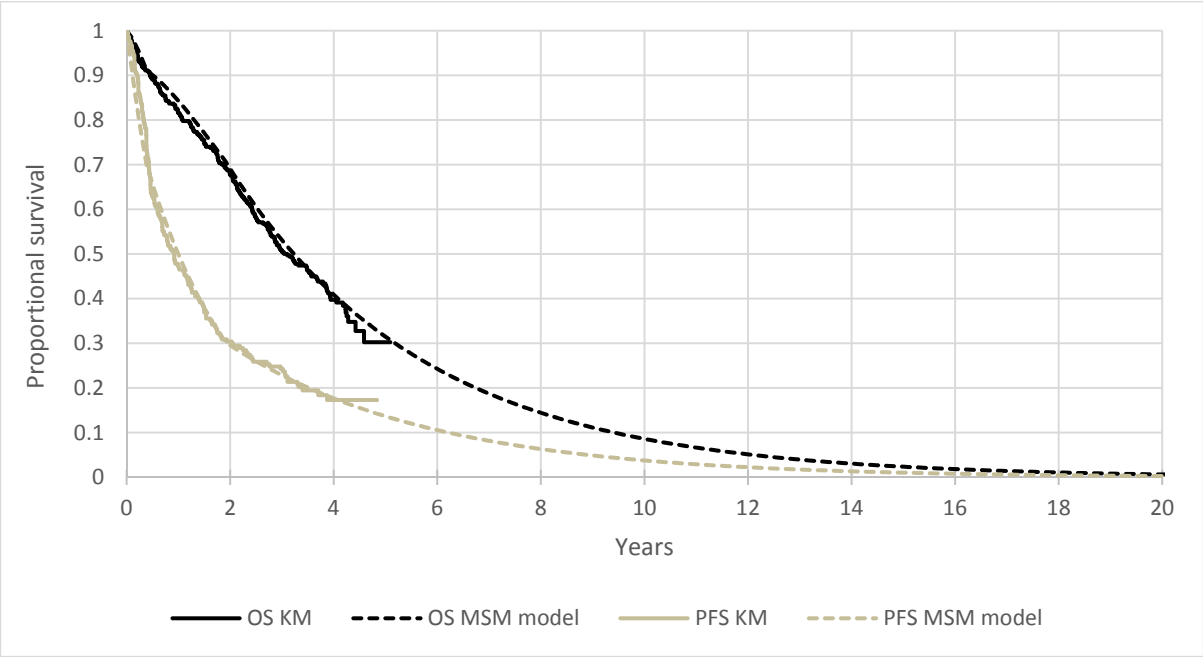
Predicting Outcomes with Output from Multi-State Model

To calculate the estimated number of people in each state at each time point, the probability transition matrices are multiplied out. For example, if the number of patients at the previous time point in states 1, 2 and 3 was x, y and z respectively, the model would estimate that there are $(p_{11} * x)$ patients in state 1, $(p_{12} * x) + (p_{22} * y)$ patients in state 2 and $(p_{13} * x) + (p_{23} * y) + (1 * z)$ patients in state 3.

Survival curves for PFS and OS were produced by calculating the estimated percentage of patients in state 1 and 2 (OS) and the estimated percentage of patients in state 1 (PFS) at each time point.

This procedure was extended beyond the end of the trial assuming that the matrix used for the last time point (728+ days) can be used to estimate survival beyond the observed trial period as shown in Figure 8.

Figure 8: Survival probabilities estimated from time-inhomogeneous* multi-state Markov models



Key: KM, Kaplan–Meier; OS, overall survival; PFS, progression-free survival.
Notes: *The hazard of moving between states is free to vary at 168 and 728 days in the statistical model

Covariate Adjustment

As described above, Multi-state Markov models allow for covariate adjustment by permitting the transition intensities to change according to levels of a covariate i.e. allowing the risk of moving between states to change depending on covariate values. Four covariates were found to be statistically significant in the model and were included in the final model. These were:

- prior treatments [1 vs 2 or 3],
- baseline beta-2 microglobulin [>2.5 or ≤2.5 mg/L],
- prior doxorubicin [Yes or No],
- worsening extramedullary plasmacytoma disease [Yes or no]

As such, 16 (one for each covariate combination; 2^4) transition probability matrices were outputted from the Multi-state Markov model for each time interval (1-167, 168-727 and 728+ days). A group prognosis method (Ghali et al. (2001)) was utilised to create an overall survival and progression-free survival curve for all patients. This involves outputting one matrix for each time point for each combination of covariates (i.e. 16 groups of patients in total).

Using the frequency of people in each combination of covariates (see supplementary table 1) survival curves are produced for each group of patients using the procedure described in the step above, a weighted survival curve is then produced simply by multiplying each survival curve by the proportion of patients within the relevant group and summing.

This procedure allows the survival curves to be calculated for OS and PFS for the group of patients of interest, e.g. second line patients. When survival curves have been estimated for second line patients only, survival curves have been weighted according to the % of patients in each of the 8 groups of second line patients.

How comparative effectiveness is estimated through calculation and comparison of adjusted survival curves

Comparative effectiveness was estimated by firstly calculated the projected OS and PFS for LEN+DEX in the group of patients most reflective of the patient characteristics in the clinical trial. This was done using the group prognosis method (as detailed above) by estimating the proportion of patients within each of the 16 groups based upon the information published within the clinical trial. Where no information was published the proportion of patients was assumed the same as within MM-009 and MM-010. This process produced the estimated survival for LEN+DEX if it was given to the same patients as received the comparator treatment within the study publication.

Following this the median OS and PFS was derived from the parametric survival curve simply by looking up the time-point at which 50% of patients had either died, or in the case of PFS either progressed or died.

The median for LEN+DEX was then compared to the median for the comparator and the hazard ratio calculated as the ratio of the two (this assumes underlying constant hazards for both arms – see the comment on this below).

Applying hazard ratios calculated from medians

As described above, for the Multi-State Markov model utilised, the sojourn time in each state is exponentially-distributed with mean -1/qrr. As such, there are parallels

between the output of an MSM and a piecewise exponential survival curve fitted to the data in a standard way. Technically applying a hazard ratio calculated from a median is only correct when this is applied to an exponential curve (as the assumption is that we can reflect comparative mean survival using only the point estimate for survival at the 50% mark i.e. that hazards are constant across time). This means that applying a hazard ratio calculated from a median is only technically correct for a time homogenous MSM model.

A balance must therefore be struck between:

- Modelling survival for LEN+DEX accurately – given that time homogenous models (and the exponential curve previously) do not fit the data well
- The technical incorrectness of a applying a HR calculated from a median

Analysis has been provided to the Committee as part of the response to the previous documentation demonstrating the impact of applying a HR calculated from a median compared to projecting use outcomes for curves fitted independently to survival curves for the comparator therapy using Petrucci 1989[1] (MP) and Ahn 2014[7] (BOR retreatment) datasets.

This comparison (presented again below) demonstrated that survival curves produced using hazard ratios produced from medians are generally similar / more optimistic in their projections of comparator outcomes than the survival curves produced using curves fit to Kaplan Meier data.

We therefore don't consider that any bias implicit in the requirement to use medians for comparison to some data sources is in favour of LEN+DEX (in fact the opposite is likely to be the case).

Figure 9: Comparison of survival curves fit to digitised data versus hazard ratios produced by median estimates using Petrucci 1989

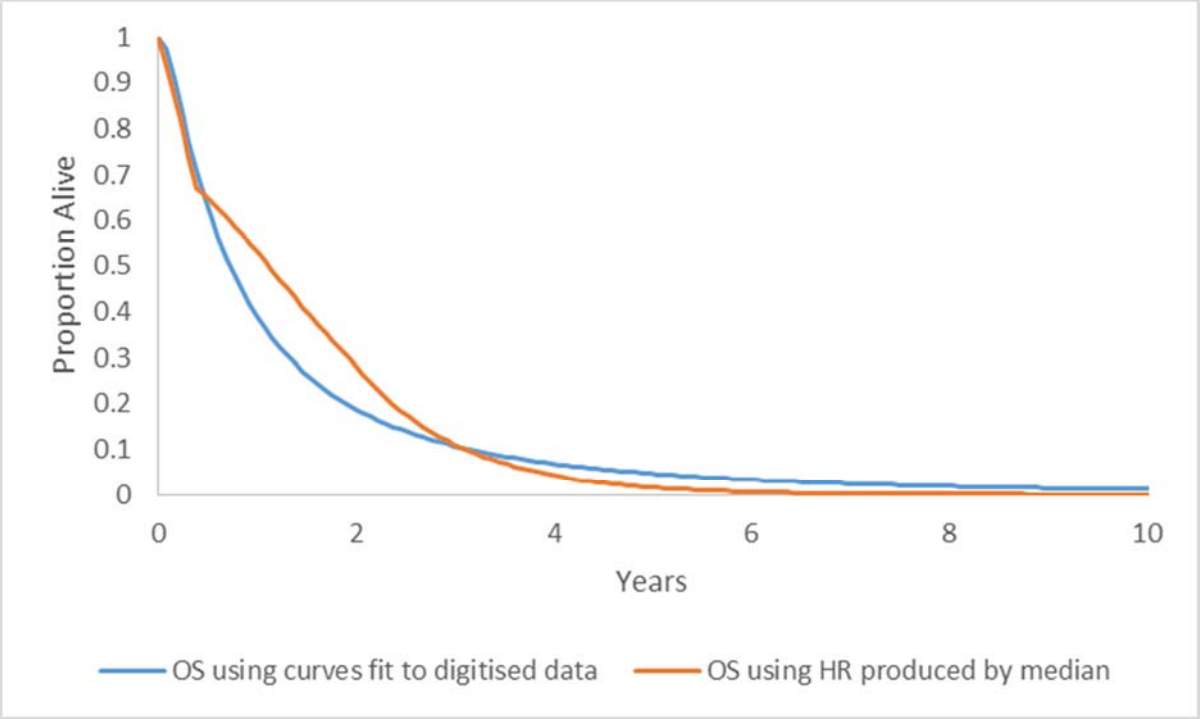
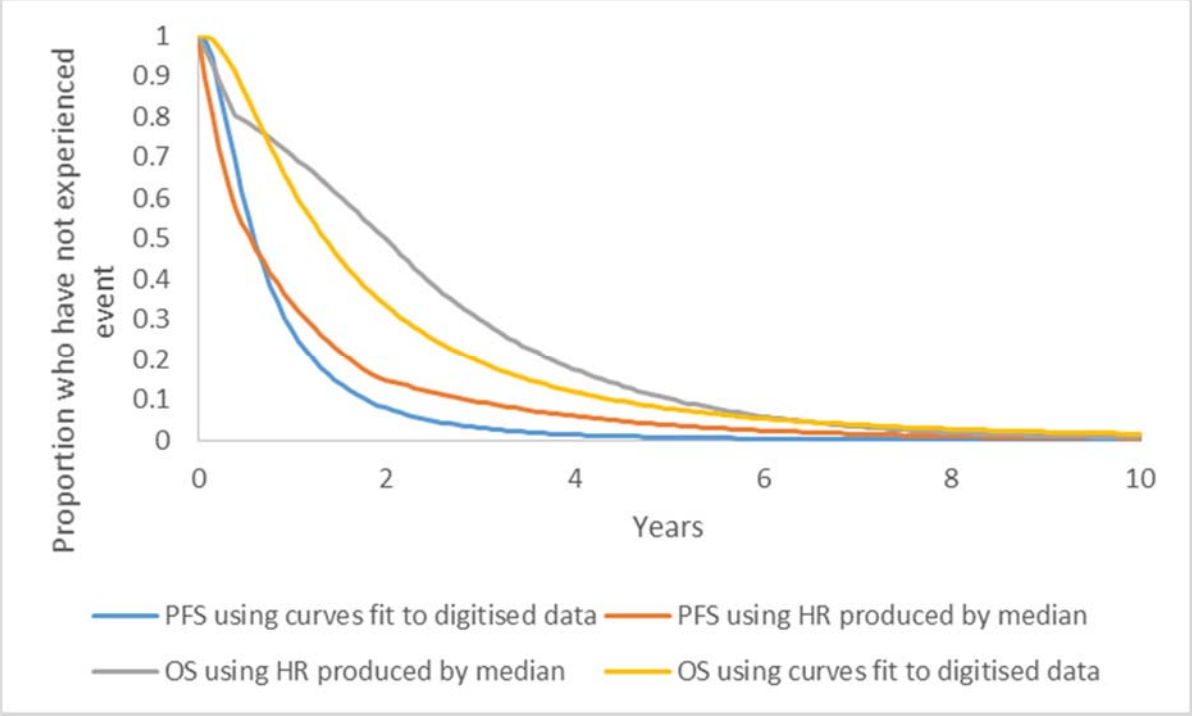


Figure 10: Comparison of survival curves fit to digitised data versus hazard ratios produced by median estimates using Ahn 2014



Supplementary Table 1: Combinations of covariates and frequencies

Group	Prior treatment	Beta-2 microglobulin	Prior doxorubicin	Worsening EPD	% all patients (n=353)	% 1 prior treatment patients (n=124)
1	1	< 2.5	Yes	YES	0.003	0.008
2	2 or 3	<2.5	Yes	YES	0.000	NA
3	1	>2.5	Yes	YES	0.008	0.024
4	2 or 3	>2.5	Yes	YES	0.008	NA
5	1	<2.5	No	YES	0.000	0.000
6	2 or 3	<2.5	No	YES	0.000	NA
7	1	>2.5	No	YES	0.008	0.024
8	2 or 3	>2.5	No	YES	0.006	NA
9	1	<2.5	Yes	NO	0.071	0.202
10	2 or 3	<2.5	Yes	NO	0.108	NA
11	1	>2.5	Yes	NO	0.088	0.250
12	2 or 3	>2.5	Yes	NO	0.266	NA
13	1	<2.5	No	NO	0.054	0.153
14	2 or 3	<2.5	No	NO	0.057	NA
15	1	>2.5	No	NO	0.119	0.339
16	2 or 3	>2.5	No	NO	0.204	NA

References

[1] Petrucci T, Avvisati G, Tribalto M, et al. Intermediate-dose (25 mg/m2) intravenous melphalan for patients with multiple myeloma in relapse or refractory to standard treatment. Eur J Haematol 1989;42:233-237

[2] Alexaniam R, Barcogie B, Dixon D. High-Dose Glucacorticoid Treatment of Resistant Myeloma. Annals of Internal Medicine. 1986;105: 8-11

[3] Alexaniam R, Dimopoulos M, Delasalle K, Barlogie B. Primary Dexamethasone Treatment of Multiple Myeloma. Blood, 1992;80(4): 887-890

[4] Facon T, Mary J-Y, Pegourie B, et al. Dexamethasone-based regimens versus melphalan-prednisone for elderly multiple myeloma patients ineligible for high-dose therapy. BLOOD, 2006; 107(4): 1292-1298

[5] Reyal Y, Popat R, Cheesman S, et al. Real world experience of bortezomib re-treatment for patients with multiple myeloma at first relapse. British Journal of Haematology. 2016; doi: 10.1111/bjh.14086

[6] Ghali WA, Quan H, Brant R, et al. Comparison of 2 methods for calculating adjusted survival curves from proportional hazards models. JAMA. 2001;286(12):1494-7.

[7] Ahn JS, Jung SH, Lee SS, et al. Clinical outcome of bortezomib retreatment in patients with relapsed or refractory multiple myeloma. Biomed Res Int. 2014: 145843.

[8] Ocio E, Richardson PG, Rajkumar SV, et al. New Drugs and Novel Mechanisms of Action in Multiple Myeloma in 2013: A Report from the International Myeloma Working Group (IMWG). Leukemia. 2014;28(3): 525–542

[9] Uy GL, Goyal SD, Fisher NM, et al. Bortezomib administered pre-auto-SCT and as maintenance therapy post transplant for multiple myeloma: a single institution phase II study. Bone Marrow Transplantation 2009;43: 793–800

[10] Schütt P, Brandhorst D, Stellberg W. Immune parameters in multiple myeloma patients: influence of treatment and correlation with opportunistic infections. Leukemia & Lymphoma, 2006; 47(8): 1570 – 1582

[11] Stadtmauer EA, Weber DM, Niesvizky R, et al. Lenalidomide in combination with dexamethasone at first relapse in comparison with its use as later salvage therapy in relapsed or refractory multiple myeloma. European Journal of Haematology. 2009;82: 426–432

[12] Kay R. A Markov model for analysing cancer markers and disease states in survival studies. Biometrics; 1986;42(4):855-65.

[13] Jackson, C. Multi-state models for panel data: The msm package for R. Journal of Statistical Software, 2011;38(8):

The clinical and cost-effectiveness of lenalidomide for people who have received at least one prior therapy with bortezomib (partial review of TA171)

A critique of the submission from Celgene

September 2016 Addendum

Produced by

Peninsula Technology Assessment Group (PenTAG)
University of Exeter Medical School
South Cloisters
St Luke's Campus
Heavitree Road
Exeter
EX1 2LU

Authors

Martin Hoyle, Associate Professor, PenTAG, University of Exeter
Nicola Huxley, Research Fellow, PenTAG, University of Exeter

Correspondence to

Martin Hoyle
Peninsula Technology Assessment Group (PenTAG)
University of Exeter Medical School
South Cloisters
St Luke's Campus
Heavitree Road
Exeter
EX1 2LU
m.w.hoyle@exeter.ac.uk

Date completed

14th September 2016

Source of funding

This report was commissioned by the NIHR HTA Programme as project number 13/07/01.

Declared competing interests of the authors

None

Acknowledgments

We acknowledge the excellent administrative support of Sue Whiffin and Jenny Lowe (University of Exeter) and Linda Long (University of Exeter) for checking the data reported from Reyal et al. (2016). We also acknowledge support and guidance given by Professor Chris Hyde (University of Exeter) relating to the critique of the company's clinical evidence. We also acknowledge the clinical advice provided by Claudius Rudin (Royal Devon and Exeter Hospital) and the support of Professor Rob Anderson (University of Exeter) who provided cover for the Director role.

Rider on responsibility for report

The views expressed in this report are those of the authors and not necessarily those of the NIHR HTA Programme. Any errors are the responsibility of the authors.

This report should be referenced as follows

Hoyle M. & Huxley N. The clinical and cost-effectiveness of lenalidomide for people who have received at least one prior therapy with bortezomib (partial review of TA171): A Single Technology Appraisal (September 2016 Addendum). Peninsula Technology Assessment Group (PenTAG), 2016.

Contributions of authors

Martin Hoyle	Contributed to the critique of Celgene's economic evaluation, developed the PenTAG base case and contributed to the writing and editing of this report.
Nicola Huxley	Provided project management, led the checking of the economic model submitted by Celgene. Contributed to the writing and editing of the report.

Contents

Abbreviations

AE	Adverse event
AFT	Accelerated failure time
ASH	American Society of Hematology
ASCT	Autologous stem cell transplantation
BOR	Bortezomib
CDF	Cancer Drugs Fund
CEAC	Cost-effectiveness acceptability curve
CI	Confidence interval
CIC	Commercial in confidence
CR	Complete response
CSR	Case study report
DEX	Dexamethasone
EBMT	European Society for Blood and Marrow Transplantation
ECOG	Eastern Cooperative Oncology Group
EMA	European Medicines Agency
EORTC	European Organisation for Research and Treatment of Cancer
ERG	Evidence review group
FDA	Food and Drug Administration
G-CSF	Granulocyte-colony stimulating factor
HR	Hazard ratio
HRQoL	Health related quality of life
ICER	Incremental cost effectiveness ratio
IDMC	Independent Data Monitoring Committee
IMiD	Immunomodulatory drug
ITT	Intention to treat
KM or K-M	Kaplan-Meier
LEN	Lenalidomide
LEN+DEX	Lenalidomide plus dexamethasone

MA	Meta-analysis
MR	Minimal response
MM	Multiple myeloma
MP	Melphalan plus prednisolone
nCR	Near complete response
NE	Not estimable
NMB	Net monetary benefit
ORR	Overall response rate
OS	Overall survival
PAS	Patient access scheme
PD	Progressive disease
PFS	Progression-free survival
PR	Partial response
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PSA	Probabilistic sensitivity analysis
QALY	Quality adjusted life year
QLQ	Quality of Life Questionnaire
rrMM	Relapsed or refractory multiple myeloma
SCT	Stem cell transplant
SD	Stable disease
SE	Standard error
SPC	Summary of product characteristics
SR	Systematic review
TTF	Time to treatment failure
TTP	Time to progression
VGPR	Very good partial response
WTP	Willingness to pay

Background & Summary

This addendum presents a critique of the most recent submission by Celgene as part of the NICE Single Technology Appraisal ID667 "Lenalidomide for the treatment of multiple myeloma in people who have received at least one prior therapy with bortezomib (partial review of TA171)". This appraisal considers the use of lenalidomide for the treatment of multiple myeloma in people who have received at least one prior therapy with bortezomib and who are unable to receive thalidomide.

This submission, Celgene's "June 2016 submission", was received by PenTAG on 21st June 2016 and is intended to supplement Celgene's original submission received November 2013 and the additional work received both in 2014 and 3rd February 2016.

The most recent NICE appraisal meeting for this HTA was held in April 2016. We wrote Addenda dated 7th March and 1st April 2016 in preparation for this meeting. This addendum supplements these previous two Addenda.

In Section 1, we report that Celgene have changed their opinion on the relative importance of the two comparators, bortezomib retreatment (BOR) and melphalan plus prednisolone (MP). They now no longer consider BOR as a comparator, but continue to consider MP.

Sections 2-6 concern the questions that NICE asked Celgene in response to the Appraisal meeting of April 2016.

Unfortunately, we have found four serious errors in Celgene's revised economic model (Section 2). We explain how we have corrected for these errors.

Section 3 concerns Celgene's adjustments for use of 3rd-line lenalidomide. We still consider these adjustments flawed.

In Section 4, we discuss the usefulness of some new trial data on the effectiveness of BOR. Celgene do not use this data in their economic analysis, as they no longer consider BOR as a comparator. In Section 7, we discuss how we have used this data in scenario analyses.

In Section 5, we report that, in response to a request from NICE, Celgene have provided an explanation of the multistate model.

In Section 6, we discuss Celgene's response to NICE's request to run a scenario analysis using a mortality hazard ratio of 1 after stopping treatment.

Finally, in Section 7, we present our revised estimates of cost-effectiveness in the light of the conclusions from the April committee meeting and Celgene's June 2016 submission. All ICERs are highly uncertain, but our revised base case ICER for

- LEN+DEX vs. MP lies between £34,000 and £90,000 per QALY,
- LEN+DEX vs. BOR lies between a value which is >£31,000 per QALY and a value which is >£83,000 per QALY.

1 Comparator treatments

In their submission for the April 2016 NICE committee meeting, Celgene compared the cost-effectiveness of lenalidomide plus dexamethasone (LEN+DEX) versus bortezomib re-treatment (BOR); and LEN+DEX vs. melphalan plus prednisolone (MP).

Our clinical expert advised that BOR and MP are rarely used in the NHS, and that bendamustine was the most appropriate comparator. However, we believe the opinion of the April NICE appraisal committee was not to reintroduce bendamustine as a comparator in this appraisal.

In their June 2016 submission, Celgene now present MP as the main comparator. They have justified this with the explanation that bortezomib re-treatment is no longer funded via the CDF, and that NHS England have informed hospitals that they will not be funding re-treatment via TA129 (which in 2007, NICE recommended BOR retreatment for progressive multiple myeloma for people whose multiple myeloma has relapsed for the first time after having one treatment, and who have had a bone marrow transplant, unless it is not suitable for them). As such, Celgene now present some of their analyses only for the comparison between LEN+DEX and MP.

It is our understanding that the April NICE appraisal committee members were aware of the concerns around funding for BOR, but still considered BOR to be a relevant comparator for people who are suited to BOR. We also understood that BOR is not appropriate for about half of patients either due to lack of response or adverse reactions. The committee also considered MP to be a relevant comparator, although may be used for few patients. Therefore, in this Addendum, we present cost-effectiveness estimates for both LEN+DEX vs. BOR and LEN+DEX vs. MP.

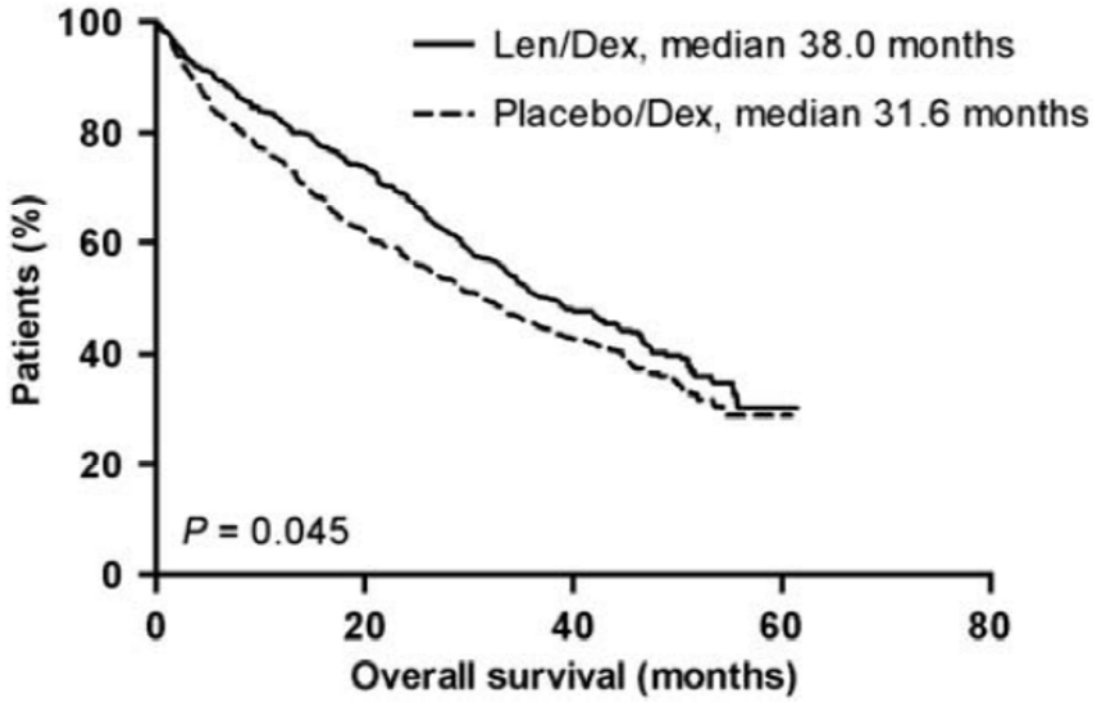
2 Validation of Celgene's model (using data for dexamethasone)

The clinical effectiveness evidence for BOR and MP is of very low quality, as both are taken from small non-randomised studies. The committee were therefore concerned that the relative effectiveness of BOR and MP versus LEN+DEX should have face validity.

In their request for additional evidence from Celgene, NICE mentioned that the April 2016 committee noted that Celgene's model predicts a mean survival benefit of 2.2 years for LEN+DEX vs. BOR, whereas the MM-009 and MM-010 trials show a median survival benefit of 6.4 months for LEN+DEX vs. DEX, and that this raised concerns about the external validity of Celgene's model, as the committee believed that BOR is more effective than DEX.

We agree that the MM-009 and MM-010 trials show a median survival benefit of 6.4 months for LEN+DEX vs. DEX for all patients in these trials combined (2nd- and subsequent lines of treatment) (Figure 1).

Figure 1. OS for pooled data from MM-009 and MM-010 RCTs for 2nd- and subsequent lines of patients combined



Source: Figure taken from Fig. 3 of Dimopoulos et al (2009)[1]

NICE therefore asked Celgene to provide mean life years (LYs) and quality-adjusted life years (QALYs) using their model for the DEX arm with clinical data taken from the MM-009 and MM-010 trials. We understand that 48% of patients in the DEX arm switched to LEN+DEX after disease progression or study unblinding. NICE requested no adjustment for this switching.

Celgene provided the requested information (Table 1). The results for the BOR and MP arms are reported under the scenario that no lenalidomide is used 3rd line. This was not the

original base case presented by Celgene; and no explanation is given for why this is the case. As such, we present the undiscounted life years under Celgene's base case for BOR and MP. This includes 100% lenalidomide use in third line for the BOR and MP arms and so has results that are slightly more comparable to the DEX only arm. We note that when LEN is used 3rd line, LYG is still higher in DEX arm than BOR and MP arms, despite 100% of patients receiving LEN in BOR and MP (compared to 48% in DEX).

Table 1. Clinical outcomes by treatment reported by Celgene

	Mean OS		Mean PFS		Mean PD	
	Celgene June 2016 submission	**Celgene February 2016 base case**	**Celgene June 2016 submission**	**Celgene February 2016 base case**	**Celgene June 2016 submission**	**Celgene February 2016 base case**
DEX	3.97	NA	0.99	NA	2.98	NA
MP	1.14	3.15	0.43[a]	0.43[a]	0.72	2.73
BOR	2.70	3.69	1.87	1.97	0.83	1.71
LEN+DEX	5.87		2.96		2.91	

Key: BOR, bortezomib; DEX, dexamethasone; LEN, lenalidomide; MP, melphalan plus prednisolone; PD progressive disease.

Notes: All values reported are undiscounted life years. In the Celgene submission, LYs are reported undiscounted, QALYs discounted by 3.5% per annum. a) PFS for MP does differ between these submissions, but the difference is <0.01

We understand that the NICE committee were expecting mean OS for BOR (with 100% of patients subsequently taking LEN+DEX) to be greater than for DEX (with 48% of patients subsequently taking LEN+DEX), given the belief that BOR is more effective than DEX, and a greater proportion of patients in the BOR arm subsequently take LEN+DEX. This is not the finding from Celgene's model, which casts doubt on Celgene's estimate of effectiveness for BOR. Indeed, Celgene admit that, in the light of these results, there are "*evidently issues with the validity of the comparative effectiveness compared to MP and BORT retreatment*"(p3 Celgene submission June 2016). Furthermore, Celgene seem unclear about the expected relative effectiveness of MP and DEX, as, the previous quote contradicts their later statement "*ordering of efficacy which would be expected based on our clinical knowledge of the treatments; LEN+DEX, BORT retreatment, DEX, MP*." (p3 Celgene submission June 2016).

Celgene appear to use this lack of face validity to abandon their attempts to model the BOR arm, stating "*Based upon this information, we believe that the comparison to MP at least should be considered further*" (p3 Celgene submission June 2016).

Celgene then provide evidence for the equivalence of the effectiveness of DEX and MP.

Celgene claim that the only study which compares the effectiveness of MP and DEX is Facon et al. (2006).[2] They say this study found no significant difference in OS between the following treatments: MP, DEX, MP+DEX, and DEX–interferon alpha, and we agree.

Given time constraints, we have not checked whether other relevant studies exist. However, as Celgene explain, patients in Facon et al. (2006) were 1st-line, whereas patients in the current HTA are people with for whom thalidomide is contraindicated and whose disease has progressed after ≥1 prior treatment with bortezomib. The trial was conducted in 1995-8, and there were 122 patients in the MP arm and 127 in the DEX arm.

We find that the paper also states that PFS was significantly longer for patients on MP. If anything, this study then suggests that MP may be more effective than DEX for the current HTA, although the evidence is weak.

Celgene claim it is beneficial to use the clinical effectiveness evidence from the DEX arm because:

- There is randomisation between LEN+DEX and DEX. We agree that this is advantageous.
- The impact of 3rd-line LEN+DEX on OS can be accurately captured. We agree that this is advantageous and note that Celgene correctly model a cost corresponding to 48% of patients in the MP arm subsequently receive LEN+DEX, as 48% of patients in the DEX arm in the MM RCTs subsequently received LEN+DEX.
- Uncertainty in comparator effectiveness is removed as PFS and OS are available from a mature dataset with patient level data, meaning that comparison using medians alone is no longer required. We agree that this is advantageous.

Celgene then estimate the following ICERs for LEN+DEX vs. MP given the following two scenarios:

- ICER = £23,000 per QALY given PFS for MP is unchanged and mean PD for MP is set equal to that of DEX. Contrary to this, we find the ICER to be £21,000 per QALY when setting the comparator to MP in cell D28 "Controls" tab and "Set post progression equal to DEX" to "Yes" in cell D61.
- ICER = £24,000 per QALY given mean PFS and PD for MP are set equal to that of DEX. Celgene consider this the most appropriate option. Contrary to this, we find the ICER to be £20,000 per QALY when setting the comparator to MP in cell D28 "Controls" tab and "Set pre-progression and post progression equal to DEX" to "Yes" in cell D63.

These ICERs are very similar to their base case ICER of £24,000 per QALY.

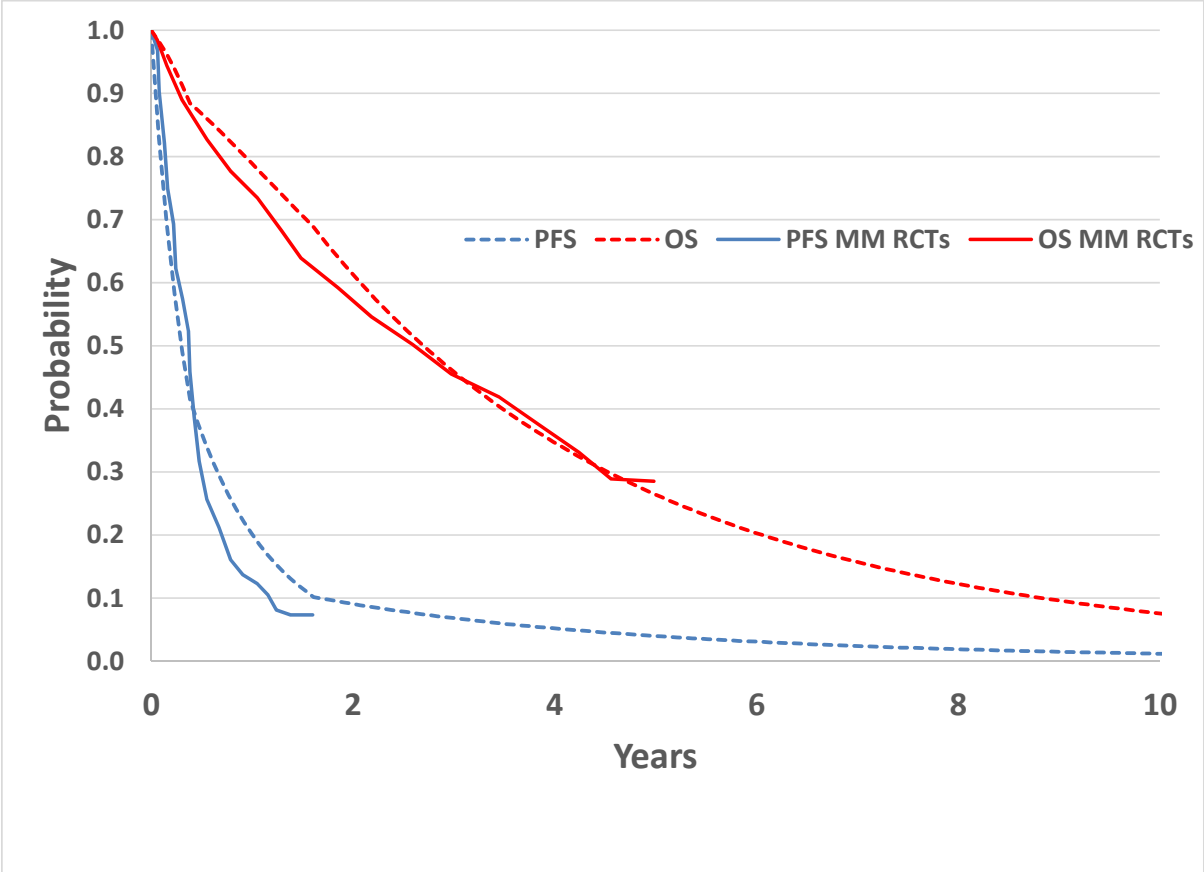
Celgene do not provide similar estimates for the cost-effectiveness of LEN+DEX vs. BOR because, as discussed above, they now dismiss BOR as a comparator. However, given that we still consider BOR as a comparator, we provide estimates of cost-effectiveness of LEN+DEX vs. BOR, given BOR PFS and OS is set equal to that for DEX in Section 7.2, p26.

Unfortunately, we find the following four serious problems with Celgene's scenario of assuming PFS and OS for MP is set equal to that for DEX.

2.1 PFS for DEX too long-tailed

We find that the modelled OS for DEX appears visually to fit the data from the MM RCTs well (Figure 2). The modelled PFS also appears visually to fit the data from the MM RCTs reasonable well (Figure 2).

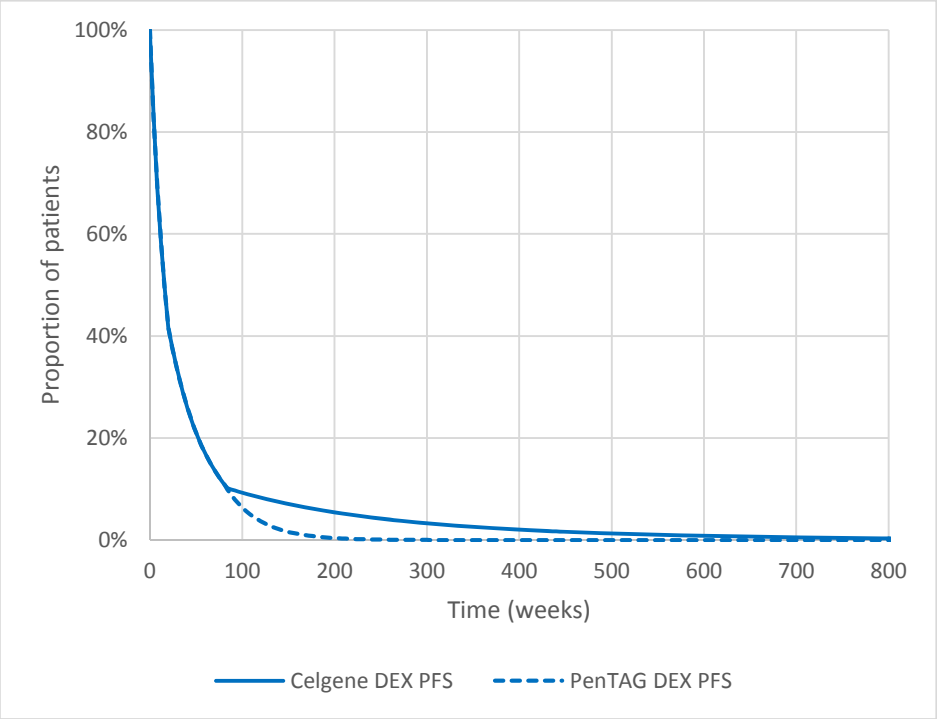
Figure 2. PFS and OS for DEX from MM-009 and MM-010 RCTs vs. Celgene model fit



However, extrapolation of the modelled PFS for DEX appears implausible. Specifically, the PFS tail appears implausibly long (Figure 2). We have corrected for this as follows. First, we assume the same PFS up to 1.5 years. Thereafter, we assume an exponential distribution with rate parameter consistent with PFS at 1.5 years (Figure 3).

We have amended the model to give the option of correcting for this error. For this change alone, Celgene's ICER for LEN+DEX vs. MP increases from £20,000 to £23,000 per QALY. The ICER increases because the estimated total acquisition and administration costs of MP then fall, as they are proportional to the time in PFS.

Figure 3. PFS DEX Celgene vs. PenTAG curve fits



Key: DEX, dexamethasone; PFS, progression free survival

2.2 Celgene DEX PFS and OS data incorrect line of treatment

Next, Celgene have modelled PFS and OS for DEX based on data from all lines of treatment, including 2nd- and subsequent lines, from the MM RCTs. However, we require 2nd-line data only to be consistent with the modelled PFS and OS for LEN+DEX which was fit to just 2nd-line data from the MM RCTs.

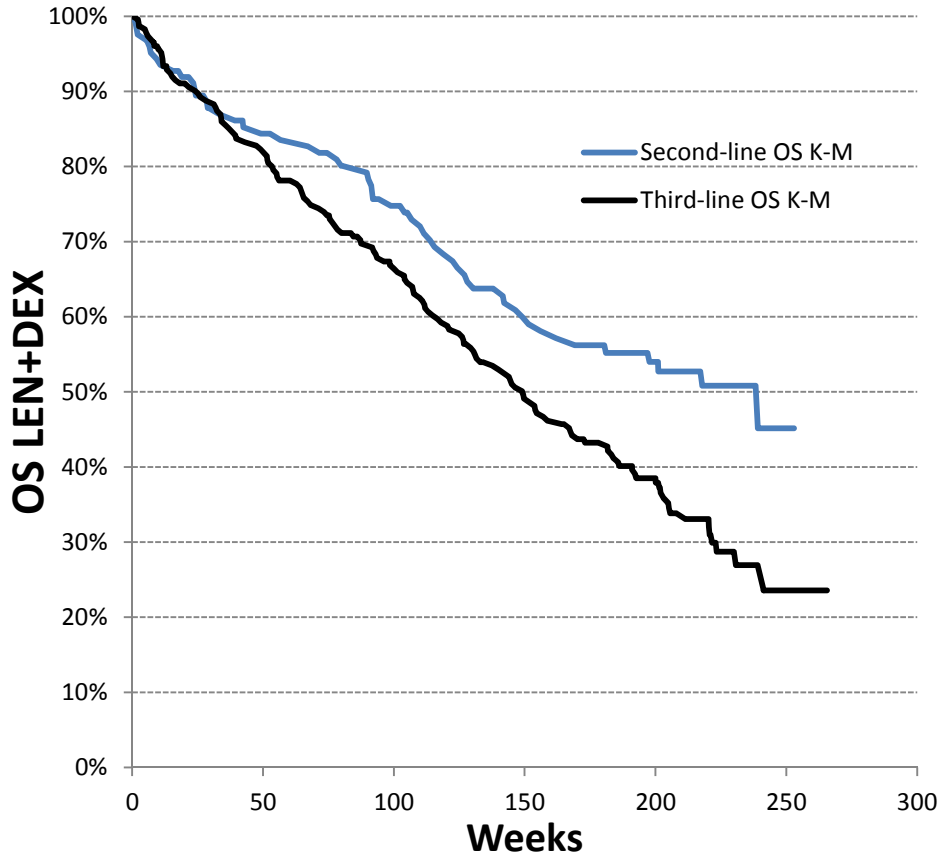
The underlying publications do not give OS for LEN+DEX or DEX split by line of treatment.[1, 3, 4] However, in their model, Celgene provide Kaplan-Meier graphs for the pooled data from the two RCTs for LEN+DEX by line of treatment (Figure 4). Even though there are rather a small number of 2nd-line patients, 65 (compared to 288 patients in subsequent lines), OS for LEN+DEX is clearly better 2nd- than subsequent lines: median OS is 238 and 149 weeks for 2nd- and subsequent lines respectively. The published literature does provide some relevant data for PFS of DEX (Table 2), but not for OS.[3, 4] This data shows similar median PFS for DEX according to line of treatment (Table 2). Given that median PFS for 2nd-line LEN+DEX was not reached, but was reached for subsequent lines, it appears that PFS for LEN+DEX was longer 2nd-line compared to subsequent lines.

Given all this information, it is reasonable to use 2nd- and subsequent lines data combined for PFS for DEX (as Celgene have done), in place of just 2nd-line data. Therefore, we do not adjust the PFS data for DEX (corresponding to all lines of treatment combined) provided by Celgene.

However, it is not clear whether to adjust the OS data for DEX provided by Celgene. On the one hand, one could assume the same OS for DEX 2nd-line line as for all lines, given that

PFS for DEX is independent of line of treatment. On the other hand, one could argue that OS for DEX 2nd-line is likely to be longer tailed than for all lines combined, given that this was observed for LEN+DEX.

Figure 4. OS for LEN+DEX by line of treatment



Key: DEX, dexamethasone; LEN, lenalidomide; OS, overall survival;

Table 2. Median PFS (months) for LEN+DEX and DEX from MM-090 and MM-010

	RCT	2nd-line	≥3 line
LEN+DEX	MM-090	Not reached	10.2
DEX		5.1	4.6
LEN+DEX	MM-010	Not reached	11.1
DEX		4.7	4.7

Key: DEX, dexamethasone; LEN, lenalidomide.

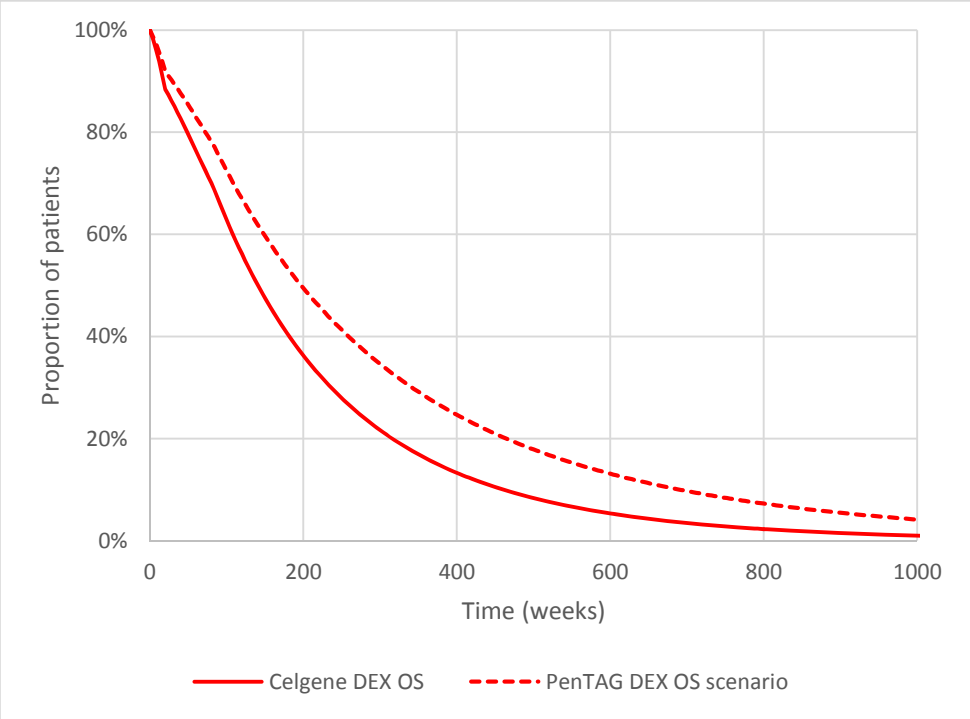
Given that we do not have the data for OS for DEX 2nd-line, in our base case, we cannot be certain that Celgene have underestimated OS for DEX in their mode, and so we do not

adjust the OS for DEX as provided by Celgene. However, in an important sensitivity analysis, we extend the OS for DEX provided by Celgene as follows.

Given that median OS for LEN+DEX 2nd -line is 238.3 weeks and median OS for all patients combined, 165 weeks,[1] this gives a ratio of median OS 2nd-line vs. all lines of 238.3 / 165 = 1.44. We use this ratio to adjust OS for DEX in the following way, implemented in tab "MSM_2L (DEX)" in the model. Assuming an accelerated failure time model, an increase in median OS by a factor of 1.44 translates to an increase in mean OS also by a factor of 1.44. We find Celgene's estimated mean OS for DEX is 3.89 years, which gives our estimated mean OS for DEX as 3.89 x 1.44 = 5.61 years. Our estimated OS curve for DEX was then set equal to Celgene's curve raised to the power of a certain hazard ratio (Figure 5). The hazard ratio was estimated as 0.69 using the Excel "Solver" function in such a way as to yield the correct mean OS of 5.61 years.

When this scenario is implemented in isolation, Celgene's ICER for LEN+DEX vs. MP increases substantially, from £20,000 to £35,000 per QALY.

Figure 5. Modelled DEX OS: Celgene vs. PenTAG scenario analysis



2.3 Error in cost of acquisition of MP

Celgene introduced a wiring error in the calculation of the acquisition cost of MP when the clinical outcomes of MP are set equal to those for DEX. This error substantially affects the estimated cost-effectiveness of LEN+DEX vs. MP.

The wiring error, which affects the cost of MP, occurs in the "PF.Comparator" sheet of the model, in column CU. Celgene have added a function so that different costs are applied

depending on the relevant scenario chosen in the “Controls” sheet. This allows for the PF.Comparator sheet to present the calculation for a range of comparators in different scenarios, depending upon the scenario chosen in the “Controls” sheet. Unfortunately the coding for the new scenarios where MP is set to have post-progression modelled the same as DEX, has missed a set of brackets and therefore, instead of applying the cost of prednisolone to the proportion of patients still on MP in each cycle, a single fixed cost of prednisolone is applied for each cycle of the model. This results in the dramatic increase in cost of MP acquisition. For example, the incorrect per patient cost of acquisition of MP is £22,390, whereas it should be £2,344.

To correct this, in column CU, cell CU17, we replaced AJ17*AO17*Cost.Cycle.Melph+Cost.Cycle.PredIV with AJ17*AO17*(Cost.Cycle.Melph+Cost.Cycle.PredIV) and copied down for all cells in this column.

This then replicates Celgene’s calculation in their February 2016 base case.

We have amended the model to give the option of correcting for this error. Given this change alone, Celgene’s ICER for LEN+DEX vs. MP increases substantially, from £20,000 to £46,000 per QALY, when MP is assumed to have the pre- and post-progression survival as DEX.

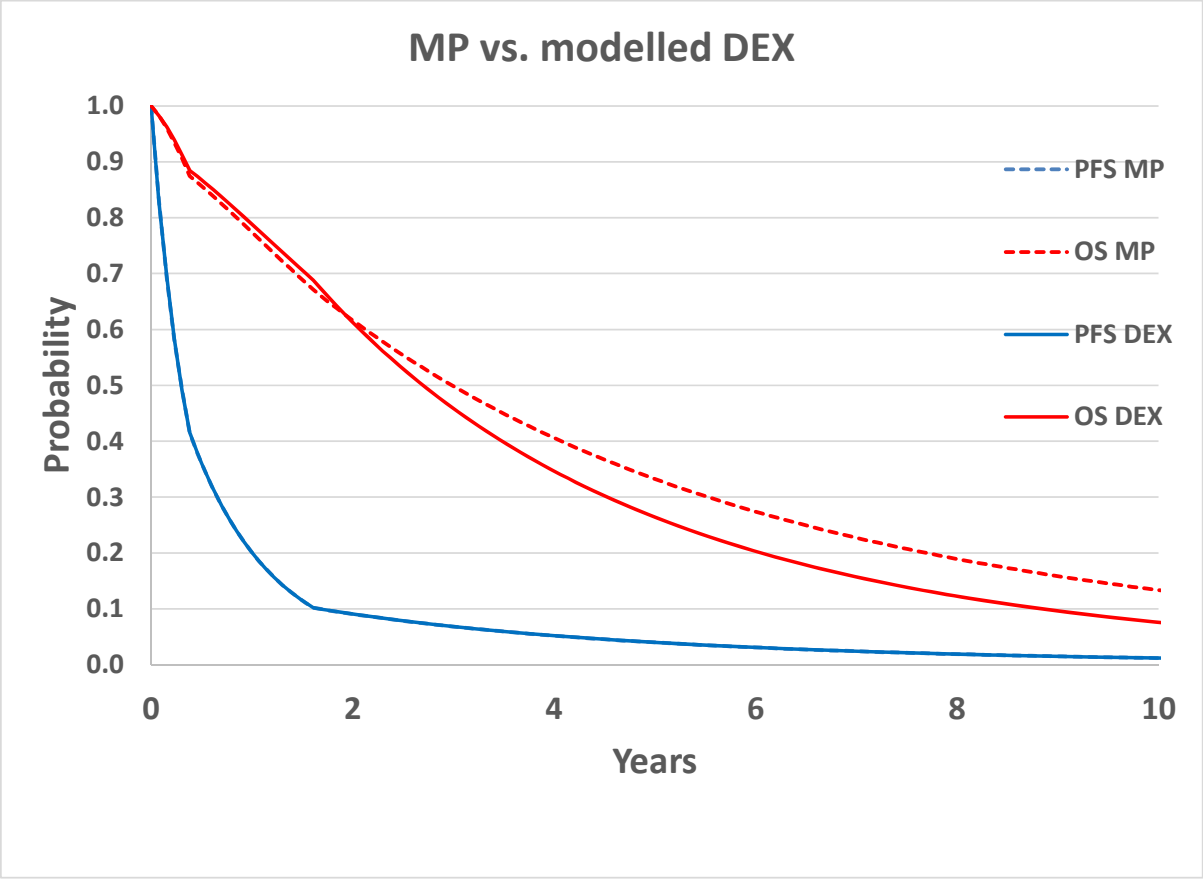
2.4 OS for DEX not used correctly for MP

Celgene then attempt to use the modelled PFS and OS for DEX in the MP arm, i.e. PFS for MP is set equal to PFS for DEX and similarly for OS. In their model, this is supposed to be achieved by setting the comparator to MP in cell D28 “Controls” tab and “Set pre-progression and post progression equal to DEX” to “Yes” in cell D63.

However, these changes do not have the desired effect, as OS for MP does not equal that for DEX, although PFS is correct (Figure 6).

We have amended the model to give the option of correcting for this error. For this change alone, Celgene’s ICER for LEN+DEX vs. MP decreases from £20,000 to £15,000 per QALY because OS for MP is then shorter-tailed.

Figure 6. PFS and OS DEX vs. MP arm with DEX data



3 Subsequent treatments

NICE asked Celgene to explain how the comparator arms of the model are adjusted to reflect the costs and benefits of third- and fourth-line treatments.

In response, Celgene discuss the technical details involved in calculating the costs and life years associated with subsequent treatments.

Whilst we thank Celgene for their clarification of their model, we are still not satisfied with the way they have modelled subsequent treatments.

Celgene then comment on two elements of our base case, namely:

- Removing all costs of 3rd-line LEN+DEX.
- Removing adjustment to OS for MP and BOR in respect of the use of 3rd-line LEN+DEX.

They find our calculations to be technically accurate, but say these assumptions are not in line with current treatment pathways. In response, we argue that subsequent treatments should be modelled correctly (which they are not), or not at all.

NICE then tell Celgene that we find the impact of their adjustment to OS for BOR and MP difficult to reconcile with their addition of the substantial cost of 3rd-line LEN treatment. Indeed, we understand that the April 2016 NICE committee also found these adjustments implausible.

For instance, in Celgene's base case, in the BOR arm, when adjustment is made for 3rd-line use of LEN, mean OS increases only moderately, from 3.59 to 3.69 years, but there is an additional substantial cost of £20,400 in respect of 3rd-line use of LEN.

If the PFS and OS 2nd-line hazard ratios are set to 1, the 3rd line LEN adjustment *reduces* BOR OS (4.37 LYs compared to 5.87 without adjustment), whereas we would expect OS to increase. Furthermore, there appear to be fewer life years gained in progressed disease for the BOR arm compared with the base case (1.43 compared to 1.71 LYs in the base case), but higher costs (£32,549 compared to £31,918 in the base case).

In response, Celgene now admit that their method of adjusting OS for the BOR arm for 3rd-line LEN is "*causing spurious results*" such as we and NICE outline above. They also say "*This problem does not exist for the comparison versus MP, where the third-line LEN outcomes are consistently better than those with MP (for which a published KM was actually available for OS)*". By this, we assume Celgene believe their adjustment for the MP arm for 3rd-line LEN is plausible.

By contrast, we believe that their adjustment for subsequent treatments is not plausible, because of the problems outlined for the adjustment for the BOR and the logical inconsistencies outlined above.

4 New clinical evidence for bortezomib retreatment

NICE then asked Celgene to provide data from a study on BOR retreatment from UCLH. Celgene responded to say that data from this study has been published this year as Reyal et al. (2016).[5]

Celgene give the patient baseline characteristics and data on response to BOR in Tables 4 and 5 of their response document respectively. We agree that they have correctly reproduced these Tables from Reyal et al. (2016).[5]

Celgene do not comment on this data, as they now believe that it is inappropriate to consider BOR as a comparator (Section 0, p8).

The current HTA concerns adults with multiple myeloma for whom thalidomide is contraindicated and whose disease has progressed after ≥1 prior treatment with BOR. All patients in Reyal et al. (2016) had indeed progressed after ≥1 prior treatment with BOR. [5] Although there is no mention of contraindication to thalidomide, we consider this a minor weakness only, as this criterion was not enforced in the selection of the data for LEN+DEX, taken from the MM RCTs.

The study is limited as it included only 23 patients from a single centre, and there is no data on OS. However, it does include recent NHS patients.

There was a median of 5 BOR treatment cycles in the study by Reyal et al. (2016). This compares with Celgene's assumed mean of 6.6 cycles, and the estimated 3.8 cycles from Taverna et al. (2012)[6], which we use Section 7.1, p23.

Next, median PFS in Reyal et al. (2016) was 14.4 months. This compares with Celgene's modelled median of 10.1 months and mean of 23.7 months.

We later use data from Reyal et al (2016) to inform the effectiveness of BOR in a scenario analysis (Section 7.1, p23).

5 Explanation of Celgene's 2016 model

NICE asked Celgene to provide a clear explanation of the multistate model in language suitable for non-specialists.

Celgene have done as requested.

6 Scenario assuming no survival benefit after stopping treatment

NICE requested the following from Celgene:

"*a scenario analysis using a hazard ratio of 1 after stopping treatment, because there is no evidence of an ongoing survival benefit beyond that time. Please present this analysis for both comparators (bortezomib retreatment and melphalan)*."

In response, "*Celgene do not believe this is a valid comparison as the MM-009 and MM-010 data shows a PPS benefit for LEN+DEX over DEX, when the DEX arm is adjusted for cross-over to LEN*." In response, in the current STA, LEN+DEX is compared to MP and BOR, not DEX, and whilst there may be a PPS benefit for LEN+DEX over DEX, this does not necessarily mean there is a PPS benefit of LEN+DEX over MP or BOR. Also, Celgene support their claim by their observation that the MSM modelled PPS curves for LEN+DEX and DEX, based on data from MM-009 and MM-010 are almost identical, even though a large proportion of patients in the DEX arm subsequently received LEN. We do not have the underlying data to check this. Also, instead of presenting the MSM modelled PPS, it would seem more appropriate to present the Kaplan-Meier curves for PPS for LEN+DEX and DEX from the MM-009 and MM-010 trials.

Celgene continue: "*in responders, the lack of cumulative toxicity and immunomodulatory activity of LEN resulting in enhanced antitumour effects mediated by T and natural killer (NK) cells, may prolong remissions. This beneficial immunomodulatory effect is less likely with immunosuppressive regimens containing bortezomib or conventional chemotherapy which are limited to a fixed treatment durations due to cumulative toxicities*.". We are not qualified to comment on this claim.

Celgene then say that they could not implement the request from the committee because of the structure of their model. Instead, they assumed equal PPS across arms and they explain how this was achieved in their model. Celgene report the resulting ICERs in the table below.

Table 3. Celgene reported ICERs

	LEN+DEX vs BOR	LEN+DEX vs MP
PPS for comparator equal to PPS for LEN+DEX	£31,000	£22,000
PPS for comparator equal to PPS for LEN+DEX, no cost of subsequent LEN treatment	£71,000	£36,000
PFS for comparator equal to PFS for DEX, PPS for comparator equal to PPS for LEN+DEX	Not reported	£23,000
PFS for comparator equal to PFS for DEX, PPS for comparator equal to PPS for LEN+DEX, no cost of subsequent LEN treatment	Not reported	£38,000

Key: BOR, bortezomib; DEX, dexamethasone; LEN, lenalidomide.

In Section 7.3, p27, we describe our methods of implementing the request of the NICE committee.

7 PenTAG revised estimates of cost-effectiveness

We repeat our caveat from our Addendum of 7th March 2016 that all ICERs are highly uncertain, in part because:

- Given that modelled TTF, PFS and OS for LEN are based on immature data, substantial extrapolation is required.
- The utilities are highly uncertain.

Additional sources of uncertainty are discussed below.

7.1 Scenario assuming PFS and OS for MP and BOR taken from single arm studies

In our previous addendum, in common with Celgene, the clinical effectiveness data for:

- MP was taken from Petrucci et al (1989).[7]
- BOR was taken from Taverna et al (2012).[6]

In this section, the clinical effectiveness for MP is unaltered. Taverna et al (2012) [6] is still used for BOR.

7.1.1 Effectiveness of BOR informed by Reyal et al. (2016)

In addition, in a separate analysis, we use data from Reyal et al. (2016)[5] to estimate the effectiveness and treatment duration for BOR.

As discussed in Section 4, p19, there was a median of 5 BOR treatment cycles in Reyal et al. (2016). This compares with Celgene's assumed mean of 6.6 cycles, and the estimated 3.8 cycles from Taverna et al. (2012)[6]. Next, median PFS in Reyal et al. (2016) was 14.4 months. This compares with Celgene's estimate of a median of 10.1 months a mean of 23.7 months. In our new scenario analysis, we assume 5 BOR treatment cycles, and a median PFS for BOR of 14.4 months (both in common with Reyal et al (2016)). We also assume an accelerated failure time model for BOR PFS, which gives a mean PFS of 23.7 x (14.4 / 10.1) = 33.8 months = 2.8 years. We then have a range of options for OS for BOR. We could either leave Celgene's mean OS of 3.7 years unchanged, by reducing mean PPS, or we could assume PPS is unchanged with the extended PFS, giving a mean OS of 4.5 years, or we could inflate OS as for PFS to a mean of 3.7 x (14.4 / 10.1) = 5.3 years. We have chosen the second method: extending PFS, but leaving PPS unchanged.

7.1.2 PenTAG previous changes to Celgene model

As in our Addendum of 7th March 2016, our updated ICERs comprise three changes to the Celgene 2016 model.

- No subsequent (3rd- and 4th-line) treatment costs, Section 4.2.2.5, p.36 of our Addendum of 7th March 2016).
- No adjustment to OS for BOR and MP for subsequent (3rd- and 4th-line) treatments. This is achieved by estimating OS for BOR and MP based on the HR alone, as assumed in the 2014 model (Section 4.2.2.4, p.34 of our Addendum of 7th March 2016). As discussed in Section 3, p18, we still believe that Celgene's adjustment for OS for 3rd-line LEN in the BOR and MP arms is not credible.
- Reduce the mean duration of BOR from 6.6 to 3.8 treatment cycles (Section 4.2.2.2, p.32 of our Addendum of 7th March 2016). We believe that the April 2016 NICE appraisal committee agreed with this change. They agreed that it is preferable to use the treatment duration from Taverna et al (2012), [6] given that this is also the source for PFS and OS.

7.1.3 PAS for BOR

In addition, we include a further change, which has only a minor effect on the ICER for LEN+DEX vs. BOR. BOR has a patient access scheme in which Janssen, the manufacturer of BOR, refunds the drug costs for people whose disease does not respond after 4 cycles of treatment. Janssen previously estimated that it would refund 15% of the total costs of BOR. Celgene assume that the refund is claimed for 55% of eligible patients, giving an overall discount of 55% x 15% = 8.3%.

Before the April 2016 appraisal committee meeting, we ran a scenario analysis in which we assumed that the refund would be claimed by all eligible patients, giving an overall discount of 15%. In this case, our base-case ICER for LEN+DEX vs. BOR increased slightly, from £43,700 to £44,300 per QALY gained. At the April 2016 meeting, the appraisal committee preferred to assume a 15% discount for bortezomib.

Therefore, we assume the full 15% discount in our revised base case.

7.1.4 Equal mortality post progression

We also consider a scenario analysis in which we assume equal mortality post progression. The rationale for this analysis is discussed in Section 7.3, p27.

7.1.5 ICERs

The resulting ICERs for LEN+DEX vs. MP are given in Table 4 and for LEN+DEX vs. BOR in Table 5.

In addition to factors discussed at the start of Section 7, p23, we repeat that all ICERs are highly uncertain, in part because:

- The underlying clinical data is not randomised.

- The quality of the clinical data used to inform PFS and OS for BOR and MP is extremely low.
- The nature of subsequent treatments is uncertain.

For LEN+DEX vs. BOR, our preferred scenarios are 1 & 2 & 3 & 4 (ICER £44,000 per QALY) and 1 & 3 & 4 & 6 (ICER £76,000 per QALY) shown in bold in Table 5.

Table 4. Impact on the ICER for LEN+DEX vs. MP of additional analyses undertaken by PenTAG

	ICER (£/QALY) LEN+DEX vs. MP
Celgene 2016 model	**£24,000**
1: No 3rd- or 4th-line treatment costs	£37,000
2: OS for MP with Celgene 3rd-line LEN adjustment removed	£19,000
3: Equal mortality between treatment arms after progression	£36,000
1 & 2	**£26,000**

Key: DEX, dexamethasone; LEN, lenalidomide; MP, melphalan plus prednisolone; OS, overall survival.

Table 5. Impact on the ICER for LEN+DEX vs. BOR of additional analyses undertaken by PenTAG

	ICER (£/QALY) LEN+DEX vs. BOR
Celgene 2016 model	**£20,000**
1: No 3rd- or 4th-line treatment costs	£36,000
2: Reduce mean duration of BOR from 6.6 to 3.8 treatment cycles	£29,000
3: OS for BOR with Celgene 3rd-line LEN adjustment removed	£20,000
4: BOR PAS discount	£21,000
1 & 2	£45,000
1 & 3	£35,000
2 & 3	£28,000
1 & 2 & 3 & 4	**£44,000**
5: Equal mortality between treatment arms after progression	£68,000
2 & 5	£84,000
2 & 4 & 5	£86,000
6: Reduce mean duration of BOR from 6.6 to 5 cycles, extend PFS (based on Reyal 2016)	£46,000
1 & 3 & 4 & 6	**£76,000**
5 & 6	>£1,000,000

Key: BOR, bortezomib; DEX, dexamethasone; LEN, lenalidomide; OS, overall survival; PAS, patient access scheme

7.2 Scenario assuming PFS and OS for MP and BOR equal to DEX

When PFS and OS for MP is set equal to PFS and OS for DEX, the ICERs given various changes are given in the table below.

Our base case lies between Scenarios 1 & 2 & 3 and 1 & 2 & 3 & 4, with corresponding ICER between £34,000 and £90,000 per QALY (Section 7.4, p28). We are unable to give a precise value because we are unsure of OS for DEX 2nd-line (Section 2.2, p13).

Table 6. Impact on the ICER for LEN+DEX vs. MP of additional analyses undertaken by PenTAG assuming PFS and OS for MP equal to DEX

	ICER (£/QALY) LEN+DEX vs. MP
Celgene current analysis	**£20,000**
1: Error MP acquisition cost (Section 2.3, p15)	£46,000
2: Error PFS and OS for MP set equal to values for DEX (Section 2.4, p16)	£15,000
3: DEX PFS tail shortened	£23,000
1 & 2	£32,000
1 & 2 & 3	**£34,000**
4: DEX OS longer tailed adjusting for line of treatment	£35,000
1 & 2 & 3 & 4	**£90,000**
5: Equal mortality between treatment arms after progression	£22,000
1 & 2 & 5	£38,000
1 & 2 & 3 & 5	£36,000

Key: DEX, dexamethasone; LEN, lenalidomide; MP, melphalan plus prednisolone.

As discussed above, Celgene do not consider a scenario in which PFS and OS for BOR is set equal to PFS and OS for DEX, because they no longer consider BOR to be a valid comparator. Contrary to this, and as discussed above, we do believe that it remains a valid comparator.

Although clinical opinion suggests that BOR is more effective than DEX, we now present an analysis in which PFS and OS for BOR is set equal to PFS and OS for DEX. This then gives lower bounds for the ICERs for LEN+DEX vs. BOR, as shown in the table below.

The technical description of the changes to Celgene's model is as follows. Despite modelling the BOR arm, MP is selected as the comparator in the "Controls" sheet. Also in this sheet, the option "Set pre-progression and post progression equal to DEX" is set to "Yes". In sheet "PenTAG base case", "Model BOR" is set to "Yes". There is a very minor loss of accuracy in mean QALYs for BOR as disutilities due to adverse events are assumed equal to those for MP, rather than BOR. The mean duration of BOR treatment is unchanged as either 6.6 or

3.8 treatment cycles, despite the change in BOR PFS. Also, the minor correction to the BOR PAS (Section 7.1.3, p24) is not considered.

Table 7. Impact on the ICER for LEN+DEX vs. BOR of additional analyses undertaken by PenTAG assuming PFS and OS for BOR equal to DEX

	ICER (£/QALY) LEN+DEX vs. BOR
Celgene current analysis	**Not modelled**
PenTAG amendments to model BOR arm	>£35,000
1: Error PFS and OS for MP/BOR set equal to values for DEX (Section 2.4, p16)	>£25,000
2: DEX PFS tail shortened	>£32,000
3: Reduce mean duration of BOR from 6.6 to 3.8 treatment cycles	>£48,000
1 & 2	>£23,000
1 & 2 & 3	**>£31,000**
4: DEX OS longer tailed adjusting for line of treatment	>£65,000
1 & 2 & 3 & 4	**>£83,000**
5: Equal mortality between treatment arms after progression	>£30,000
1 & 5	>£32,000
1 & 2 & 5	>£27,000
1 & 2 & 3 & 5	>£34,000

Key: BOR, bortezomib; DEX, dexamethasone; LEN, lenalidomide.

7.3 Scenario assuming no survival benefit after stopping treatment

As explained in Section 6, p21, the April 2016 NICE appraisal committee requested the following from Celgene:

"*a scenario analysis using a hazard ratio of 1 after stopping treatment, because there is no evidence of an ongoing survival benefit beyond that time. Please present this analysis for both comparators (bortezomib retreatment and melphalan)*."

The motivation for this request is not clear to us. The statement suggests that there is evidence of a survival benefit of LEN+DEX over BOR or MP whilst patients are treated with LEN+DEX. However, we are not aware of such evidence.

Celgene replied that they could not implement the request from the committee because of the structure of their model. Instead, they assumed equal PPS across arms.

We find that adjustment of Celgene's model to give equal rates of mortality after treatment cessation would indeed be difficult. Instead, we present the following approximation, which is relatively simple to implement. We assume that the proportion of patients still alive at treatment cessation is equal across treatment arms (e.g. 100%), and we ignore differences

due to discounting in costs and benefits after treatment cessation due to differences in timing of events. Then, we artificially forces PFS to equal time on treatment, and calculate the ICERs from the total costs and QALYs incurred up to progression (set to treatment discontinuation). The resulting ICERs are given in Table 8.

Table 8. Impact on the ICERs for LEN+DEX assuming no survival benefit after treatment cessation

	ICER (£/QALY) LEN+DEX vs. BOR	ICER (£/QALY) LEN+DEX vs. MP
Celgene 2016 model	**£20,000**	**£24,000**
1: Model up to treatment discontinuation	£45,000	£59,000
2: Reduce mean duration of BOR from 6.6 to 3.8 treatment cycles (Taverna 2012)	£29,000	n/a
1 & 2	**£49,000**	**£59,000**

Key: DEX, dexamethasone; LEN, lenalidomide; MP, melphalan plus prednisolone.

For completeness, we also estimates ICERs given no difference in survival after progression. Similar to the previous analysis where mortality after treatment cessation is assumed equal between treatments, in order to estimate ICERs given no difference in survival after progression, we assumed the same proportion of patients are alive on progression across treatment arms, and we ignore any differential effects of discounting on costs and QALYs in PD. The resulting ICERs are given in Section 7.1, p23 and Section 7.2, p26 above.

7.4 PenTAG base case

7.4.1 LEN+DEX vs. MP

For LEN+DEX vs. MP, our base case analysis, given in Section 7.2, p26, assumes PFS and OS for MP equal to DEX from MM-090 and MM-010. The advantages of this approach are:

- The comparison between LEN+DEX and MP is based on randomised data.
- There is evidence that the effectiveness of MP is similar to DEX (Section 2, p9).
- A large proportion, 48%, of patients in the MP arm are subsequently treated with LEN, which reflects current clinical practice.

One disadvantage is that we are not directly using any clinical effectiveness evidence for MP. Instead, we are assuming the effectiveness of MP equals DEX, based on the study Facon et al. (2006).[2]

The corresponding ICERs are given in Table 6, p26. Our base case ICER for LEN+DEX vs. MP lies between £34,000 and £90,000 per QALY, depending on OS for DEX 2nd-line.

These values are greater than the £26,000 per QALY (Table 4, p25) assuming the clinical effectiveness of MP is taken from Petrucci et al (1989).[7] However, the values are consistent with the ICER of £36,000 per QALY assuming equal mortality between treatment arms after

progression (Table 4, p25) and £59,000 per QALY assuming no survival benefit after treatment cessation (Table 8, p28).

7.4.2 LEN+DEX vs. BOR

We find it harder to identify a base case scenario with which to compare LEN+DEX vs. BOR.

Assuming PFS and OS for BOR equal to DEX from MM-090 and MM-010 has the following advantages:

- The comparison between LEN+DEX and BOR is based on randomised data.
- A large proportion of patients, 48%, in the BOR arm are subsequently treated with LEN, which reflects current clinical practice.

Disadvantages are:

- We have no evidence for the relative effectiveness of BOR vs. DEX.
- We do not directly using any clinical effectiveness evidence for BOR. Instead, we assume BOR and DEX are equally effective.

However, given that it is likely that BOR is more effective than DEX, we can present lower bounds for the resulting ICERs. Our base case lies between the following two scenarios in Table 7, p27: Scenario 1 & 2 & 3 (ICER > £31,000 per QALY) and Scenario 1 & 2 & 3 & 4 (ICER > £83,000 per QALY).

This is consistent with the ICERs of:

- £44,000 per QALY taking BOR effectiveness from Taverna et al (2012) (Table 5, p25).
- £76,000 per QALY taking BOR effectiveness from Taverna et al (2012) and Reyal et al (2016) (Table 5, p25).
- £68,000 per QALY assuming equal mortality between treatment arms after progression, survival from Taverna et al (2012) (Table 5, p25).
- >£34,000 per QALY assuming equal mortality between treatment arms after progression, PFS set to DEX (Table 7, p27).
- £49,000 per QALY assuming no survival benefit after treatment cessation (Table 8, p28).

References

1. Dimopoulos MA, Chen C, Spencer A, Niesvizky R, Attal M, Stadtmauer EA, et al. Long-term follow-up on overall survival from the MM-009 and MM-010 phase III trials of lenalidomide plus dexamethasone in patients with relapsed or refractory multiple myeloma. Leukemia. 2009;23(11):2147-52.
2. Facon T, Mary JY, Pegourie B, Attal M, Renaud M, Sadoun A, et al. Dexamethasone-based regimens versus melphalan-prednisone for elderly multiple myeloma patients ineligible for high-dose therapy. Blood. 2006;107(4):1292-8.
3. Dimopoulos M, Spencer A, Attal M, Prince HM, Harousseau JL, Dmoszynska A, et al. Lenalidomide plus dexamethasone for relapsed or refractory multiple myeloma. N Engl J Med. 2007;357(21):2123-32.
4. Weber DM, Chen C, Niesvizky R, Wang M, Belch A, Stadtmauer EA, et al. Lenalidomide plus dexamethasone for relapsed multiple myeloma in North America. N Engl J Med. 2007;357(21):2133-42.
5. Reyal Y, Popat R, Cheesman S, Rismani A, D'Sa S, Rabin N, et al. Real world experience of bortezomib re-treatment for patients with multiple myeloma at first relapse. Br J Haematol. 2016.
6. Taverna C, Voegeli J, Trojan A, Olie RA, von Rohr A. Effective response with bortezomib retreatment in relapsed multiple myeloma--a multicentre retrospective survey in Switzerland. Swiss Med Wkly. 2012;142:w13562.
7. Petrucci MT, Avvisati G, Tribalto M, Cantonetti M, Giovangrossi P, Mandelli F. Intermediate-dose (25 mg/m2) intravenous melphalan for patients with multiple myeloma in relapse or refractory to standard treatment. Eur J Haematol. 1989;42(3):233-7.

Celgene Ltd
1 Longwalk Road
Stockley Park, Uxbridge
UB11 1DB, UK
Telephone +44 (0)208 831 8620
Facsimile +44 (0)208 831 8301

Dr Melinda Goodall,
Associate Director – Committee B,
National Institute for Health and Care Excellence
10 Spring Gardens | London SW1A 2BU | United Kingdom
Tel: 44 (0)20 7045 2248 | Fax: 44 (0)20 7061 9830

Dear Dr Goodall,

RE: **Multiple myeloma - lenalidomide (post bortezomib) (part review TA171) [ID667]**

Having reviewed the ERG report we have identified factual inaccuracies as a result of misinterpretation of the model and the use of incorrect and out-of-date data cuts. We would like to thank NICE for the opportunity to provide clarification prior to the Committee meeting.

Of the four main 'errors' the ERG identify, we believe that they are incorrect on three of them. This, in combination with additional issues, renders the scenario analyses provided within the ERG addendum uninformative to the decision problem.

Although correcting a number of these misinterpretations disadvantages Celgene, we felt it was important to provide the Committee with accurate information for decision making. The below document provides details of the inaccuracies within the addendum.

Yours Sincerely,

[redacted]

Factual inaccuracies within the ERG report

The 'Errors' Highlighted by the ERG

Section 2.1. – PFS for Dex too long tailed

We believe based upon the description in the ERG report that Figure 2 (reproduced below) uses PFS KM data from the 2005 datacut from Dimopolous et al. Within the analyses we provided we use the later datacut (2008) from the MM009 / MM010 trials which provides additional data with which to more accurately model PFS.

Figure 1 – Presentation of Figure 2. PFS and OS for DEX from MM-009 and MM-010 RCTs vs. Celgene model fit from the ERG report

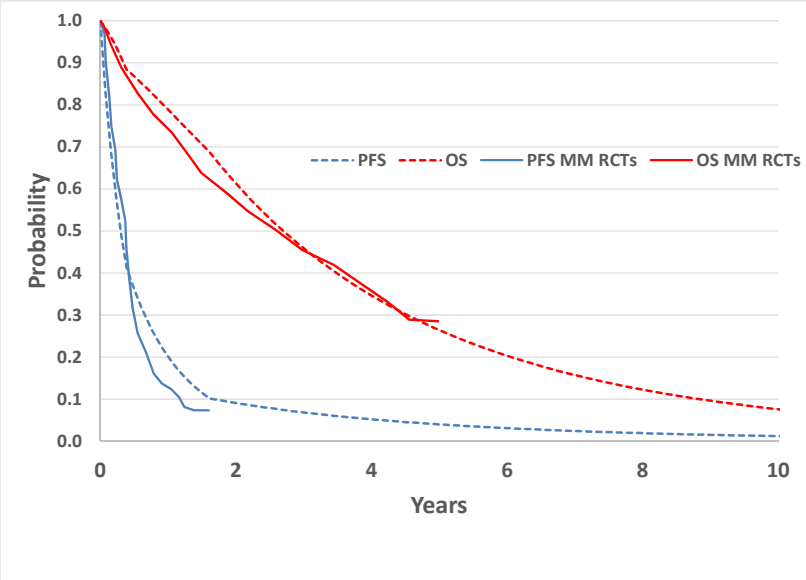
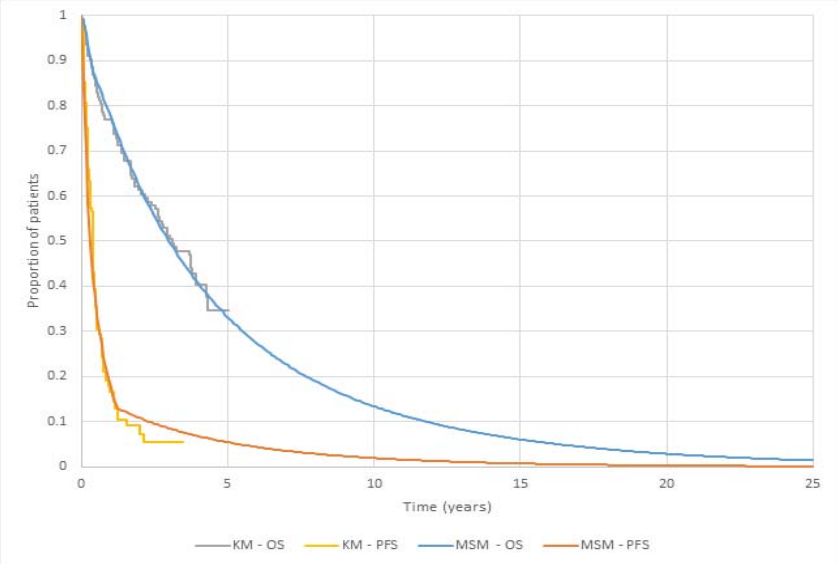


Figure 2 – The fitted KM data using the 2008 datacut



When the PFS KM data from the 2008 datacut is compared to the fitted model for dexamethasone (see above) the model provides a better fit to the observed KM data.

It can be seen that the KM using the 2008 cut provides approximately 2 years' worth of additional information compared to the 2005 cut.

It should be noted that the KM data supplied here is for patients receiving 2nd line lenalidomide only. This compares to the fitted curves which were produced using the MSM procedure as detailed in our previous response:

1. Data from all lines of treatment is used to fit the MSM model
2. A covariate is included for line of treatment (1 vs 2+ prior lines)
3. When we produce the survival curves using the corrected group prognosis method we only include patient information for patients at 2nd line
4. This produces fitted survival curves for 2nd line patients

Section 2.2 – Celgene DEX PFS and OS data incorrect line of treatment

This is a misunderstanding on how the model operates. As detailed above, the model uses data for all lines and then applies a covariate for line of treatment. The corrected group prognosis method is used to predict the final survival curves based upon the covariate adjusted data giving 0 weighting to anyone not at 2nd line (i.e. predicts for 2nd line patients only).

When considering the ERG exploratory analysis, as previously recognised by both the ERG and Committee, there is uncertainty associated with the calculation of hazard ratios from medians and additional uncertainty created in the assumption that the treatment effect between lenalidomide and dexamethasone is constant across lines of therapy. The addition of this uncertainty is unnecessary in this case as patient level data exists with which to project outcomes.

Section 2.3 – Error in cost acquisition of MP

The ERG are correct. We'd like to note this only applies when DEX data is used.

Section 2.4 – OS for DEX not used correctly in the model

When DEX data is used for MP the model is using the correct data, however, as the ERG rightly point out, the PPS data is not applied correctly when DEX is selected as a comparator within the model.

This was due to lack of inclusion of the comparator being selected as DEX in the statement which determines which PPS data is used in the comparator arm included in the MSM_2L (DEX):

For example cell X7 reads

"=IF(OR(MP_pps_DEX="Yes",MP_equal_to_dex="Yes"),ET7,MSM_2L!X7)"

As can be seen from the formula (and explained above), when the comparator is selected to use PPS data from DEX (as was the case when comparison is conducted vs MP and BOR) then the correct data was used. This means that the ERG are incorrect in their conclusion that new analyses are required when predicting MP outcomes using DEX PPS.

We believe this probably resulted from us providing incorrect data for DEX in Table 1 of our response to the additional questions from NICE. In response to ERG comments we have now also included information both with and without subsequent use of LEN and table 1 should read as follows:

Table 1 – Clinical Outcomes from the model compared to DEX

	With no subsequent LEN				With subsequent LEN			
	PF LY	PP LY	LY	QALYs	PF LY	PP LY	LY	QALYs
DEX	0.99	3.87	4.86	2.77	0.99	3.87	4.86	2.77
MP	0.43	0.72	1.14	0.78	0.43	2.73	3.15	1.88
BORT	1.87	0.83	2.7	1.86	1.97	1.71	3.69	2.37
Rd	2.96	2.91	5.87	3.55	2.96	2.91	5.87	3.55

It should be noted that none of the ICERs that were produced used DEX as a comparator therefore the error identified does not impact any of the ICER information submitted.

Additional information

Funding of bortezomib retreatment

The following statement within Section 1 of the report is incorrect:

*"TA129 (which in 2007, NICE recommended BOR **retreatment** for progressive multiple myeloma for people whose multiple myeloma has relapsed for the first time after having one treatment, and who have had a bone marrow transplant, unless it is not suitable for them)"*

TA129 in fact recommended BOR monotherapy based upon data for BOR used for the first time (i.e. not retreatment).

Additionally, we now have information demonstrating that use of BOR retreatment is not allowed by NHS England via TA129 (provided as an Appendix for the Committee's reference).

Bortezomib information from the Reyal trial

In addition to BOR no longer being a funded comparator, the additional information from this trial is not seen as relevant to this appraisal. The study has been conducted in a transplant eligible population which is outside of the scope and 11 of the 23 patients in the study received a 2nd transplant in the retreatment arm of the study. This is likely to considerably bias outcomes and reduce DoT as the SPC states that patients should receive 4 cycles for induction prior to transplant.

Approach to modelling equal effectiveness post second-line treatment

We agree that the approach presented by the ERG for this scenario is at best an approximation. We would also like to restate that we do not consider this scenario to be relevant; due to the immunomodulatory properties of LEN, clinical expectation is that benefit is likely to continue to be seen post treatment.

The main issue with the approach suggested by the ERG is that it ignores the evidence presented which clearly demonstrates that as would be expected TTF is

less than PFS (given that patients can discontinue for reasons other than progression or death). This artificially inflates the cost of treatment with LEN.

Impact on the ICER

The impact on the ICER of the amendments detailed above is provided here. BOR results are not provided, as BOR retreatment should no longer be considered a valid comparator as it cannot be accessed.

Table 2 – ICERs from scenarios

	Description	ICER vs MP	Comparator		
			Total LYs	Total PF LYs	Total PP LYs
1	**Base case** (using data from the Petrucci 1989 trial for MP, Taverna 2012 for BOR 100% use of LEN at 3L)	£23,572	3.15	0.43	2.73
2	**No use of LEN 3L** (using data from the Petrucci 1989 trial for MP, Taverna 2012 for BOR 0% use of LEN at 3L)	£23,742	1.14	0.43	0.72
3	**DEX data used for PPS** (with MP costing error fixed, assumes 48% use of LEN at 3L)	£39,545	4.32	0.43	3.89
4	**DEX data used for PFS and PPS** (with MP costing error fixed, assumes 48% use of LEN at 3L)	£45,897	4.86	0.99	3.87
5	**DEX data used for PFS; MP data used for OS*** (with MP costing error fixed, 0% use of LEN at 3L)	£24,439	1.14	0.99	0.16
6	**DEX data used for PFS; MP data used for OS*** (with MP costing error fixed, 100% use of LEN at 3L)	£24,472	3.15	0.99	2.17
7	**PPS assumed equal across arms, subsequent LEN** (no PPS benefit for LEN, 100% use of LEN at 3L)	£22,172	3.13	0.43	2.71
8	**PPS assumed equal across arms, no subsequent LEN** (no PPS benefit for LEN, 0% use of LEN at 3L)	£35,830	3.13	0.43	2.71

*New analyses

Therefore the ICER range against MP is between £24,000 and £46,000. We have also provided additional scenarios where the PFS of DEX is applied with the PPS of MP to highlight that the variability in the ICER range is dependent upon the proportion of patients accessing LEN at 3L and effectiveness of LEN at 3L.

The clinical and cost-effectiveness of lenalidomide for people who have received at least one prior therapy with bortezomib (partial review of TA171)

A critique of the submission from Celgene

4th October 2016 Addendum

Produced by

Peninsula Technology Assessment Group (PenTAG)
University of Exeter Medical School
South Cloisters
St Luke's Campus
Heavitree Road
Exeter
EX1 2LU

Authors

Martin Hoyle, Associate Professor, PenTAG, University of Exeter
Nicola Huxley, Research Fellow, PenTAG, University of Exeter

Correspondence to

Martin Hoyle
Peninsula Technology Assessment Group (PenTAG)
University of Exeter Medical School
South Cloisters
St Luke's Campus
Heavitree Road
Exeter
EX1 2LU
m.w.hoyle@exeter.ac.uk

Date completed

4th October 2016

Source of funding

This report was commissioned by the NIHR HTA Programme as project number 13/07/01.

Declared competing interests of the authors

None

Acknowledgments

We acknowledge the excellent administrative support of Sue Whiffin and Jenny Lowe (University of Exeter) and Linda Long (University of Exeter) for checking the data reported from Reyal et al. (2016). We also acknowledge support and guidance given by Professor Chris Hyde (University of Exeter) relating to the critique of the company's clinical evidence. We also acknowledge the clinical advice provided by Claudius Rudin (Royal Devon and Exeter Hospital) and the support of Professor Rob Anderson (University of Exeter) who provided cover for the Director role.

Rider on responsibility for report

The views expressed in this report are those of the authors and not necessarily those of the NIHR HTA Programme. Any errors are the responsibility of the authors.

This report should be referenced as follows

Hoyle M. & Huxley N. The clinical and cost-effectiveness of lenalidomide for people who have received at least one prior therapy with bortezomib (partial review of TA171): A Single Technology Appraisal (October 2016 Addendum). Peninsula Technology Assessment Group (PenTAG), 2016.

Contributions of authors

Martin Hoyle	Contributed to the critique of Celgene's economic evaluation, developed the PenTAG base case and contributed to the writing and editing of this report.
Nicola Huxley	Provided project management, led the checking of the economic model submitted by Celgene. Contributed to the writing and editing of the report.

Background

This addendum presents a response to comments the ERG received from Celgene on 3rd October 2016, as part of the NICE Single Technology Appraisal ID667 “Lenalidomide for the treatment of multiple myeloma in people who have received at least one prior therapy with bortezomib (partial review of TA171)”.

Celgene provided comments in response to our most recent critique of their submission, dated 14th September 2016. This new addendum both provides a response to Celgene’s comments and, where appropriate, updates our 14th September 2016 critique.

Due to time constraints, this document does not comment upon any new ICERs Celgene have provided.

1 PFS for DEX

In our Addendum dated 14th September 2016, we said we believed that Celgene has assumed a tail for PFS for DEX that was too long.

In their recent response, Celgene state that we compared their curve fit against a 2005 data cut from the MM RCTs of LEN+DEX vs. DEX. This is correct. They also claim that a 2008 data cut was available with longer follow up. We accept this.

By comparing our fit and Celgene's to the 2008 data cut (Figure 1 and Figure 2), we still consider our curve fit (Figure 2) to be more appropriate than that of Celgene (Figure 1). However, we also note that the ICERs for LEN+DEX vs. MP and LEN+DEX vs. BOR in Section 4.2 (Scenario assuming PFS and OS for MP and BOR equal to DEX), are rather insensitive to the choice of fit.

Figure 1. PFS DEX Celgene curve fit

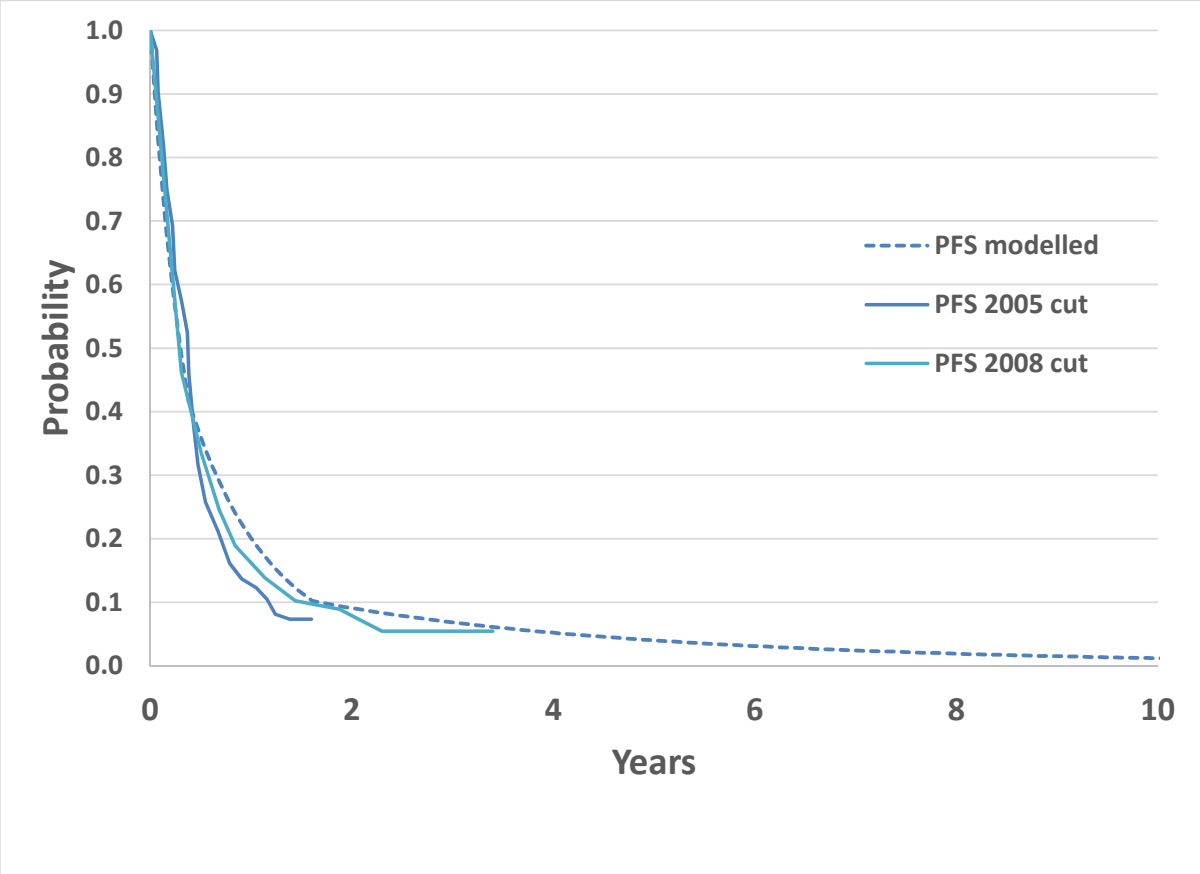
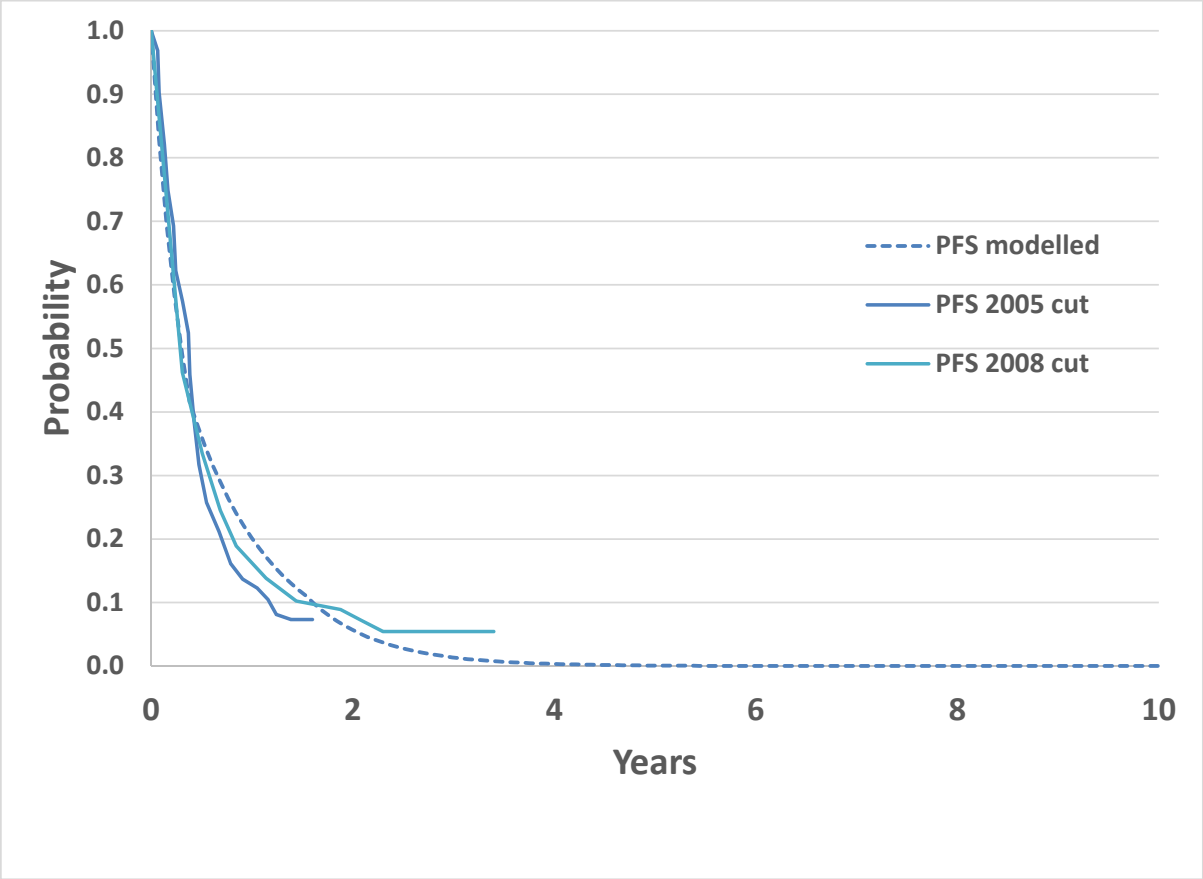


Figure 2. PFS DEX PenTAG curve fit



2 OS for DEX

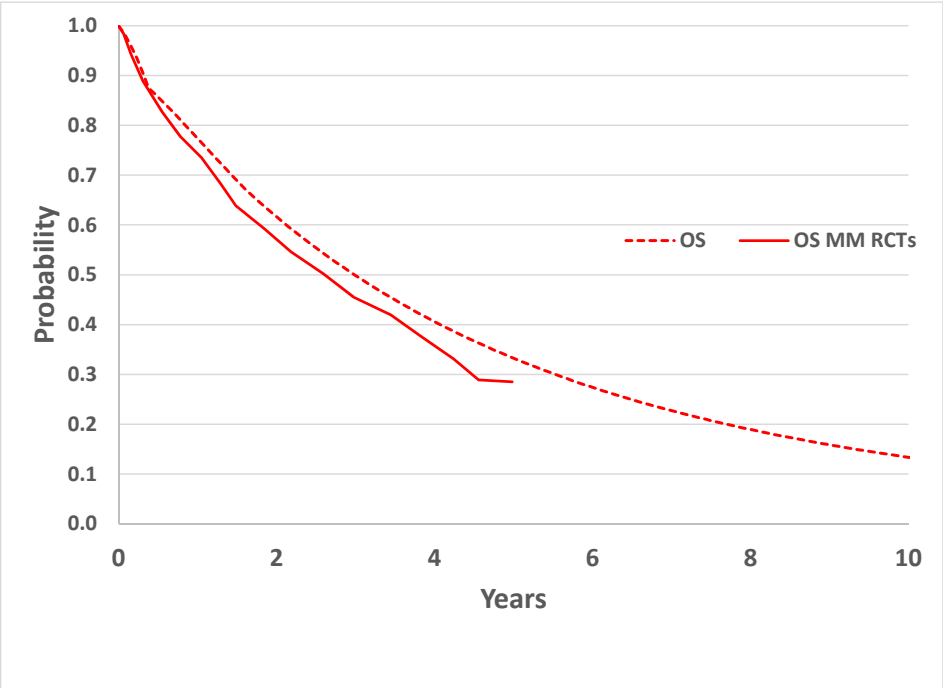
In our Addendum of 14th September 2016, we claimed that OS for DEX had not been implemented correctly in Celgene's model for the scenario in which OS for MP is set equal to that for DEX from the MM RCTs.

We also claimed that Celgene had used OS for DEX for 2nd and 3rd-lines combined from the MM RCTs, whereas we need to consider 2nd line data only. We therefore presented a scenario analysis in which we attempted to correct for this.

In response, Celgene now concede that when DEX is used as a comparator, they model OS for DEX incorrectly. However, they claim that in the scenario analysis in which OS for MP is set equal to that for DEX, there is no error, i.e. OS for MP is then correctly set equal to that for DEX. They also disagree with our assertion that they fit OS for DEX to 2nd- and 3rd-line data combined from the MM RCTs, claiming that they fit to 2nd-line data.

We are now prepared to accept all of Celgene's responses above, subject to the important caveat that we have no way to validate their modelled OS for 2nd-line DEX. We find that their modelled OS for 2nd-line DEX is slightly longer tailed than the Kaplan-Meier for 2nd and 3rd-line combined from the MM RCTs (Figure 3). However, we can only check their modelled OS for DEX if we have access to the Kaplan-Meier data for 2nd-line use from the MM RCTs (as we do for the LEN+DEX arm, see Fig. 4 our Addendum of 14th Sept 2016). Given this uncertainty, we retain our scenario analysis of estimating 2nd-line OS for DEX in Section 4.2.

Figure 3. OS DEX: Celgene fit for 2nd-line vs. MM RCT Kaplan-Meier data for 2nd and 3rd-lines combined



3 Reyal et al (2016) data on BOR retreatment

In our Addendum of 14th Sept, in some scenario analyses, we incorporated some of the data from Reyal et al (2016).

In response, Celgene say that this data is not relevant, because about half of the 23 patients had a transplant and such patients are not within the scope of the current appraisal.

Although we do not have a sense for the impact of this deviation from the scope, we are prepared to accept Celgene's argument.

4 PenTAG revised estimates of cost-effectiveness

4.1 Scenario assuming PFS and OS for MP and BOR taken from single arm studies

Our opinion on these analyses is unchanged from our Addendum of 14th Sept, except that we now no longer consider the scenario analyses using data from Reyal et al (2016).

4.2 Scenario assuming PFS and OS for MP and BOR equal to DEX

When PFS and OS for MP is set equal to PFS and OS for DEX, the ICERs given various changes are given in the table below.

These tables are to be compared with Tables 6 and 7 in our previous Addendum. We have now removed our previous correction "*Error PFS and OS for MP set equal to values for DEX (Section 2.4, p16)*". We repeat our concern that we have been unable to validate Celgene's modelled OS for 2nd-line DEX.

Assuming Celgene's modelled OS for 2nd-line DEX is correct, our base case ICER for LEN+DEX vs. MP is now £48,000 per QALY, and for LEN+DEX vs. BOR, >£44,000 per QALY.

Table 1. Impact on the ICER for LEN+DEX vs. MP of additional analyses undertaken by PenTAG assuming PFS and OS for MP equal to DEX

Scenario	ICER (£/QALY) LEN+DEX vs. MP
Celgene current analysis	**£20,000**
1: Error MP acquisition cost	£46,000
2: DEX PFS tail shortened	£23,000
1 & 2	**£48,000**
3: DEX OS longer tailed PenTAG adjustment for line of treatment	£35,000
1 & 2 & 3	£90,000
4: Equal mortality between treatment arms after progression	£22,000
1 & 4	£37,000
1 & 2 & 4	£35,000

Key: DEX, dexamethasone; LEN, lenalidomide; MP, melphalan plus prednisolone.

Table 2. Impact on the ICER for LEN+DEX vs. BOR of additional analyses undertaken by PenTAG assuming PFS and OS for BOR equal to DEX

Scenario	ICER (£/QALY) LEN+DEX vs. BOR
Celgene current analysis	**Not modelled**
PenTAG amendments to model BOR arm	>£35,000
1: DEX PFS tail shortened	>£32,000
2: Reduce mean duration of BOR from 6.6 to 3.8 treatment cycles	>£48,000
1 & 2	**>£44,000**
3: DEX OS longer tailed PenTAG adjustment for line of treatment	>£65,000
1 & 2 & 3	>£83,000
4: Equal mortality between treatment arms after progression	>£30,000
1 & 4	>£27,000
1 & 2 & 4	>£33,000

Key: BOR, bortezomib; DEX, dexamethasone; LEN, lenalidomide.

4.3 Scenario assuming no survival benefit after stopping treatment

Our opinion on these analyses is unchanged from our Addendum of 14th Sept.

4.4 PenTAG base case

As in our Addendum of 14th September 2016, we still consider our base case for LEN+DEX vs. MP and LEN+DEX vs. BOR as corresponding to the scenarios assuming PFS and OS for MP and BOR equal to DEX, i.e. Section 4.2 above.