

Single Technology Appraisal

Lenalidomide for the maintenance treatment of multiple myeloma after autologous stem cell transplantation [ID475]

Committee Papers

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

SINGLE TECHNOLOGY APPRAISAL

Lenalidomide for the maintenance treatment of multiple myeloma after autologous stem cell transplantation [ID475]

Contents:

The following documents are made available to consultees and commentators:

- 1. Response to consultee, commentator and public comments on the Appraisal Consultation Document (ACD)**
- 2. Comments on the Appraisal Consultation Document from Celgene UK - A BMS company**
 - Addendum 1
 - Addendum 2
- 3. Consultee and commentator comments on the Appraisal Consultation Document** from:
 - a. Myeloma UK & response appendix
- 4. Comments on the Appraisal Consultation Document received through the NICE website**
- 5. Evidence Review Group critique of company comments on the ACD**
- 6. Evidence Review Group critique – factual accuracy check**

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

Lenalidomide maintenance treatment after an autologous stem cell transplant for newly diagnosed multiple myeloma

Single Technology Appraisal

Response to consultee, commentator and public comments on the Appraisal Consultation Document (ACD)

Type of stakeholder:

Consultees – Organisations that accept an invitation to participate in the appraisal including the companies, national professional organisations, national patient organisations, the Department of Health and Social Care and the Welsh Government and relevant NHS organisations in England. Consultees can make a submission and participate in the consultation on the appraisal consultation document (ACD; if produced). All non-company consultees can nominate clinical experts and/or patient experts to verbally present their personal views to the Appraisal Committee. Company consultees can also nominate clinical experts. Representatives from NHS England and clinical commissioning groups invited to participate in the appraisal may also attend the Appraisal Committee as NHS commissioning experts. All consultees have the opportunity to consider an appeal against the final recommendations, or report any factual errors, within the final appraisal document (FAD).

Clinical and patient experts and NHS commissioning experts – The Chair of the Appraisal Committee and the NICE project team select clinical experts and patient experts from nominations by consultees and commentators. They attend the Appraisal Committee meeting as individuals to answer questions to help clarify issues about the submitted evidence and to provide their views and experiences of the technology and/or condition. Before they attend the meeting, all experts must either submit a written statement (using a template) or indicate they agree with the submission made by their nominating organisation.

Commentators – Commentators can participate in the consultation on the ACD (if produced), but NICE does not ask them to make any submission for the appraisal. Non-company commentator organisations can nominate clinical experts and patient experts to verbally present their personal views to the Appraisal Committee. Commentator organisations representing relevant comparator technology companies can also nominate clinical experts. These organisations receive the FAD and have opportunity to report any factual errors. These organisations include comparator technology companies, Healthcare Improvement Scotland any relevant National Collaborating Centre (a group commissioned by NICE to develop clinical guidelines), other related research groups where appropriate (for example, the Medical Research Council and National Cancer Research Institute); other groups such as the NHS Confederation, the NHS Commercial Medicines Unit, the Scottish Medicines Consortium, the Medicines and Healthcare Products Regulatory Agency, the Department of Health and Social Care, Social Services and Public Safety for Northern Ireland).

Public – Members of the public have the opportunity to comment on the ACD when it is posted on the Institute's web site 5 days after it is sent to consultees and commentators. These comments are usually presented to the appraisal committee in full, but NICE reserves the right to summarise and edit comments received during consultations, or not to publish them at all, where in the reasonable opinion of NICE, the comments are voluminous, publication would be unlawful or publication would be otherwise inappropriate.

Please note: Comments received in the course of consultations carried out by NICE are published in the interests of openness and transparency, and to promote understanding of how recommendations are developed. The comments are published as a record of the submissions that NICE has received, and are not endorsed by NICE, its officers or advisory committees.

Comment number	Type of stakeholder	Organisation name	Stakeholder comment Please insert each new comment in a new row	NICE Response Please respond to each comment
1	Consultee (company)	Celgene	<p><i>The company should have presented evidence from other trials of lenalidomide maintenance treatment</i></p> <p>A complete and thorough presentation of clinical effectiveness evidence from the CALGB-100104 and GIMEMA trials of lenalidomide maintenance treatment, in addition to further details from Myeloma XI, are included in an addendum to this response.</p> <p>The review of the data showed that</p> <ol style="list-style-type: none"> 1. CALGB and Myeloma XI were well conducted studies and included valid comparisons of lenalidomide maintenance compared with placebo / observation. Both studies showed a statistically significant benefit on overall survival and progression-free survival with lenalidomide maintenance. 2. CALGB and Myeloma XI baseline characteristics and populations were presented in detail. Whilst the two studies differed in the baseline distribution of ISS scores, and somewhat by gender and age, the differences were potentially due to methods used in data collection and definitions, as well as some differences in prior therapies (induction) received in the studies. The review identified factors important for consideration that informed the subsequent matched-adjusted analyses of CALGB and Myeloma XI. 3. The methods and study design used for GIMEMA are affected by an error that caused the study to provide a biased estimate of the treatment efficacy with lenalidomide maintenance compared with placebo and specifically, with respect to the Appraisal decision problem. <p>Overall, the CALGB and Myeloma XI were found to be robust studies, largely comparable and pertinent with the decision problem.</p>	<p>Thank you for your comment. The FAD has been updated to reflect that the committee saw evidence from all trials of lenalidomide maintenance therapy that met the systematic literature review criteria (see section 3.4 of the FAD).</p>
2	Consultee (company)	Celgene	<p><i>The company's model structure does not allow assumptions about subsequent treatments to be explored.</i></p> <p><i>The committee concluded that the company's model structure had limitations. It also concluded that there was likely to be uncertainty around the cost-effectiveness estimate because assumptions about the effects of subsequent therapies on survival could not be fully explored.</i></p>	<p>Thank you for your comment. The committee considered the limitations associated with the partitioned survival analysis model structure (see section 3.6 of the FAD) and considered the different assumptions relating to costs of subsequent treatments (see sections 3.11, 3.12, and 3.13 of the</p>

Comment number	Type of stakeholder	Organisation name	Stakeholder comment Please insert each new comment in a new row	NICE Response Please respond to each comment
			<p>The model structure was pertinent with regards to the exploration of the impact of subsequent therapies.</p> <p>In the Addendum to this Response, it was shown that the subsequent therapies used in CALGB and Myeloma XI were largely similar, and both were reflective of clinical settings where monoclonal antibodies were not in use as they were not licensed yet. This is reflective of subsequent therapies in a world without Cancer Drugs Fund access.</p> <p>Extensive clinical validation was conducted on the scenarios included in the model, which confirmed that the distributions used are reflective of current practice without CDF funded therapies.</p> <p>Because of concordance between clinical efficacy incorporated in the model and subsequent therapies costed as part of the longer term follow up, the model structure is adequate to reflect valid subsequent therapies scenarios.</p> <p>The proportions of subsequent therapies in the model were tested in scenario analyses, using clinically validated proportions for second ASCT and lenalidomide subsequent use in further lines. The cost-effectiveness of lenalidomide maintenance was confirmed across a large range of likely scenarios.</p>	FAD).
3	Consultee (company)	Celgene	<p><i>The company's methods and rationale for pooling Myeloma XI and CALGB 100104 data, and adjusting for treatment switching, are unclear</i></p> <p><u>Pooling and adjustment methods</u> New statistical models were developed to account for potential differences in populations and study design between CALGB and Myeloma XI. Propensity scores weighting (PW) and matched-adjusted indirect comparison (MAIC) adjusted comparisons were developed. The new analyses provide a more extensive comparison of the two trials, which was used to inform an extensive set of cost-effectiveness scenarios in the model.</p> <p>All models used for matching produced highly concordant results and showed that the case for the two studies being in comparable populations remains strong.</p> <p>In the analyses, differences were identified in the distribution of some potential prognostic factors. Nevertheless, the matched-adjusted OS Kaplan-Meier for placebo remained stable in all analyses and showed a high degree of concordance with both the unadjusted CALGB Kaplan-Meier and the Kaplan-Meier for the observation arm in Myeloma XI.</p> <p>It is therefore unlikely that the prognosis for people in both placebo/observation</p>	Thank you for your comment. The committee considered the methods used to pool and adjust data from Myeloma XI and CALGB 100104 (see sections 3.7, 3.8 and 3.9 of the FAD).

Comment number	Type of stakeholder	Organisation name	Stakeholder comment Please insert each new comment in a new row	NICE Response Please respond to each comment
			<p>arms of the two studies may be impacted by factors other than those accounted for in the matching models.</p> <p>This would confirm the initial assumption that the populations in the two studies were not largely different with respect to baseline characteristics, and that the untreated population in the two studies has a similar prognosis.</p> <p>The robustness of the KM curve with respect to matching adjustment also constitutes indirect support to the difference between the lenalidomide arm in CALGB and that in Myeloma XI could be largely attributed in lenalidomide dosing. For this reason, we also extended the pooled model as part of the new analyses, using a treatment by trial interaction term as well as matching for population and trial design characteristics. The pooled model confirms the comparability of the studies and has the advantage of controlling for the difference in treatment dosage between CALGB and Myeloma XI (21/28 days and 28/28 days).</p> <p>The matched-adjusted analyses were robust to the methods employed, to the choice of matching variables and provide extensive validation for the comparability of Myeloma XI and CALGB; based on the large majority of the analyses and scenarios from the matched adjusted comparisons, the cost-effectiveness of lenalidomide maintenance remained robust and within acceptability ranges.</p> <p>Given all variations and methodological approaches considered, the cost-effectiveness of lenalidomide maintenance remained favourable.</p>	
4	Consultee (company)	Celgene	<p><i>The justification for using the rank preserving structural failure model should be provided</i></p> <p>An extensive presentation of methods, feasibility and results for the RPSFM approach used in the submission is presented.</p> <p>In brief, at a time when the CALGB study hit the superiority boundary in the early follow up (2 years), all patients initially randomised to placebo that had not progressed after ASCT (thus still fulfilling eligibility for maintenance) were offered to switch to active maintenance therapy.</p> <p>Of more than 100 patients in the placebo arm at the time who had not progressed, the majority chose to receive maintenance. The number of patients who did not switch was small and such that an analysis based on data from non-switchers would be possible.</p> <p>Furthermore, switching in CALGB was unrelated with disease outcomes and because it was conditional on patients not having progressed, the conditions for</p>	Thank you for your comment. The committee considered the methods used to adjust for treatment switching in the CALGB 100104 trial (see section 3.7 of the FAD).

Comment number	Type of stakeholder	Organisation name	Stakeholder comment Please insert each new comment in a new row	NICE Response Please respond to each comment
			<p>non-informative censoring would probably be violated.</p> <p>Extensive assessment of the common treatment effect assumption, on which RPSFM relies, was also conducted and presented.</p> <p>The RPSFM method was considered the most suitable approach for the reasons above.</p>	
5	Consultee (company)	Celgene	<p><i>Survival extrapolations should use Myeloma XI data as the main source of evidence but could be supplemented with CALGB 100104 data</i></p> <p>The model was extended to include a range of matched-adjusted analyses. Two statistical approaches were taken, propensity scores weights and MAIC; the former using patient level data for both studies and the latter matching CALGB to aggregate data from Myeloma XI. The methods are explained in detail in the Addendum to this response.</p> <p>The matched-adjusted analyses were incorporated into the cost-effectiveness model; the following scenarios were generated:</p> <ul style="list-style-type: none"> • Use of treatment effects from Myeloma XI until month 60, and thereafter using the treatment effects for the matched-adjusted CALGB extrapolations; • Use of treatment effects from Myeloma XI at all time; • Use of treatment effects from the revised pooled analysis of CALGB and Myeloma XI at all time points, incorporating covariates to control for study, treatment, and study-by-treatment interaction. <p>The aim of the scenarios was to assess the cost-effectiveness robustness to CALGB analyses.</p> <p>Statistical goodness of fit and clinical plausibility were used to identify the potentially relevant extrapolations in the model; however, it was preferred to present the results of cost-effectiveness for all clinically plausible distributions, with the exception of distributions that clearly were not appropriate.</p> <p>The revised estimates of cost-effectiveness show that lenalidomide maintenance is robustly cost-effective across all scenarios, for all plausible distributions, and</p>	<p>Thank you for your comment. The committee considered the methods used to pool and adjust data from Myeloma XI and CALGB 100104 (see sections 3.8 and 3.9 of the FAD).</p>

Comment number	Type of stakeholder	Organisation name	Stakeholder comment Please insert each new comment in a new row	NICE Response Please respond to each comment
			<p>for all matching and adjustment methods considered.</p> <p>Using estimates from the pooled model and using the Myeloma XI specific pooled effect also confirms the cost-effectiveness results, regardless of extrapolation distributions chosen.</p>	
6	Consultee (company)	Celgene	<p><i>The treatment effect of lenalidomide maintenance may wane over time and this should be included in the model</i></p> <p>The case for waning effects being likely assumptions for lenalidomide used in maintenance was explored. It was concluded that most common reasons to suspect waning of treatment effect are not likely to occur with maintenance.</p> <ol style="list-style-type: none"> 1. Maintenance with lenalidomide continues until disease progression or unacceptable toxicity. Capping rule for the duration of maintenance are not applicable 2. Non-compliance was unlikely with Myeloma XI. Maintenance with lenalidomide is spaced with 'off treatment' intervals such that tolerability if actively pursued. 3. In the case of prolonged non-compliance, it is unlikely that a patient would remain in a prolonged state of pre-progression, captured in PFS. 4. Myeloma XI and CALGB have long follow-up; PFS data in CALGB are mature, making extrapolation necessary from a late point in time in the model, reducing uncertainty. 5. The analysis of CALGB and Myeloma XI showed is robust evidence of the pertinence of proportional hazard which is not at odds with the assumption that treatment effect is decreasing. <p>For these reasons, and based on clinical opinion received, it is believed that waning of treatment effect would not be plausible for maintenance with lenalidomide.</p> <p>Nevertheless, scenario analyses were used to explore the impact of waning treatment effect, which showed that cost-effectiveness estimates are robust to most conservative treatment effect waning assumptions.</p>	Thank you for your comment. The committee considered whether a treatment waning effect should be included in the model (see section 3.10 of the FAD).
7	Consultee (company)	Celgene	<p><i>Costs of subsequent treatments are highly uncertain so scenarios should be presented</i></p> <p>Subsequent therapies in Myeloma XI and CALGB were assessed and tabulated for comparison, to verify the material similarity between subsequent treatment used in CALGB and Myeloma XI.</p> <ul style="list-style-type: none"> • Both studies show that most participants who progressed received subsequent therapies • The mix of subsequent therapies in the two studies is comparable, 	Thank you for your comment. The committee considered the different assumptions relating to costs of subsequent treatments (see sections 3.11, 3.12, and 3.13 of the FAD).

Comment number	Type of stakeholder	Organisation name	Stakeholder comment Please insert each new comment in a new row	NICE Response Please respond to each comment
			<p>owing to the range of drugs available during study conduction</p> <ul style="list-style-type: none"> • Both studies included subsequent therapies based on real clinical choices in a context where monoclonal antibodies' availability was limited or not an option (pre-registration) • Both studies provide clear, concordant, real world estimates of subsequent rates of second ASCT, which are low and similar by arm. <p>The scenarios applied in the model closely reflected the data observed in Myeloma XI and in CALGB, and therefore were deemed in line with clinical efficacy reflected in the data from these two studies; and furthermore, to closely reflect real clinical choices in a context before monoclonal antibodies had become available.</p> <p>The scenarios were extensively validated by means of clinical opinion.</p> <p>It was concluded that they closely reflect the scenarios that would still be observed in clinical practice in the absence of CDF drugs.</p> <p>The scenarios showed that:</p> <ul style="list-style-type: none"> • The cost-effectiveness of lenalidomide in maintenance is robust to uncertainty regarding the most plausible rates of subsequent therapies after first progression; although the therapeutic pathway in MM is rapidly evolving, the most commonly used therapies at the time of CALGB and Myeloma XI have remained actual and therefore the distributions reflected in the model are valid and closely concordant with the efficacy data; • All scenarios used are highly concordant with the exclusion of therapies currently available in the UK via the Cancer Drug fund; • The cost-effectiveness is robust to variations in current, plausible values for second ASCT; • The cost-effectiveness improves with higher rates of lenalidomide and carfilzomib used in second line; the values used in the model are highly concordant with clinical data and with the expected use of these therapies in clinical reality 	
8	Consultee (company)	Celgene	<p><i>Myeloma XI trial data should be used to estimate relative dose intensity</i></p> <p>An extensive description of methods used to estimate RDI was added to the Addendum to this response. It was found that the RDI from Myeloma XI is approximately ■. This value was applied in the model.</p>	Thank you for your comment. The committee considered the company's approach to estimating relative dose intensity based on Myeloma XI data (see section 3.14 of the FAD).

Comment number	Type of stakeholder	Organisation name	Stakeholder comment Please insert each new comment in a new row	NICE Response Please respond to each comment
			<p>Data from Myeloma XI show that in real clinical practice, the dosing of maintenance with lenalidomide is adapted to patients by means of the extensive use of treatment delays and dose reduction. In practice, this approach supports a cost-effective use of lenalidomide.</p> <p>A costing methodology was used in the model that accounted for the cost of all prescribed drugs, from the NHS perspective. From this viewpoint, the costing methods accounted for all drug that is wasted as part of delivery of treatment.</p> <p>Because there was no information regarding whether patients are compliant in everyday use is impossible to ascertain; nevertheless, this would not be a factor in the costing of treatment from the perspective of the UK NHS.</p> <p>The reduced RDI was the result of spaced intervals between a cycle and another and all daily doses dispensed to patients are fix dose tablets.</p> <p>Therefore, it was concluded that all possible sources of wastage have been accounted for in the model.</p>	
9	Consultee (company)	Celgene	<p><i>A model scenario reflecting a 1-day to 28-day lenalidomide treatment regimen.</i></p> <p>A model scenario using 28/28 days of maintenance administration was applied in the model in a scenario analysis.</p> <p>The scenario showed that the large majority of results obtained for the 21/28 days dose remain valid for the 28/28 days dose.</p>	Thank you for your comment. The committee considered the 21- and 28-day dosing scenarios (see sections 3.2 and 3.15 of the FAD).
10	Consultee (Patient organisation)	Myeloma UK	<p>Myeloma UK Response to Lenalidomide Maintenance ACD Consultation</p> <p>Myeloma UK is very disappointed that lenalidomide maintenance for newly diagnosed multiple myeloma patients who are eligible for high-dose therapy and stem cell transplantation (HDT-SCT) has not been approved for routine commissioning.</p> <p>Given the scale of new modelling and evidence requested from the company, and the need for this to be informed by expert clinical opinion and relevant patient insight, we ask that clinicians and Myeloma UK are invited to take part in the second Committee meeting.</p>	Thank you for your comment. After considering additional evidence provided by the company, the committee considered lenalidomide maintenance therapy to be cost-effective (see section 3.15 of the FAD), so the treatment has been recommended.
11	Consultee (Patient organisation)	Myeloma UK	<p><u>Has all of the relevant evidence been taken into account?</u></p> <p>No. We note the requests in the Appraisal Consultation Document (ACD) for new modelling and evidence to be presented by the company to address areas of uncertainty identified by the Committee.</p>	Thank you for your comment. Comment noted.

Comment number	Type of stakeholder	Organisation name	Stakeholder comment Please insert each new comment in a new row	NICE Response Please respond to each comment
			<p>We welcome the following findings in the ACD based on the evidence presented:</p> <ul style="list-style-type: none"> - lenalidomide is the only potential option for maintenance treatment for multiple myeloma after an autologous stem cell transplant. This underlines the high level of unmet need at this point in the pathway - Lenalidomide maintenance would be the standard treatment for newly diagnosed myeloma patients who are eligible for an HD-SCT - The demonstrable need for and benefits of this treatment are strengthened by the fact that first remission is often the best remission, as patients are at their fittest and can maintain their highest quality of life and, unfortunately, the numbers of patients able to access further lines of treatment diminishes - The dosing schedule used in clinical practice would be 21 out of 28 days. We understand why, for completeness, the Committee has asked for further data on the 28 days schedule since it reflects the marketing authorisation. However, we note that the unanimous support of clinicians for the 21-day schedule was strengthened by NHS England's confirmation that it would commission a 21-day schedule. There can be no practical doubt that the 21-day schedule is the one that would be used in clinical practice - Lenalidomide is an effective maintenance treatment for people who have had an autologous stem cell transplant. We emphasise that lenalidomide is not just an effective treatment, it is highly effective; the magnitude of additional clinical benefit it delivers is rarely seen in the treatment of myeloma, which remains an incurable cancer - Myeloma XI should be used in relation to relative dose intensity 	
12	Consultee (Patient organisation)	Myeloma UK	<p><u>Are the summaries of clinical and cost effectiveness reasonable interpretations of the evidence?</u></p> <p>Given the scale of further information requested from the company at this point in the appraisal process, it is not possible to come to a definitive view on whether the summaries are reasonable.</p> <p>We understand why the Committee has requested further evidence from the company to enable further scrutiny and understanding of uncertainty, for example to enable assumptions about the effects of subsequent therapies to be fully explored.</p> <p>However, this work must be seen in the context of the continuously and rapidly evolving myeloma treatment pathway. (Impacted by routine commissioning, CDF approvals and also by recent approvals of alternative oral treatments during COVID- 19.)</p>	<p>Thank you for your comment. The committee considered areas of outstanding uncertainty when making its decision. The committee was aware that it was difficult to make assumptions about the treatments used later in the pathway (see sections 3.11, 3.12 and 3.13 of the FAD).</p> <p>After considering additional evidence provided by the company, the committee considered lenalidomide maintenance therapy to be cost-effective (see section 3.15 of the FAD), so the treatment has been recommended.</p>

Comment number	Type of stakeholder	Organisation name	Stakeholder comment Please insert each new comment in a new row	NICE Response Please respond to each comment
			<p>In this context, it is inevitable that trial design and subsequent data will be unable to fully reflect UK clinical practice at any given time. The Committee has also acknowledged the challenges in modelling the costs of subsequent therapies given that Cancer Drugs Fund (CDF) approved therapies should not be used in economic modelling.</p> <p>A proportionate approach to the inevitable uncertainty which arises from the welcome development of the myeloma treatment pathway is therefore vital. Not to do this would be unreasonable.</p> <p>We note that the issue of treatment sequencing and the inability to reflect real world use of CDF funded drugs is impacting all myeloma appraisals and that current procedures may not be sustainable. There is a danger that, in the absence of new approaches to dealing with this complexity, NICE decisions will become increasingly disconnected from established real world practice and therefore difficult to present as meaningfully “reasonable”.</p>	
13	Consultee (Patient organisation)	Myeloma UK	<p><u>Are the provisional recommendations sound and a suitable basis for guidance to the NHS?</u></p> <p>No. Notwithstanding the further information that has been requested from the company, we do not accept that a decision not to recommend lenalidomide maintenance post HDT-SCT is sound.</p> <p>There is a clear and significant unmet need for lenalidomide maintenance post HDT-SCT for the treatment of myeloma which is standard best practice internationally.</p> <p>In Myeloma XI we have a mature and comprehensive UK data set which, alongside the CALGB trial, provides compelling evidence of the significant benefit delivered by lenalidomide maintenance post HDT-SCT.</p> <p>This is an incredibly effective, life-extending, safe treatment, administered orally which, in the current COVID-19 environment, delivers further benefits to patients, families and to the NHS.</p> <p>It would be deeply concerning if, with this rich data, the company, Evidence Review Group (ERG) and NICE were not collectively able to resolve sufficiently these areas of uncertainty to enable a positive recommendation.</p> <p>In the next stage of the appraisal it is therefore essential that the company provide the additional evidence the Committee has requested; and that, in interpreting the evidence, the Committee takes into account the inevitable limitations of both the data, and of appraisal methods and processes in providing</p>	Thank you for your comment. After considering additional evidence provided by the company, the committee considered lenalidomide maintenance therapy to be cost-effective (see section 3.15 of the FAD), so the treatment has been recommended.

Comment number	Type of stakeholder	Organisation name	Stakeholder comment Please insert each new comment in a new row	NICE Response Please respond to each comment
			<p>a complete and certain picture of clinical benefit and practice.</p> <p>Not to do so runs the risk of arriving at a decision which can be justified methodologically, but which is not reasonable when it is applied in a real-world context.</p>	
14	Consultee (Patient organisation)	Myeloma UK	Myeloma UK have sought comments directly from patients about the impact of the draft no recommendation and these have been included Appendix A of our response to the ACD.	Thank you for your comment. Comment noted.
15	Web comment (public)	(Web commenter 1)	<p>I am a patient who was diagnosed with multiple myeloma in October 2019. I underwent VDT and was ready for an SCT in April 2020 but this was delayed due to Covid 19. I was prescribed thalidomide as maintenance treatment but my Paraprotein levels started to rise so I received a further 1.5 cycles of VDT which brought the levels down again and I underwent an SCT in August this year. I note the Committee accepts that Lenalidomide as a maintenance treatment after SCT improves PFS and OS. The uncertainty and reason for the draft no seems to be about the cost benefits of treatment and other issues surrounding the models used, and the use of existing data by the pharmaceutical company in support of their application for approval.</p> <p>As a patient all I can do is outline my views.</p> <p>I am of working age. When I have recovered from my SCT I intend to return to work. As myeloma is incurable, I recognise, as all patients do, that myeloma will return some day. What we don't know is when it will return so it is like having the sword of Damocles hanging above you 24/7. I cannot underestimate the effect of this on patients.</p> <p>I have only received VDT to date, which was tough going and prevented me from working. I worry that when I relapse, (assuming my SCT has worked, I won't know until after a bone marrow biopsy in November) I will be unable to work again. Many patients in a similar position will have to apply for state benefits once they relapse, and there is also the cost to the NHS of treatment for active disease which the data you have considered indicates will be incurred sooner without maintenance treatment. I do hope the relevant data will be provided to you about the possible costs.</p> <p>The evidence you have already reviewed strongly supports the proposition that the first remission after SCT is the longest and provides the best quality of life for most patients.</p> <p>Lenalidomide will provide most patients, if it's available as maintenance treatment after SCT on the NHS, with:</p> <p>A) Longer quality of life, which is so important to us.</p> <p>B) More time in work for those of working age. For many of us continuing to work and be able to financially support ourselves and our families for as long as possible is very important. It also benefits the economy, and reduces reliance on</p>	<p>Thank you for your comment. The committee considered lenalidomide would be an effective treatment and would be widely used amongst people who have had an autologous stem cell transplant (see sections 3.1 and 3.3 of the FAD).</p> <p>After considering additional evidence provided by the company, the committee considered lenalidomide maintenance therapy to be cost-effective (see section 3.15 of the FAD), so the treatment has been recommended.</p> <p>The views of clinical experts and patient/carer representatives were considered by the Appraisal Committee when formulating its recommendations.</p>

Comment number	Type of stakeholder	Organisation name	Stakeholder comment Please insert each new comment in a new row	NICE Response Please respond to each comment
			<p>state benefits. C) Longer time off intensive treatment before the disease becomes active once more. Intensive treatment believe me is usually tough for most of us.</p> <p>The data you have reviewed indicates the vast majority of patients would take Lenalidomide as a maintenance treatment if it was available on the NHS. This treatment is available privately in the UK. It is available in Europe. It should be available on the NHS to all myeloma patients who have had an SCT, there should be no arbitrary cut off date, and whose PFS and OS it is considered clinically, taking into account the risks to the patient from receiving Lenalidomide, will be extended by receiving maintenance treatment.</p>	
16	Web comment (public)	(Web commenter 2)	<p>Has all of the relevant evidence been taken into account? Yes</p>	Thank you for your comment. Comment noted.
17	Web comment (public)	(Web commenter 2)	<p>Are the summaries of clinical and cost effectiveness reasonable interpretations of the evidence? "No - see detailed comments regarding the intrinsic flaw in the argument that the data does not include current England-specific next treatment comparisons. There is a ""catch 22"" of wanting long term survival data yet wanting that data to reflect current treatment options. The committee needs to decide which argument it is going to use to decline a drug but cannot use opposing reasons in different appraisals. The use of CDF approved therapies should be considered as this reflects current practice in England. The exclusion thereby makes any conclusions not relevant to current practice. A simple review of market share will reveal this to be the case."</p>	<p>Thank you for your comment. The committee was aware of the difficulties associated with making assumptions about subsequent therapies (see section 3.11 of the FAD).</p> <p>After considering additional evidence provided by the company, the committee considered lenalidomide maintenance therapy to be cost-effective (see section 3.15 of the FAD), so the treatment has been recommended.</p>
18	Web comment (public)	(Web commenter 2)	<p>Are the recommendations sound and a suitable basis for guidance to the NHS? No. The recommendations fail to recognise the most significant step forward in the management of myeloma in this country in the last decade. The UK is one of the only developed countries in the world to not enable access to a clearly highly effective maintenance option. There are no robust arguments against approval. I implore the committee to rethink its decision in the interests of patient care and clinical outcomes rather than use fatuous arguments to block access to a highly effective drug.</p>	Thank you for your comment. After considering additional evidence provided by the company, the committee considered lenalidomide maintenance therapy to be cost-effective (see section 3.15 of the FAD), so the treatment has been recommended.
19	Web comment (public)	(Web commenter 2)	Lenalidomide 10 mg days 1-21 is the dose that is used across the UK with many patients still currently receiving it in the Myeloma XI clinical trial and used in the private sector in the UK for those who have access to this. There is substantial familiarity with its use across the country and of the management of possible toxicity.	Thank you for your comment. The committee considered the dosing schedule that would most likely be used in clinical practice (see section 3.1 of the FAD).
20	Web comment (public)	(Web commenter 2)	Meta-analysis of all the trials mentioned using different dosing schedules of lenalidomide show a clear advantage of lenalidomide over no maintenance.	Thank you for your comment. Comment noted.
21	Web comment (public)	(Web commenter 2)	Data from the Myeloma XI study (Jackson et al, Lancet Oncology 2019) indicates a manageable toxicity profile with lenalidomide maintenance. There is significant familiarity in the UK managing these now and therefore it is not the case that	Thank you for your comment. The committee considered that lenalidomide is likely to have an acceptable safety profile (see section 3.5 of the FAD).

Comment number	Type of stakeholder	Organisation name	Stakeholder comment Please insert each new comment in a new row	NICE Response Please respond to each comment
			there is clinical uncertainty about the safety profile of the drug amongst those who treat people with myeloma.	
22	Web comment (public)	(Web commenter 2)	The ERG argument is intrinsically flawed. The Myeloma XI trial is the largest of its kind and is directly relevant to UK practice. In a rapidly evolving field such as myeloma it cannot be expected that next line treatment options will remain static over time as newer treatments are developed. It is integral to the presentation of long term survival data that patients have to have been treated several years ago in order to generate robust survival modelling. If this argument were followed to its logical conclusion, no drug could ever be evaluated unless the survival associated with that drug were measured in a matter of months because treatment options would have changed over that time. Patients should not be penalised because the modelling preferred by the ERG cannot take account of this.	Thank you for your comment. The committee considered the company had appropriately used the committee's preferred approach to extrapolate survival in its updated analyses (see section 3.8 of the FAD).
23	Web comment (public)	(Web commenter 2)	There is no clinical rationale to suggest that 28 day continuous dosing is any more or less effective than 21/28 dosing. That is mere speculation and should not be used in an appraisal where robust evidence should be examined.	Thank you for your comment. Comment noted.
24	Web comment (public)	(Web commenter 2)	Given the range of therapies available on the CDF in myeloma, it is clinically inappropriate not to include these options as they do reflect current standard of care in myeloma with proven clinical efficacy and good UK clinical experience. This exclusion needs urgent revision as makes assumptions about next treatment irrelevant to the population in question treated in England and any modelling entirely fictitious. Carfilzomib is available as a second line treatment option in myeloma - see TA457.	Thank you for your comment. The committee was aware of the difficulties associated with making assumptions about subsequent therapies (see section 3.11 of the FAD).
25	Web comment (public)	(Web commenter 2)	Agree regarding the use of Myeloma XI for relative dose intensity which is current UK specific data using a single agent as maintenance rather than multiagent approach in a highly selected population of patients with relapsed disease who are likely to have higher compliance with therapy.	Thank you for your comment. The committee considered the company's approach to estimating relative dose intensity based on Myeloma XI data (see section 3.14 of the FAD).
26	Web comment (public)	(Web commenter 3)	Has all of the relevant evidence been taken into account? No comment	N/A
27	Web comment (public)	(Web commenter 3)	Are the summaries of clinical and cost effectiveness reasonable interpretations of the evidence? As someone whose father has recently been diagnosed with Active Multiple Myeloma and will need Lenalidomide as maintenance treatment in order to live for longer, I don't believe that these summaries are "reasonable". If "reasonable" is defined as: "having sound judgement; fair and sensible" I would absolutely argue that these summaries are unreasonable. They couldn't be understood as fair or sensible by anyone who is really thinking about the consequences. The recommendations state that the science shows that Lenalidomide maintenance treatment prolongs life and at a high quality. There isn't another drug or any other treatment available that does this, so the only alternative is to die faster. Thousands of people would live quality lives for years longer than they are currently able to with this drug available on the NHS. If there wasn't a pandemic, people would be able to access this drug as part of a "trial" - my father would be	Thank you for your comment. The committee considered lenalidomide would be an effective treatment and would be widely used amongst people who have had an autologous stem cell transplant (see sections 3.1 and 3.3 of the FAD). After considering additional evidence provided by the company, the committee considered lenalidomide maintenance therapy to be cost-effective (see section 3.15 of the FAD), so the treatment has been recommended. The views of clinical experts and patient/carer representatives were considered by the Appraisal

Comment number	Type of stakeholder	Organisation name	Stakeholder comment Please insert each new comment in a new row	NICE Response Please respond to each comment
			<p>one of them. Therefore, to not approve this drug at a time when there is a pandemic on seems inhumane to me. Particularly given that it is taken as a capsule at home, which means it doesn't require a hospital visit. My understanding is that this drug is prescribed as maintenance treatment as a matter of practise if a patient is being treated privately, rather than through the NHS. I don't think it is "reasonable" that someone who has a lot of money could afford to keep themselves alive for longer than someone who doesn't have a lot of money. If the cost-effectiveness needs to be understood more clearly from additional modelling, I think this should be done asap to turn this decision into a "yes". If Celgene / BMS are charging more than the NHS modelling can approve, they should offer this treatment to the NHS at a lower fee, or NICE and the NHS should find a way to consider this through a model that allows for a higher threshold. There is something wrong here when a treatment that works so well is being withheld from people who undeniably need it and can't get it any other way. I understand that Celgene / BMS have stakeholders to answer to and that this drug is a "blockbuster" for them, but they will still make profits, while allowing thousands of families to know their loved ones for longer. I honestly believe that to say "no" to approval of this drug would go against the Hippocratic Oath which surely used to be at the heart of the pharmaceutical industry as well as healthcare more generally. If it is coming out of patent soon anyway, why allow thousands of people to suffer and die prematurely, just to make more money in the meantime? How can anyone in a decision-making position allow that to happen? If NICE's mission statement is to: "Protect people's money, ensure their safety and improve their experiences", they must be able to find a way to approve Lenalidomide for maintenance on the NHS. Otherwise will be forced into paying life-changing prices, selling their homes, to try and live for longer. This isn't "safe" for mental or physical health and, rather than "improving experiences", would most definitely negatively affect their experiences and those of their wider circle of friends and family too. The science says that this drug really helps people who are in desperate need, the only potential problem is the price. Please, please reconsider this draft no and find a way through the corporate red tape and stakeholder interests to remember that real people's lives are at stake here. My dad is one of them and I beg you to reconsider.</p>	<p>Committee when formulating its recommendations.</p>
28	Web comment (public)	(Web commenter 3)	<p>Are the recommendations sound and a suitable basis for guidance to the NHS? Absolutely not. The recommendations acknowledge that people with this disease who aren't very rich or who didn't get onto a trial before Covid-19 struck, will die faster than is necessary. They show that this drug would increase how long people live and that it extends the time before the condition gets worse. Also that prolonging the first remission is a key factor in optimising patient survival. The recommendations show that there is no other treatment available, so the only alternative to the patient is to die faster, knowing that their life wasn't worth the cost to the system. I understand that other people need NHS funds too and that there is only so much to go around. But how much more effective does a</p>	<p>Thank you for your comment. The committee considered lenalidomide would be an effective treatment and would be widely used amongst people who have had an autologous stem cell transplant (see sections 3.1 and 3.3 of the FAD).</p> <p>After considering additional evidence provided by the company, the committee considered lenalidomide maintenance therapy to be cost-effective (see section 3.15 of the FAD), so the treatment has been</p>

Comment number	Type of stakeholder	Organisation name	Stakeholder comment Please insert each new comment in a new row	NICE Response Please respond to each comment
			<p>treatment need to be, for it to be worth it? I understand that there are “limitations to the cost effectiveness-model”, but surely at a really basic human level, someone has to see that these limitations must be overcome in order to save lives? If the drug was less expensive, the recommendations say that the NHS would recommend this treatment as standard for patients with MM after an autologous stem cell transplant. Which means it works! Somebody just needs to find a way to make the costs make sense in the system that has been set up. Please find it in your heart to do this work and find a way. There must be a way for the economic modelling system to say “yes” to prolonging the lives of thousands of real people like my dad. Thank you so much for anything you can do.</p>	<p>recommended.</p> <p>The views of clinical experts and patient/carer representatives were considered by the Appraisal Committee when formulating its recommendations.</p>
29	Web comment (public)	(Web commenter 3)	<p>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?</p> <p>No comment</p>	N/A

Lenalidomide for the maintenance treatment of multiple myeloma after autologous stem cell transplantation [ID475]

Consultation on the appraisal consultation document – deadline for comments 5pm on Tuesday 29 September email: NICE DOCS

<p>x</p>	<p>Please read the checklist for submitting comments at the end of this form. We cannot accept forms that are not filled in correctly.</p> <p>The Appraisal Committee is interested in receiving comments on the following:</p> <ul style="list-style-type: none"> • 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? <p>NICE is committed to promoting equality of opportunity, eliminating unlawful discrimination and fostering good relations between people with particular protected characteristics and others. Please let us know if you think that the preliminary recommendations may need changing in order to meet these aims. In particular, please tell us if the preliminary recommendations:</p> <ul style="list-style-type: none"> • could have a different impact on people protected by the equality legislation than on the wider population, for example by making it more difficult in practice for a specific group to access the technology; • could have any adverse impact on people with a particular disability or disabilities. <p>Please provide any relevant information or data you have regarding such impacts and how they could be avoided or reduced.</p>
<p>Organisation name – Stakeholder or respondent (if you are responding as an individual rather than a registered stakeholder please leave blank):</p>	<p>Celgene Ltd.</p>
<p>Disclosure Please disclose any past or current, direct or indirect links to, or funding from, the tobacco industry.</p>	<p>[Insert disclosure here]</p>
<p>Name of commentator person completing form:</p>	<p>Emanuela Castelnovo</p>

Lenalidomide for the maintenance treatment of multiple myeloma after autologous stem cell transplantation [ID475]

Consultation on the appraisal consultation document – deadline for comments 5pm on Tuesday 29 September email: NICE DOCS

Comment number	Comments
1	<p style="text-align: center;">Insert each comment in a new row. Do not paste other tables into this table, because your comments could get lost – type directly into this table.</p> <p><i>The company should have presented evidence from other trials of lenalidomide maintenance treatment</i></p> <p>A complete and thorough presentation of clinical effectiveness evidence from the CALGB-100104 and GIMEMA trials of lenalidomide maintenance treatment, in addition to further details from Myeloma XI, are included in an addendum to this response.</p> <p>The review of the data showed that</p> <ol style="list-style-type: none"> 1. CALGB and Myeloma were well conducted studies and included valid comparisons of lenalidomide maintenance compared with placebo / observation. Both studies showed a statistically significant benefit on overall survival and progression-free survival with lenalidomide maintenance. 2. CALGB and Myeloma XI baseline characteristics and populations were presented in detail. Whilst the two studies differed in the baseline distribution of ISS scores, and somewhat by gender and age, the differences were potentially due to methods used in data collection and definitions, as well as some differences in prior therapies (induction) received in the studies. The review identified factors important for consideration that informed the subsequent matched-adjusted analyses of CALGB and Myeloma XI. 3. The methods and study design used for GIMEMA are affected by an error that caused the study to provide a biased estimate of the treatment efficacy with lenalidomide maintenance compared with placebo and specifically, with respect to the Appraisal decision problem. <p>Overall, the CALGB and Myeloma XI were found to be robust studies, largely comparable and pertinent with the decision problem.</p>
2.	<p><i>The company’s model structure does not allow assumptions about subsequent treatments to be explored. The committee concluded that the company’s model structure had limitations. It also concluded that there was likely to be uncertainty around the cost-effectiveness estimate because assumptions about the effects of subsequent therapies on survival could not be fully explored.</i></p> <p>The model structure was pertinent with regards to the exploration of the impact of subsequent therapies.</p>

Lenalidomide for the maintenance treatment of multiple myeloma after autologous stem cell transplantation [ID475]

Consultation on the appraisal consultation document – deadline for comments 5pm on Tuesday 29 September email: NICE DOCS

	<p>In the Addendum to this Response, it was shown that the subsequent therapies used in CALGB and Myeloma XI were largely similar, and both were reflective of clinical settings where monoclonal antibodies were not in use as they were not licensed yet. This is reflective of subsequent therapies in a world without Cancer Drugs Fund access.</p> <p>Extensive clinical validation was conducted on the scenarios included in the model, which confirmed that the distributions used are reflective of current practice without CDF funded therapies.</p> <p>Because of concordance between clinical efficacy incorporated in the model and subsequent therapies costed as part of the longer term follow up, the model structure is adequate to reflect valid subsequent therapies scenarios.</p> <p>The proportions of subsequent therapies in the model were tested in scenario analyses, using clinically validated proportions for second ASCT and lenalidomide subsequent use in further lines. The cost-effectiveness of lenalidomide maintenance was confirmed across a large range of likely scenarios.</p>
3	<p><i>The company's methods and rationale for pooling Myeloma XI and CALGB 100104 data, and adjusting for treatment switching, are unclear</i></p> <p><u>Pooling and adjustment methods</u></p> <p>New statistical models were developed to account for potential differences in populations and study design between CALGB and Myeloma XI. Propensity scores weighting (PW) and matched-adjusted indirect comparison (MAIC) adjusted comparisons were developed. The new analyses provide a more extensive comparison of the two trials, which was used to inform an extensive set of cost-effectiveness scenarios in the model.</p> <p>All models used for matching produced highly concordant results and showed that the case for the two studies being in comparable populations remains strong.</p> <p>In the analyses, differences were identified in the distribution of some potential prognostic factors. Nevertheless, the matched-adjusted OS Kaplan-Meier for placebo remained stable in all analyses and showed a high degree of concordance with both the unadjusted CALGB Kaplan-Meier and the Kaplan-Meier for the observation arm in Myeloma XI.</p> <p>It is therefore unlikely that the prognosis for people in both placebo/observation arms of the two studies may be impacted by factors other than those accounted for in the matching models.</p> <p>This would confirm the initial assumption that the populations in the two studies</p>

Lenalidomide for the maintenance treatment of multiple myeloma after autologous stem cell transplantation [ID475]

Consultation on the appraisal consultation document – deadline for comments 5pm on Tuesday 29 September email: NICE DOCS

	<p>were not largely different with respect to baseline characteristics, and that the untreated population in the two studies has a similar prognosis.</p> <p>The robustness of the KM curve with respect to matching adjustment also constitutes indirect support to the difference between the lenalidomide arm in CALGB and that in Myeloma XI could be largely attributed in lenalidomide dosing. For this reason, we also extended the pooled model as part of the new analyses, using a treatment by trial interaction term as well as matching for population and trial design characteristics. The pooled model confirms the comparability of the studies and has the advantage of controlling for the difference in treatment dosage between CALGB and Myeloma XI (21/28 days and 28/28 days).</p> <p>The matched-adjusted analyses were robust to the methods employed, to the choice of matching variables and provide extensive validation for the comparability of Myeloma XI and CALGB; based on the large majority of the analyses and scenarios from the matched adjusted comparisons, the cost-effectiveness of lenalidomide maintenance remained robust and within acceptability ranges.</p> <p>Given all variations and methodological approaches considered, the cost-effectiveness of lenalidomide maintenance remained favourable.</p>
4	<p><i>The justification for using the rank preserving structural failure model should be provided</i></p> <p>An extensive presentation of methods, feasibility and results for the RPSFM approach used in the submission is presented.</p> <p>In brief, at a time when the CALGB study hit the superiority boundary in the early follow up (2 years), all patients initially randomised to placebo that had not progressed after ASCT (thus still fulfilling eligibility for maintenance) were offered to switch to active maintenance therapy.</p> <p>Of more than 100 patients in the placebo arm at the time who had not progressed, the majority chose to receive maintenance. The number of patients who did not switch was small and such that an analysis based on data from non-switchers would be possible.</p> <p>Furthermore, switching in CALGB was unrelated with disease outcomes and because it was conditional on patients not having progressed, the conditions for non-informative censoring would probably be violated.</p> <p>Extensive assessment of the common treatment effect assumption, on which RPSFM relies, was also conducted and presented.</p> <p>The RPSFM method was considered the most suitable approach for the reasons</p>

Lenalidomide for the maintenance treatment of multiple myeloma after autologous stem cell transplantation [ID475]

Consultation on the appraisal consultation document – deadline for comments 5pm on Tuesday 29 September email: NICE DOCS

	above.
5	<p><i>Survival extrapolations should use Myeloma XI data as the main source of evidence but could be supplemented with CALGB 100104 data</i></p> <p>The model was extended to include a range of matched-adjusted analyses. Two statistical approaches were taken, propensity scores weights and MAIC; the former using patient level data for both studies and the latter matching CALGB to aggregate data from Myeloma XI. The methods are explained in detail in the Addendum to this response.</p> <p>The matched-adjusted analyses were incorporated into the cost-effectiveness model; the following scenarios were generated:</p> <ul style="list-style-type: none"> • Use of treatment effects from Myeloma XI until month 60, and thereafter using the treatment effects for the matched-adjusted CALGB extrapolations; • Use of treatment effects from Myeloma XI at all time; • Use of treatment effects from the revised pooled analysis of CALGB and Myeloma XI at all time points, incorporating covariates to control for study, treatment, and study-by-treatment interaction. <p>The aim of the scenarios was to assess the cost-effectiveness robustness to CALGB analyses.</p> <p>Statistical goodness of fit and clinical plausibility were used to identify the potentially relevant extrapolations in the model; however, it was preferred to present the results of cost-effectiveness for all clinically plausible distributions, with the exception of distributions that clearly were not appropriate.</p> <p>The revised estimates of cost-effectiveness show that lenalidomide maintenance is robustly cost-effective across all scenarios, for all plausible distributions, and for all matching and adjustment methods considered.</p> <p>Using estimates from the pooled model and using the Myeloma XI specific pooled effect also confirms the cost-effectiveness results, regardless of extrapolation distributions chosen.</p>
6	<p><i>The treatment effect of lenalidomide maintenance may wane over time and this should be included in the model</i></p> <p>The case for waning effects being likely assumptions for lenalidomide used in</p>

Lenalidomide for the maintenance treatment of multiple myeloma after autologous stem cell transplantation [ID475]

Consultation on the appraisal consultation document – deadline for comments 5pm on Tuesday 29 September email: NICE DOCS

	<p>maintenance was explored. It was concluded that most common reasons to suspect waning of treatment effect are not likely to occur with maintenance.</p> <ol style="list-style-type: none"> 1. Maintenance with lenalidomide continues until disease progression or unacceptable toxicity. Capping rule for the duration of maintenance are not applicable 2. Non-compliance was unlikely with Myeloma XI. Maintenance with lenalidomide is spaced with ‘off treatment’ intervals such that tolerability if actively pursued. 3. In the case of prolonged non-compliance, it is unlikely that a patient would remain in a prolonged state of pre-progression, captured in PFS. 4. Myeloma XI and CALGB have long follow-up; PFS data in CALGB are mature, making extrapolation necessary from a late point in time in the model, reducing uncertainty. 5. The analysis of CALGB and Myeloma XI showed is robust evidence of the pertinence of proportional hazard which is not at odds with the assumption that treatment effect is decreasing. <p>For these reasons, and based on clinical opinion received, it is believed that waning of treatment effect would not be plausible for maintenance with lenalidomide.</p> <p>Nevertheless, scenario analyses were used to explore the impact of waning treatment effect, which showed that cost-effectiveness estimates are robust to most conservative treatment effect waning assumptions.</p>
7	<p><i>Costs of subsequent treatments are highly uncertain so scenarios should be presented</i></p> <p>Subsequent therapies in Myeloma XI and CALGB were assessed and tabulated for comparison, to verify the material similarity between subsequent treatment used in CALGB and Myeloma XI.</p> <ul style="list-style-type: none"> • Both studies show that most participants who progressed received subsequent therapies • The mix of subsequent therapies in the two studies is comparable, owing to the range of drugs available during study conduction • Both studies included subsequent therapies based on real clinical choices in a context where monoclonal antibodies’ availability was limited or not an option (pre-registration) • Both studies provide clear, concordant, real world estimates of subsequent rates of second ASCT, which are low and similar by arm. <p>The scenarios applied in the model closely reflected the data observed in Myeloma XI and in CALGB, and therefore were deemed in line with clinical efficacy reflected in the data from these two studies; and furthermore, to closely reflect real clinical choices in a context before monoclonal antibodies had become available.</p>

Lenalidomide for the maintenance treatment of multiple myeloma after autologous stem cell transplantation [ID475]

Consultation on the appraisal consultation document – deadline for comments 5pm on Tuesday 29 September email: NICE DOCS

	<p>The scenarios were extensively validated by means of clinical opinion.</p> <p>It was concluded that they closely reflect the scenarios that would still be observed in clinical practice in the absence of CDF drugs.</p> <p>The scenarios showed that:</p> <ul style="list-style-type: none"> • The cost-effectiveness of lenalidomide in maintenance is robust to uncertainty regarding the most plausible rates of subsequent therapies after first progression; although the therapeutic pathway in MM is rapidly evolving, the most commonly used therapies at the time of CALGB and Myeloma XI have remained actual and therefore the distributions reflected in the model are valid and closely concordant with the efficacy data; • All scenarios used are highly concordant with the exclusion of therapies currently available in the UK via the Cancer Drug fund; • The cost-effectiveness is robust to variations in current, plausible values for second ASCT; • The cost-effectiveness improves with higher rates of lenalidomide and carfilzomib used in second line; the values used in the model are highly concordant with clinical data and with the expected use of these therapies in clinical reality
8	<p><i>Myeloma XI trial data should be used to estimate relative dose intensity</i></p> <p>An extensive description of methods used to estimate RDI was added to the Addendum to this response. It was found that the RDI from Myeloma XI is approximately 86%. This value was applied in the model.</p> <p>Data from Myeloma XI show that in real clinical practice, the dosing of maintenance with lenalidomide is adapted to patients by means of the extensive use of treatment delays and dose reduction. In practice, this approach supports a cost-effective use of lenalidomide.</p> <p>A costing methodology was used in the model that accounted for the cost of all prescribed drugs, from the NHS perspective. From this viewpoint, the costing methods accounted for all drug that is wasted as part of delivery of treatment.</p> <p>Because there was no information regarding whether patients are compliant in everyday use is impossible to ascertain; nevertheless, this would not be a factor in the costing of treatment from the perspective of the UK NHS.</p> <p>The reduced RDI was the result of spaced intervals between a cycle and another</p>

Lenalidomide for the maintenance treatment of multiple myeloma after autologous stem cell transplantation [ID475]

Consultation on the appraisal consultation document – deadline for comments 5pm on Tuesday 29 September email: NICE DOCS

	<p>and all daily doses dispensed to patients are fix dose tablets.</p> <p>Therefore, it was concluded that all possible sources of wastage have been accounted for in the model.</p>
9	<p><i>A model scenario reflecting a 1-day to 28-day lenalidomide treatment regimen.</i></p> <p>A model scenario using 28/28 days of maintenance administration was applied in the model in a scenario analysis.</p> <p>The scenario showed that the large majority of results obtained for the 21/28 days dose remain valid for the 28/28 days dose.</p>

Lenalidomide for the maintenance treatment of multiple myeloma after autologous stem cell transplantation [ID475]

Consultation on the appraisal consultation document – deadline for comments 5pm on Tuesday 29 September email: NICE DOCS

Checklist for submitting comments

- Use this comment form and submit it as a Word document (not a PDF).
- Complete the disclosure about links with, or funding from, the tobacco industry.
- Combine all comments from your organisation into 1 response. We cannot accept more than 1 set of comments from each organisation.
- Do not paste other tables into this table – type directly into the table.
- Please underline all confidential information, and separately highlight information that is submitted under **commercial in confidence** in turquoise and all information submitted under **academic in confidence** in yellow. If confidential information is submitted, please also send a 2nd version of your comment with that information replaced with the following text: 'academic / commercial in confidence information removed'. See the Guide to the processes of technology appraisal (section 3.1.23 to 3.1.29) for more information.
- Do not include medical information about yourself or another person from which you or the person could be identified.
- Do not use abbreviations
- Do not include attachments such as research articles, letters or leaflets. For copyright reasons, we will have to return comments forms that have attachments without reading them. You can resubmit your comments form without attachments, it must send it by the deadline.
- If you have received agreement from NICE to submit additional evidence with your comments on the appraisal consultation document, please submit these separately.

Note: We reserve the right to summarise and edit comments received during consultations, or not to publish them at all, if we consider the comments are too long, or publication would be unlawful or otherwise inappropriate.

Comments received during our consultations are published in the interests of openness and transparency, and to promote understanding of how recommendations are developed. The comments are published as a record of the comments we received, and are not endorsed by NICE, its officers or advisory committees.

**Lenalidomide for the maintenance treatment of
multiple myeloma after autologous stem cell
transplantation - [ID475]**

**Consultation on the appraisal consultation document
Addendum to the response to the consultation
document**

October 2020

Issue 1. Clinical effectiveness evidence from the CALGB-100104 and GIMEMA trials of lenalidomide maintenance treatment

1 Introduction

Myeloma XI, a UK-based trial assessing the efficacy of lenalidomide as maintenance therapy post-ASCT that follows anticipated UK clinical practice, was used as the evidence base to support technology appraisal ID475.

Data are available from two additional trials identified in the systematic literature review, CALGB 100104 and GIMEMA. We present data from the trials and conduct a critical appraisal to inform their appropriateness to support long term survival with lenalidomide as maintenance therapy in response to questions raised by the NICE committee.

Owing to its study design, survival outcomes from the GIMEMA trial are not considered appropriate to assess the efficacy of lenalidomide in maintenance, the rationale for which is discussed in this document.

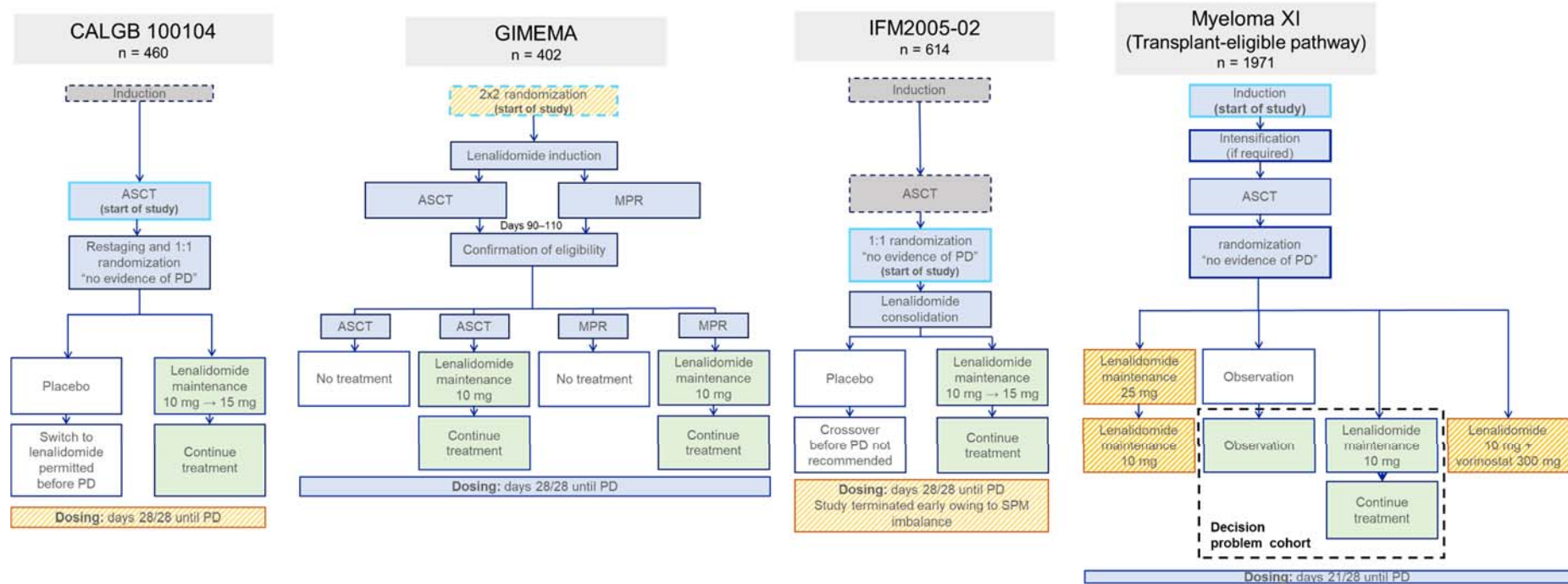
Previously, to support longer-term survival extrapolations, the company pooled survival data from the Myeloma XI and CALGB 100104 trials without adjustment for factors including potential discrepancies in baseline characteristics for participants in the trials, differences in dosage regimen and treatment-switching pre-disease progression in the CALGB 100104 trial.

A meta-analysis estimating survival outcomes in patients with newly diagnosed multiple myeloma who received lenalidomide maintenance post-ASCT was published by McCarthy *et al*, 2017.¹ This analysis includes the CALGB 100104, GIMEMA and IFM2005-02 trials,¹ and was conducted before data from Myeloma XI were available. This meta-analysis is not considered relevant to the decision problem because it includes IFM2005-02,¹ which is not considered relevant to the decision problem.

2 An overview of the study design and evidence base for CALGB 100104, GIMEMA and Myeloma XI

Three trials CALGB 100104, GIMEMA and Myeloma XI were considered by the NICE committee to present survival data relevant to the decision problem covered in ID475. A side-by-side comparison of the CALGB 100104, GIMEMA and Myeloma XI trials is presented in Figure 1 and Table 1, and a more detailed overview of the trials in the Sections that follow.

Figure 1. Graphical comparison of the trial design of CALGB 100104, GIMEMA and Myeloma XI



ASCT, autologous stem cell transplantation; MPR, melphalan, prednisolone and lenalidomide; PD, progressive disease.

Table 1. Study design of CALGB 100104, GIMEMA and Myeloma XI, summary

Study	CALGB 100104 NCT00114101 McCarthy <i>et al</i> , 2012 USA	GIMEMA NCT00551928 Palumbo <i>et al</i> , 2014 Italy and Israel	Myeloma XI NCT01554852 (maintenance phase only) Jackson <i>et al</i> , 2019 UK
Study design	Randomised, parallel, double- blind, placebo- controlled trial	2x2 factorial randomised multicentre, controlled, phase 3 trial. Allocation at study entry: <i>Randomisation 1:</i> ASCT or MPR <i>Randomisation 2:</i> lenalidomide maintenance or no maintenance	Randomised, multifactorial, adaptive, multi-centre, open-label, phase 3 trial.
Population	18–70 years of age with NDMM Stable disease or responsive to 2–12 months of any induction therapy Stable disease or responsive to ASCT in first 100 days	≤ 65 years of age with symptomatic, measurable NDMM No prior therapy	Patients ≥ 18 years of age with symptomatic or non-secretory NDMM <i>ASCT-eligible patients:</i> as per the clinical assessment of the recruiting physician
Intervention	Lenalidomide (post-induction + ASCT)	Lenalidomide (post- HDM + ASCT)	Lenalidomide or lenalidomide + vorinostat
Comparator	Placebo	No Maintenance	Observation
Study arms	Lenalidomide maintenance Placebo	MPR + lenalidomide maintenance MPR + no maintenance ASCT + lenalidomide maintenance ASCT + no maintenance	Lenalidomide Lenalidomide + vorinostat Observation
Inclusion criteria	Stable disease, or a marginal, partial or complete response in the first 100 days following stem- cell transplantation.	Completion of ASCT No evidence of progressive disease	Achievement of at least a minimal response on completion of their assigned induction therapy and had received at least 100 mg/m ² melphalan

Study	CALGB 100104 NCT00114101	GIMEMA NCT00551928	Myeloma XI NCT01554852 (maintenance phase only)
Exclusion criteria	Progressive disease	Progressive disease	Progressive disease or no change following lenalidomide induction therapy (component of KCRD) Failed response to treatment Progressive disease or relapse from complete response
Randomisation	Between 100 and 110 days post-ASCT, 1:1 randomisation to Lenalidomide maintenance placebo.	1:1:1:1 at enrolment: HDM+ASCT plus lenalidomide maintenance HDM+ASCT plus no maintenance MPR plus lenalidomide maintenance [not relevant to this submission] MPR plus no maintenance [not relevant to this submission] Randomisation blinded until the end of induction: eligibility for maintenance reassessed between day 90–110 post-ASCT.	Lenalidomide maintenance or observation
Maintenance regimen	10 mg daily on days 1–28/28 until disease progression	10 mg daily on days 1–21/28	10 mg daily on days 1–21/28
Stratification factors for randomisation	Prior use or non-use of thalidomide during induction Prior use or non-use of lenalidomide during induction Baseline serum β 2-microglobulin (> 2.5 mg/L vs ≤ 2.5 mg/L)	Age (≤ 60 years vs 61–65 years) International Staging System (ISS) disease stage at diagnosis (stage I–II vs stage III)	β 2-microglobulin (< 3.5 mg/L vs 3.5– < 5.5 mg/L vs ≥ 5.5 mg/L vs unknown) Haemoglobin (< 115 g/L vs ≥ 115 g/L for men; < 95 g/L vs ≥ 95 g/L for women) Corrected serum calcium (< 2.6 mmol/L vs ≥ 2.6 mmol/L), serum creatinine (< 140 μ mol/L vs ≥ 140 μ mol/L)

Study	CALGB 100104 NCT00114101	GIMEMA NCT00551928	Myeloma XI NCT01554852 (maintenance phase only)
			Platelets (< 150 x 10 ⁹ cells/L vs ≥ 150 150 x 10 ⁹ cells/L) Centre

ASCT, autologous stem cell transplant; HDM high-dose melphalan; MPR, melphalan, prednisone and lenalidomide; NDMM, newly diagnosed multiple myeloma.

2.1 CALGB 100104

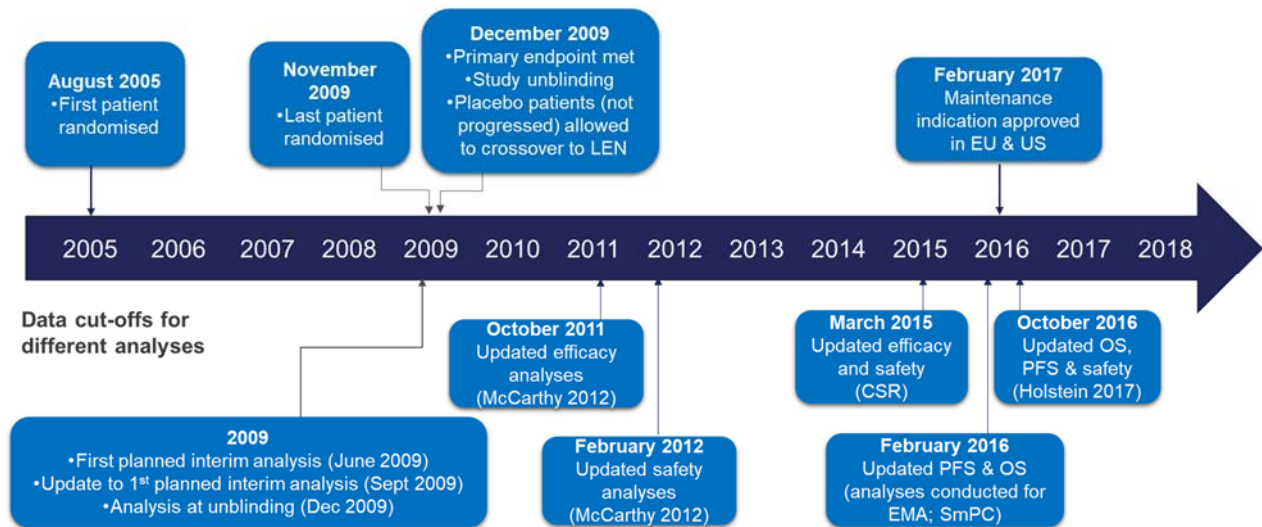
CALGB 100104 (hereafter CALGB) was a US-based, phase 3, randomised, double-blind, placebo-controlled trial that assessed the clinical efficacy of lenalidomide as maintenance therapy compared with placebo in patients who undergone induction followed by a single-cell transplantation. Eligible patients were aged 18–70 years with ECOG performance status 0–1, symptomatic disease requiring treatment, and had received any induction regimen 2–12 months of duration.²

Patients eligible for randomisation to the maintenance phase were those with stable disease, or a marginal, partial or complete response in the first 100 days following stem-cell transplantation. Between 100 and 110 days post-ASCT, eligible patients were randomised 1:1 to lenalidomide 10 mg daily on days 1–28 of a 28-day cycle until disease progression or placebo. Patients allocated to the lenalidomide arm had the option of dose escalation to 15 mg daily after 3 months. Stratification factors were prior use of thalidomide, prior use of lenalidomide and serum β 2-microglobulin (> 2.5 mg/L vs ≤ 2.5 mg/L). Once the primary endpoint was met, the study was unblinded and patients in the placebo arm whose disease had not progressed were permitted to receive lenalidomide.²

2.1.1 Evidence base

Several data cuts from CALGB are available as summarised in Figure 2. The most recent efficacy data are taken from the 19 October 2016 data cut (median follow-up: 91 months), and the most recent baseline characteristics from the 1 February 2016 data cut (median follow-up: [REDACTED]), both of which are used in this analysis.

Figure 2. Overview of key timepoints and available data from the CALGB trial



CSR, clinical study report; EMA, European Medicines Agency; LEN, lenalidomide, OS, overall survival; PFS, progression-free survival; SmPC, Summary of Product Characteristics.

Source: McCarthy 2012,² CSR³ and SmPC.⁴

2.2 GIMEMA

The GIMEMA trial was an open-label, phase 3, randomised trial with a 2x2 factorial design conducted in Italy and Israel. GIMEMA assessed the efficacy of lenalidomide maintenance therapy compared with no maintenance therapy in patients allocated to melphalan, prednisone and lenalidomide (MPR), or ASCT following high dose melphalan at randomisation 1. Eligible patients were aged 65 years or younger with a Karnofsky performance status score of at least 60% and symptomatic, measurable newly-diagnosed multiple myeloma.⁵ Patient baseline characteristics published in the meta-analysis by McCarthy *et al*, 2017¹ suggests that patients may receive up to two ASCTs, although the reporting surrounding this lacks clarity.

All patients were randomly assigned 1:1:1:1 at enrolment to one of four groups as summarised in Table 2. Stratification factors included age (≤ 60 years vs 61–65 years) and International Staging System (ISS) disease stage (stage I–II vs stage III). Patients allocated to maintenance therapy with lenalidomide received lenalidomide 10 mg on days 1–21 of a 28-day cycle. Randomisation to maintenance was conducted at study recruitment (i.e. before receiving induction) and concealed until the end of the MPR or ASCT phase when patients were assessed for maintenance therapy. The

efficacy of maintenance therapy in patients randomised to MPR is not relevant to the decision problem, and therefore not discussed further.⁵

Table 2. Summary of the treatment pathway for the subpopulations of the GIMEMA trial

Stage	At randomisation	At Stage 1	At Stage 2
ASCT no maintenance	Induction	HDM + ASCT	Placebo
ASCT & lenalidomide maintenance	Induction	HDM + ASCT	Lenalidomide maintenance
MPR no maintenance	Induction	MPR	Placebo
MPR & lenalidomide maintenance	Induction	MPR	Lenalidomide maintenance

ASCT, autologous stem cell transplant; HDM, high dose melphalan; MPR, melphalan, prednisolone and lenalidomide. Source: Palumbo *et al*, 2014.⁵

2.2.1 Evidence base

Two publications report outcomes from the GIMEMA trial:

- **Palumbo et al, 2014⁵** reported the outcome of the primary comparison of lenalidomide maintenance vs no maintenance for the ITT population irrespective of whether patients had or had not received an ASCT (data cut: 30 April 2013; median follow-up: 31 months).
- **McCarthy et al, 2017¹** reported the outcome of a meta-analysis that compared the outcomes with and without lenalidomide maintenance in patients who had received an ASCT using data from three trials (CALGB, GIMEMA and IFM 2005-02), which is misaligned with the NICE decision problem owing to the inclusion of the IFM 2005-05 trial (section 8.2).

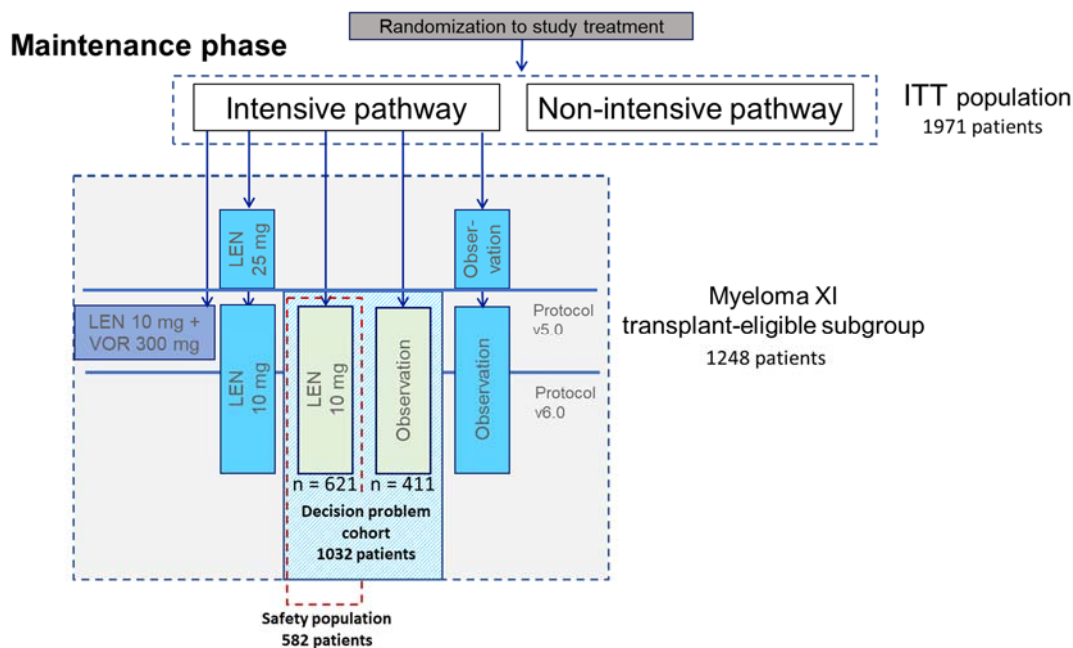
2.3 Myeloma XI

The Myeloma XI trial is an ongoing UK-based phase 3, open-label, randomised trial that assessed the efficacy of lenalidomide maintenance compared with observation in patients who may or may not have received an ASCT. Eligible patients were aged 18 years and had symptomatic or non-secretory multiple myeloma. Two treatment pathways were included in the Myeloma XI trial: the intensive pathway that included younger, fitter patients who were eligible for an ASCT and the non-intensive pathway that included patients who were ASCT-ineligible. Given that the decision problem is focused only on patients who received an ASCT, only outcomes in the subset of

patients that entered the intensive pathway of Myeloma XI (Figure 3) were considered relevant.⁶

Patients in the intensive pathway eligible for randomisation to the maintenance phase of the Myeloma XI were those who achieved at least a minimal response on completion of their assigned induction therapy and had received at least 100 mg/m² melphalan. Randomisation was stratified by β 2-microglobulin (< 3.5 mg/L vs 3.5–< 5.5 mg/L vs \geq 5.5 mg/L vs unknown), haemoglobin (< 115 g/L vs \geq 115 g/L for men; < 95 g/L vs \geq 95 g/L for women), corrected serum calcium (< 2.6 mmol/L vs \geq 2.6 mmol/L), serum creatinine (< 140 μ mol/L vs \geq 140 μ mol/L), platelets (< 150 x 10⁹ cells/L vs \geq 150 x 10⁹ cells/L).⁶

Figure 3. Summary of patients included in the Myeloma XI analysis



LEN, lenalidomide; ITT, intention-to-treat; VOR, vorinostat.

Source: Jackson *et al*, 2019⁶

2.3.1 Evidence base

The primary analysis of the maintenance phase of the Myeloma XI was published in Jackson *et al.*, 2019 (median follow-up [redacted]),⁶ and reported the efficacy of lenalidomide maintenance therapy compared with observation in the ITT population, which included both patients who had received an ASCT and those who had not, and hence was not specific to the decision problem. Therefore, a subset of data from the

Myeloma XI trial on patients who received monotherapy with lenalidomide maintenance in the post-ASCT setting (referred to as the decision problem cohort) was used to support this submission (Figure 3).

3 Comparison of patient baseline characteristics and demographics between CALGB and Myeloma XI

The company considers CALGB an appropriate evidence base to support longer term survival estimates based on study design and patient population. Previously, the company pooled data from CALGB and Myeloma XI without adjustment for differences in lenalidomide maintenance dose and population. In response to questions raised by the NICE committee, the company have compared patient demographics and baseline characteristics (sections 4 and 6) to identify key factors that are discussed further in section 8.3. The equivalent exercise was not performed for the GIMEMA and Myeloma XI trials as the company considered the GIMEMA study biased by design, and therefore unsuitable to inform an unbiased comparison.

4 Comparison of baseline demographics data from Myeloma XI and CALGB

Table 3 Table 3 presents a side-by-side comparison of baseline demographics from the ITT population of CALGB (data cutoff: 1 February 2016; data on file) and the decision problem cohort of Myeloma XI (data cutoff: 23 October 2017; data on file). Comparison of the two studies presents a slight difference in median age and sex, with a higher proportion of patients over 60 years, and a lower proportion of men in CALGB vs Myeloma XI.

Table 3. Comparison of baseline demographics: Myeloma XI and CALGB

	Myeloma XI (decision problem cohort)		CALGB (ITT population)	
Data cut	23 October 2017		1 February 2016	
Characteristic	Lenalidomide (n = 621)	Observation (n = 411)	Lenalidomide (n = 231)	Placebo (n = 229)
Median follow-up, months	██████████		81.9	81.0
Median age, years	██████████	██████████	58.0 ^a	58.0 ^a
< 60 years, n (%)			131 (56.7)	133 (58.1)
Sex, n (%)	██████████	██████████		
Male			121 (52.4)	129 (56.3)

^aAge was only reported at time of randomisation to maintenance therapy.

Pale red highlight denotes imbalance between trials.

ITT, intent to treat.

Source: data on file.

5 Comparison of disease characteristics data from Myeloma XI and CALGB

Table 4 presents a side-by-side comparison of disease characteristics from the ITT population of CALGB (data cutoff: 1 February 2016; data on file) and the decision problem cohort of Myeloma XI (data cutoff: 23 October 2017).

There is a slight imbalance in ISS stage at trial entry with a slightly higher proportion of patients at ISS stage 1 in CALGB than in Myeloma XI. There is also an imbalance in ISS stage between treatment arms in CALGB with a higher proportion of patients at ISS stage 1 in the placebo compared with the lenalidomide arm.

In Myeloma XI, ISS scores were calculated for each patient at study recruitment as part of the trial protocol whereas, ISS stage at diagnosis in CALGB is understood to have been collected retrospectively when patients underwent randomisation to maintenance. It is unclear if ISS scores in CALGB were obtained from patient medical records using values obtained at diagnosis, or from values obtained in the intervening time between diagnosis and randomisation (approximately a year following diagnosis). The retrospective collection or assessment of ISS stage may account for the high proportion of missing data. As ISS stage at any time after diagnosis or during treatment is not considered clinically meaningful by clinical experts (clinical opinion), the comparison between ISS scores in Myeloma XI and CALGB should be interpreted

with caution. The ISS scores at ASCT and randomisation are available for CALGB but as discussed previously and have not been considered here. These values can be found in Appendix A.

Table 4. Comparison of disease characteristics: Myeloma XI (at trial entry) and CALGB (at diagnosis [collected retrospectively])

	Myeloma XI (decision problem cohort)		CALGB (ITT population)	
Data cut	23 October 2017		1 February 2016	
Characteristic	Lenalidomide (n = 621)	Observation (n = 411)	Lenalidomide (n = 231)	Placebo (n = 229)
Median follow-up, months				<u>81.0⁷</u>
ISS stage				
1, n (%)				
2 (%)				
3 (%)				
Missing, n (%)				
β 2 microglobulin, n (%)				
< 3.5 mg/dL				
\geq 3.5 mg/dL				
Missing				

^aHigh percentage of missing data for ISS stage at time of diagnosis because these data were collected retrospectively in CALGB since the study started at randomisation to maintenance therapy. Although data are available at other timepoints (pre-ASCT and post-ASCT), these are not considered clinically relevant due to the impact of treatment. [%], % excluding missing cases; ISS, international staging system; ITT, intent to treat.

Pale red: imbalances between trials and dark red highlight denotes imbalances between treatment arms.

Source: data on file.

6 Comparison of induction therapy data from Myeloma XI and CALGB

Table 5 presents a side-by-side comparison of induction therapy from the ITT population of CALGB (data cutoff: 1 February 2016) and the decision problem cohort of Myeloma XI (data cutoff: 23 October 2017).

The induction pathways used in Myeloma XI and CALGB differed; whereas patients enrolled in Myeloma XI were randomised to induction, those entering the CALGB trial received induction therapy prior to the trial with no standardised protocol. Prior lenalidomide and prior thalidomide were mutually exclusive in Myeloma XI but not in CALGB. A higher proportion of patients in Myeloma XI received prior lenalidomide than in CALGB.

Table 5. Comparison of induction therapy: Myeloma XI and CALGB

	Myeloma XI (decision problem cohort)		CALGB (ITT population)	
Data cut	23 October 2017		1 February 2016	
Characteristic	Lenalidomide (n = 621)	Observation (n = 411)	Lenalidomide (n = 231)	Placebo (n = 229)
Median follow-up, months	████████		81.9 ⁷	81.0 ⁷
Prior lenalidomide:				
Yes	██████	██████	80 (34.6%)	78 (34.1%)
Prior thalidomide:				
Yes	██████	██████	102 (44.2%)	104 (45.4%)
Prior lenalidomide (no thalidomide, no bortezomib):				
Yes	██████	██████	██████	██████
Prior thalidomide (no lenalidomide, no bortezomib):				
Yes	██████	██████	██████	██████

Pale red highlight denotes imbalances between trials ; ITT, intent to treat.

Source: data on file.

6.1 Comparison of response post-ASCT from Myeloma XI and CALGB

Table 6 presents a side-by-side comparison of the post-ASCT response rates from the ITT population of CALGB (data cutoff: 19 October 2016) and the decision problem cohort of Myeloma XI (data cutoff: 23 October 2017).

A higher proportion of patients in Myeloma XI achieved a CR or VGPR compared with CALGB. Both trials assessed response in using criteria from the International Myeloma Working Group (IMWG); however, initially CALGB used criteria based on the European Group for Bone and Marrow Transplantation (EBMT) before switching to the IMWG criteria for central review in 2009.

As CALGB was a registration trial, complete response had to be confirmed with a bone marrow aspiration whereas in Myeloma XI, which was more pragmatic, complete response was confirmed in some patients without a bone marrow culture [data on file]. The percentage of patients with a complete response without bone marrow confirmation is summarised in Table 6.

There was also discrepancy in response outcomes between treatment arms in CALGB; a higher proportion of patients allocated to the placebo arm of the maintenance phase had a CR or VGPR than in the lenalidomide arm.

Table 6. Comparison of response post-ASCT prior to maintenance therapy: Myeloma XI and CALGB

	Myeloma XI (decision problem cohort)		CALGB (ITT population)	
Data cut	23 October 2017		19 October 2016	
Characteristic	Lenalidomide (n = 621)	Observation (n = 411)	Lenalidomide (n = 231)	Placebo (n = 229)
Median follow-up, months				
Response post-ASCT:				
CR				
CR without BM				
VGPR				
PR				
MR				
PD				
SD				
Unable to assess/not evaluable				
Missing				
Response category post-ASCT:				
CR or VGPR				
Not CR or VGPR				
Missing/unable to assess/not evaluable				

Pale red highlight denotes differences between trials and dark red highlight denotes imbalances between treatment arms.

ASCT, autologous stem cell transplantation; BM, bone marrow; CR, complete response; ITT, intent to treat; MD, minimal disease; NE, not evaluable; PD, progressive disease; PR, partial response; SD, stable disease; VGPR, very good partial response.

Source: data on file.

7 Summary of outcomes from the CALGB, GIMEMA and Myeloma XI trials

Clinical efficacy and safety outcomes from CALGB, GIMEMA and Myeloma XI are summarised in Table 7. It should be noted that the primary endpoint of CALGB was reported as time to progression, which was defined as time to progression or death from any cause after transplantation and aligns with the definition of PFS from the IMWG⁹ and FDA.¹⁰

Table 7. Summary of the clinical and safety outcomes from CALGB, GIMEMA and Myeloma XI

Trial, country	Intervention vs comparator	PFS	OS	Grade 3/4 AEs in lenalidomide arm, %	
CALGB USA N = 460 ⁸	Lenalidomide vs placebo	Median TTP, ^a months (95% CI)	Median OS, months (95% CI)	Anaemia, 4.8	
		57.3 (44.2–73.3) vs 28.9 (23.0–36.3)	113.8 (100.4–not reached) vs 84.1 (73.8–106.0)	Neutropenia, 15.5	
		HR, 0.57 (95% CI, 0.46–0.71); p < 0.001	HR, 0.61 (95% CI, 0.46–0.8); p < 0.0004	Thrombocytopenia, 6.9	
		Primary endpoint	Secondary endpoint		
		Median follow-up, months: 91			
GIMEMA5 Italy and Israel N = 273	Lenalidomide vs no maintenance	Median PFS, months	HR for death	Anaemia, 1.7	
		41.9 vs 21.6	0.64 (95% CI: 0.36–1.15); p = 0.14	Neutropenia, 23.3	
		HR, 0.47 (95% CI, 0.33–0.65); p < 0.001		Thrombocytopenia, 4.3	
		Primary endpoint	Secondary endpoint		
		Median follow-up, months: 51.2			
Myeloma XI ¹¹ UK N = 1032	Lenalidomide vs observation	Median PFS (95% CI), months	Median OS, months (95% CI)	██████	
		██████	██████	██████	
		Co-primary endpoint	Co-primary endpoint		
		Median follow-up, months: ██████			

^aDefined as time to progressive disease or death from any cause after transplantation. This definition aligns with the definition of PFS provided by both the IMWG⁹ and FDA.¹⁰ ^bJuly 2010 data cut-off

^cOctober 2011 data cut-off.

AE, adverse event; CI, confidence interval; FDA, Food and Drug Administration; HR, hazard ratio; IMWG, International Myeloma Working Group; NA, not applicable; NR, not reached; OS, overall survival; PFS, progression-free survival; TTP, time to progression.

8 Discussion

8.1 Summary of the CALGB, GIMEMA and Myeloma XI trials

A brief overview of the CALGB, GIMEMA and Myeloma XI trials is given in Table 8.

Table 8. Comparison of CALGB, GIMEMA and Myeloma XI

	CALGB ²	GIMEMA ⁵	Myeloma XI ⁶
UK patients as proportion of study (%)	0 ^a	0 ^a	100
Study powered for detecting survival difference?	No	No	Yes ^b
Double ASCT	No	Yes	No
Lenalidomide dose cycle	1–28/28-day cycle	1–21/28-day cycle	1–21/28-day cycle
Time at which randomised to maintenance	Post-ASCT	Study enrolment	Post-ASCT
Patients allowed to switch to active treatment before PD allowed ^c	Yes	No	No

^aCALGB 100101 was a US study and GIMEMA an Italian/Israeli.

^bCo-primary endpoint. Myeloma XI is the only RCT to date powered to detect a survival difference in patients treated with maintenance therapy.

^cConfounds survival analysis.

Shaded cells represent trial attributes consistent with decision problem, reflective of anticipated UK clinical practice or desirable statistical feature, as detailed in footnotes.

ASCT, autologous stem cell transplant; PD, progressive disease; RCT, randomised controlled trial.

8.2 GIMEMA is not suitable to support longer-term survival estimates

An overview of the study design of the GIMEMA trial is given in Figure 1 and Table 1. The evidence base for the GIMEMA trial is provided by Palumbo *et al.*, 2014,⁵ which presents a comparison between lenalidomide maintenance and no maintenance. This comparison includes patients randomised to ASCT +/- maintenance, and those randomised to MPR +/- maintenance.

Patients who received MPR are not relevant to the decision problem, and data from these patients are therefore excluded from any further analyses. This discussion will focus exclusively on the patient population randomised to ASCT +/- maintenance.

Owing to the study design of the GIMEMA trial, the comparison between maintenance and no maintenance in the groups randomised to ASCT at the first randomisation stage is invalid. The GIMEMA trial has a 2x2 factorial randomised trial (stage 1 intervention: ASCT vs MPR followed by a second random allocation to stage

2 intervention: maintenance vs no maintenance) meaning that patients were assigned to one of four study groups at study recruitment Figure 1 and Table 2.

Eligibility for maintenance for patients allocated to ASCT +/- lenalidomide maintenance was confirmed at the end of the ASCT phase. The period between randomisation at study enrolment and the start of maintenance therapy was approximately one year during which patients underwent induction therapy followed by high-dose melphalan and ASCT. Only patients who responded to induction received an ASCT, and maintenance is only given to patients who responded to ASCT. The composition of the two groups randomised to ASCT +/- maintenance include the populations outlined in Table 9. Not all patients in the ASCT +/- maintenance comparison received the intervention of interest in this appraisal. Furthermore, patients who did not receive an ASCT nor maintenance received treatment locally post-progression as per the protocol, potentially confounding the results. Data for the subpopulation of patients who received an ASCT +/- maintenance was not published as part of the primary analysis by Palumbo *et al.*, 2014.⁵

Table 9. Summary of the patient populations included in GIMEMA

Randomised	Possible outcomes during trial period		Groups analysed
ASCT + maintenance	Patients who succeeded induction, succeeded ASCT and received maintenance	Aligned with decision problem	All patients included in the analysis: ASCT + maintenance OS, PFS
	Patients who failed induction, received no ASCT, received no maintenance	* Fail the decision problem	
	Patients who responded to induction, failed ASCT, and received no maintenance	* Fail the decision problem	
ASCT + no maintenance	Patients who succeeded induction, succeeded ASCT	Aligned with decision problem	All patients included in the analysis: ASCT + no maintenance OS, PFS
	Patients who failed induction and received no ASCT	* Fail the decision problem	

ASCT, autologous stem cell transplant; ITT, intention-to-treat; OS, overall survival; PFS, progression-free survival.

Additional analyses that included unpublished data from the GIMEMA trial were performed by McCarthy *et al.*, 2017,¹ and included only patients who responded to

induction therapy and an ASCT, and are thus aligned with the decision problem. McCarthy *et al.*, 2017, reported a meta-analysis comparing lenalidomide maintenance with no maintenance therapy in patients who had received an ASCT based on evidence from three trials (CALGB, IFM2005-02 and GIMEMA), and reported the baseline characteristics (Supplementary table 1) and efficacy data for the subpopulation of patients in the GIMEMA trial relevant to the decision problem.¹ However:

1. The results of the meta-analysis include IFM2005-02; as such, they are not valid in the context of this appraisal.
2. Kaplan–Meier data were not presented for the individual studies included in the meta-analysis; therefore, the full results for the patient population who received an ASCT +/- maintenance therapy are still not available.
3. The meta-analysis reports hazard ratios for OS and PFS for each included study. The HR for OS from the GIMEMA trial was 0.72 (95% CI, 0.37–1.38); however, it is not possible to assess how this subpopulation compares with Myeloma XI owing to the differences in patient population and study design. Furthermore, as this is a post hoc analysis dependent on whether patients failed ASCT or not, there is a possibility that randomisation was broken following exclusion of patients who failed induction and ASCT.

In addition, the baseline characteristics for the subpopulation of patients in the GIMEMA trial who received ASCT +/- maintenance therapy suggest that approximately two-thirds of patients received two ASCTs (Supplementary table 1), although the study publications are not transparent in this respect. Given the differences in the study design and the wider patient population of the GIMEMA trial, as well as the mix of treatments that are broader than the decision problem, the GIMEMA and Myeloma XI trials are not comparable.

8.3 Appropriateness of data from CALGB to support longer term survival extrapolations in patients who received lenalidomide post-ASCT

The study designs of the CALGB 100104 and Myeloma XI trials are summarised in Figure 1, Table 1 and Table 8, with both trials presenting outcomes in broadly similar patient populations aligned with the decision problem.

Key differences in study design that warrant adjustment include lenalidomide dose regimen (CALGB, 10 mg on days 28/28 with the option of dose escalation to 15 mg after 3 months vs Myeloma XI, 10 mg on days 21/28) and the permitted switch to active treatment from placebo in the CALGB. Patient baseline characteristics and demographics were also compared between the CALGB and Myeloma XI trials and are further discussed.

8.4 Patient demographics between the Myeloma XI and CALGB trial

A comparison of patient demographics is summarised in Table 10. There was a slight imbalance in the proportion of patients aged 60 years and under between trials, with a lower proportion of patients in Myeloma XI under 60 years of age than in CALGB (48.0% vs 57.4%). Eligible patients were aged 17–70 years in CALGB, whereas there was no upper age limit in Myeloma XI.

There was also a higher proportion of men enrolled in Myeloma XI compared with CALGB (62.1% vs 54.3%); however, sex is not considered clinically as a prognostic factor.

Table 10. Comparison of patient demographics: Myeloma XI and CALGB, summary

Patient demographics	Demographic	Myeloma XI	CALGB
	Median age < 60 years of age, %	██████	██████
	Male sex, %	██████	██████

Source: data on file.

8.5 Imbalances in ISS stage between the Myeloma XI and CALGB trial

There was an imbalance in ISS stage at study entry/at diagnosis between Myeloma XI and CALGB. ISS data were missing for approximately one-quarter of patients in CALGB; therefore, the data presented in Table 11 is adjusted for missing data in both treatment arms. A lower proportion of patients in Myeloma XI had ISS stage 1 disease

than CALGB (36.2% vs 45.1%), balanced by a relatively higher proportion of patients at ISS stage 2 (41.7% vs 32.0%).

Table 11. Comparison of disease characteristics: Myeloma XI and CALGB, summary

Disease characteristics	Characteristic	Myeloma XI (at study entry)	CALGB (collected retrospectively at diagnosis)
	ISS stage, % ^a		
	1	██████	██████
	2	██████	██████
	3	██████	██████
	β2 microglobulin, % ^a		
	< 3.5 mg/dL	██████	██████
	≥ 3.5 mg/dL	██████	██████

^aAdjusted for missing data.

Source: data on file.

As discussed in section 5, ISS scores were calculated for each patient at study recruitment in Myeloma XI, whereas in CALGB, ISS stage at diagnosis was collected retrospectively (methods unclear) when patients underwent randomisation to maintenance. This may account for the high proportion of missing ISS data (27%) in CALGB, potentially adding uncertainty. ISS data are considered most clinically relevant at the time of diagnosis or following treatment relapse, and are therefore presented here; however, the comparison of ISS scores between Myeloma XI and CALGB should be interpreted with caution.

β2 microglobulin at a threshold of 2.5 mg/dL using data collected at registration (i.e. post-ASCT) was a stratification factor in CALGB. This threshold is not considered clinically relevant in the context of myeloma staging. β2 microglobulin is a key component of the ISS staging criteria (Table 12), together with serum albumin; however, the clinically relevant stratification for β2 microglobulin is < 3.5 mg/dL, 3.5–5.4 mg/dL and > 5.4 mg/dL as used in Myeloma XI. Therefore, an imbalance in ISS stage in CALGB may not be unexpected and may be the result of using the 2.5 mg/dL cut-off for stratification compounded with using values obtained at randomisation (i.e. post-ASCT) which were accurate but not clinically meaningful.

Table 12. ISS staging criteria

ISS stage	β 2 microglobulin, mg/L	Albumin, g/dL
1	< 3.5	\geq 3.5
2	< 3.5	< 3.5
	3.5–5.4	\geq 3.5
3	> 5.4	–

ISS, International Staging System.

Source: myeloma.org/international-staging-system-iss-revised-iss-r-iss.

There is also an imbalance in ISS score between treatment arms in CALGB, with a higher proportion of patients in the placebo arm having ISS stage 1 disease at diagnosis (51.2% vs 39.1%) than in the lenalidomide arm, where patients are more equally distributed between ISS stage 1 and 2 (ISS stage 1, 39.1%; ISS stage 2, 36.0%). It is unclear what drives this imbalance but there is a possibility that use of clinically inappropriate stratification for β 2 microglobulin may have contributed to unanticipated and unknown imbalances in other prognostic factors.

8.6 Imbalances in induction therapy between Myeloma XI and CALGB

A comparison of induction therapy is summarised in Table 13. A higher proportion of patients in Myeloma XI received prior lenalidomide as induction therapy than in CALGB (69.3% vs 34.3%; difference, +35%).

Table 13. Comparison of induction therapy: Myeloma XI and CALGB, summary

Induction therapy	Induction	Myeloma XI	CALGB
	Prior lenalidomide, %	██████	██████
	Prior thalidomide, %	██████	██████

Source: data on file.

All patients enrolled in Myeloma XI were allocated to receive lenalidomide or thalidomide as induction therapy as per the protocol (Appendix), whereas in CALGB, while patients received induction therapy prior to the study enrolment; there was no stipulation as to the specific induction therapy in CALGB, and most patients received previous treatment with a lenalidomide- or thalidomide-based regimen (~80%).

8.7 Imbalances in post-ASCT response category between Myeloma XI and CALGB

A comparison of response outcomes post-ASCT is summarised in Table 14. A higher proportion of patients in Myeloma XI achieved a CR or VGPR compared with CALGB (84.9% vs 61.1%; difference, +23.8%).

Table 14. Comparison of response post-ASCT: Myeloma XI and CALGB, summary

Response post-ASCT	Response post-ASCT, %	Myeloma XI	CALGB
	CR or VGPR		84.9%
Not CR or VGPR		15.1%	38.9%

ASCT, autologous stem cell transplantation; CR, complete response, VGPR, very good partial response.

Source: data on file.

The difference in post-ASCT response may arise from the differences in induction therapy between the two trials. Patients enrolled in Myeloma XI who achieved less than a VGPR to induction with cyclophosphamide, thalidomide and dexamethasone (CTD) or cyclophosphamide, lenalidomide and dexamethasone (CRD) were randomised to receive intensification therapy with bortezomib, cyclophosphamide and dexamethasone (VCD) or no further treatment prior to ASCT (Appendix). It is expected that patients who received intensification therapy would achieve a better response going into ASCT, and hence receipt of intensification therapy was a stratification factor for the maintenance phase of Myeloma XI. This may provide some explanation as to why a higher proportion of patients in Myeloma XI achieved a better post-ASCT response than in CALGB, where induction therapy was completed prior to trial enrolment.

Assessment of response in CALGB was initially based criteria on from the EBMT; however, following updated guidance from the IMWG in June 2009, the response criteria were adjusted. The Central Review Committee followed the IMWG criteria, which is aligned with Myeloma XI. The response criteria from the EBMT and IMWG criteria are summarised in Appendix.

In addition, as CALGB is a registration study, the response criteria may have been applied more stringently than in Myeloma XI. Myeloma XI was more pragmatic with some patients considered to achieve a complete response without bone marrow

confirmation (approximately 35% of patients with a CR); these patients could have been classified as achieving a VGPR rather than a CR if the definitions were applied more stringently [data on file].

There was discrepancy in response between treatment arms in CALGB; a higher proportion of patients allocated to the placebo arm of the maintenance phase had a CR or VGPR than in the lenalidomide arm (66.8% vs 55.4%). A higher proportion of patients in the placebo arm had ISS stage 1 disease at diagnosis than in the lenalidomide arm; therefore, the better response to ASCT seen in the placebo arm may reflect that patients in the placebo were fitter at diagnosis.

A. ISS stage pre- and post-ASCT

	Myeloma XI (decision problem cohort)		CALGB (ITT population)	
Data cut	23 October 2017		1 February 2016	
Characteristic	Lenalidomide (n = 621)	Observation (n = 411)	Lenalidomide (n = 231)	Placebo (n = 229)
ISS stage	Pre-ASCT			
1, n (%)	N/A	N/A		
2 (%)	N/A	N/A		
3 (%)	N/A	N/A		
Missing, n (%)	N/A	N/A		
ISS stage	Post-ASCT			
1, n (%)	N/A	N/A		
2 (%)	N/A	N/A		
3 (%)	N/A	N/A		
Missing, n (%)	N/A	N/A		

ASCT, autologous stem cell transplant; ISS, International Staging System; ITT, intention-to-treat.

Source: Data on file

B. Patient and disease characteristics from GIMEMA (taken from McCarthy et al. 2017¹)

Supplementary table 1. Patient demographics and baseline characteristics in the subpopulation of patients who received an ASCT as part of the GIMEMA trial

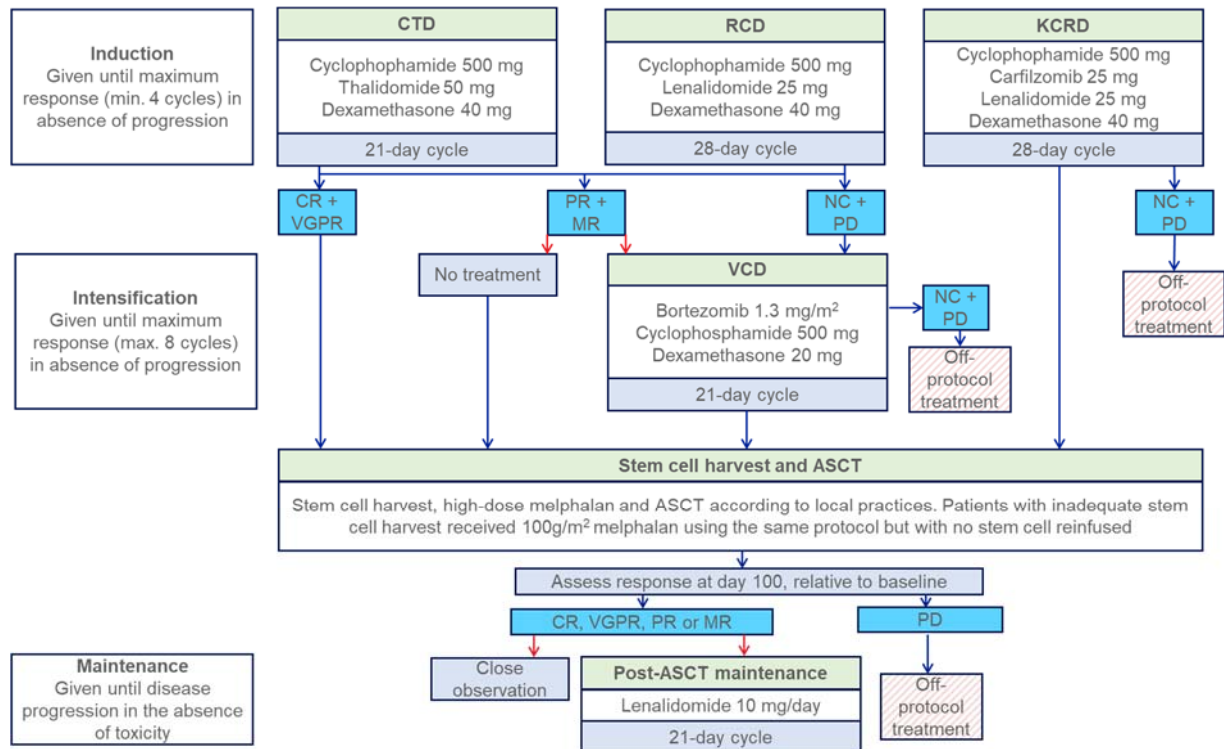
Characteristic	Lenalidomide (n = 67)	Observation (n = 67)
Median age, years (Min, Max) < 60 years, n (%)	57.9 (35.5-65.1) 43 (64.2)	56.2 (40.1-66.0) 48 (71.6)
Sex, n (%) Male	32 (47.8)	39 (58.2)
Race, n (%) White Other Missing	65 (97.0) 2 (3.0) 0	67 (100) 0 0
Extramedullary disease, n (%) Yes No Missing	10 (14.9) 57 (85.1) 0	10 (14.9) 57 (85.1) 0
ISS stage ^b , n (%) I II III Missing	35 (52.2) 24 (35.8) 8 (11.9) 0	41 (61.2) 17 (25.4) 9 (13.4) 0
Creatinine clearance categories < 50 mL/min ≥ 50 mL/min Missing	4 (6.0) 63 (94.0) 0	3 (4.5) 64 (95.5) 0
Lactate dehydrogenase, n (%) Normal > ULN	62 (92.5) 5 (7.5)	63 (94.0) 4 (6.0)
Prior induction therapy, n (%) Lenalidomide containing	67 (100)	67 (100)
Number of ASCTs, n (%) 1 2 Missing	24 (35.8) 43 (64.2) 0	23 (34.3) 43 (64.2) 1 (1.5)
Response after ASCT ^c , n (%) CR or VGPR CR PR/SD/PD Not evaluable/missing	24 (35.8) 4 (6.0) 37 (55.2) 6 (9.0)	21 (31.3) 6 (9.0) 41 (61.2) 5 (7.5)

^a Age was collected at time of diagnosis. ^b ISS Stage was calculated based on β 2-microglobulin and albumin values at diagnosis. ^c Based on central review ASCT, autologous stem cell transplantation; CR, complete response; ISS, international staging system; ITT, intent to treat; PD, progressive disease; PR, partial response; SD, stable disease; ULN, upper limit of normal; VGPR, very good partial response.

Source: McCarthy et al, 2017.¹

C. Myeloma XI, detailed study overview

Supplementary figure 1. Myeloma XI, detailed study overview



Source: Jackson *et al*, 2019.⁶

ASCT, autologous stem cell transplant; CR, complete response, CTD, cyclophosphamide, thalidomide and dexamethasone; KCRD, carfilzomib, cyclophosphamide, lenalidomide and dexamethasone; MR, minimal response, NC, no change; PD, progressive disease; PR, partial response; RCD, lenalidomide cyclophosphamide and dexamethasone; VGPR, very good partial response.

D. Comparison of EBMT and IMWG response criteria

IMWG Response	Criteria ^a	EBMT Criteria
CR	Negative immunofixation on the serum and urine and disappearance of any soft tissue plasmacytomas and $\leq 5\%$ plasma cells in the bone marrow ^b	Absence of the original monoclonal paraprotein in serum and urine by immunofixation, for a minimum of 6 weeks. The presence of oligoclonal bands consistent with oligoclonal immune reconstitution does not exclude CR. $< 5\%$ plasma cells in a bone marrow aspirate/trephine bone biopsy, if performed. If absence of monoclonal protein is sustained for 6 weeks, bone marrow not repeated, except in patients with non-secretory myeloma where the marrow examination must be repeated after an interval of at least 6 weeks to confirm CR. No increase in size or number of lytic bone lesions (development of a compression fracture does not exclude response). Disappearance of soft tissue plasmacytomas.
sCR	CR as defined above plus: Normal FLC ratio Absence of clonal cells in bone marrow ^b by immunohistochemistry or immunofluorescence ^c	N/A
VGPR	Serum and urine M-component detectable by immunofixation but not on electrophoresis, or 90 or greater in serum M-component plus urine M-component < 100 mg per 24h	N/A
PR	$\geq 50\%$ reduction of serum M-protein and reduction in 24-h urinary M-protein by $\geq 90\%$ or to < 200 mg per 24 h. If serum and urine M-protein unmeasurable, $\geq 50\%$ decrease in difference between involved and uninvolved FLC levels required in place of the M-protein criteria If serum and urine M-protein unmeasurable, and serum free light assay also unmeasurable, $\geq 50\%$ reduction in plasma cells required in place of M-protein, provided baseline bone marrow plasma cell percentage was $\geq 30\%$. In addition to the above criteria, if present at baseline, a $\geq 50\%$ reduction in soft tissue plasmacytomas required	$> 50\%$ reduction in the level of the serum monoclonal paraprotein, maintained for a minimum of 6 weeks. Reduction in 24 h urinary light chain excretion either by $> 90\%$ or to < 200 mg, maintained for a minimum of 6 weeks. For patients with non-secretory myeloma only, $> 50\%$ reduction in plasma cells in a bone marrow aspirate and on trephine biopsy, if biopsy is performed, maintained for a minimum of 6 weeks. $> 50\%$ reduction in the size of soft tissue plasmacytomas (by radiography or clinical examination). No increase in size or number of lytic bone lesions (development of a compression fracture does not exclude response).

(Cont'd)

IMWG Response	Criteria ^a	EBMT Criteria
MR		25–49% reduction in the level of the serum monoclonal paraprotein maintained for a minimum of 6 weeks. 50–89% reduction in 24 h urinary light chain excretion, which still exceeds 200 mg/24 h, maintained for a minimum of 6 weeks. For patients with non-secretory myeloma only, 25–49% reduction in plasma cells in a bone marrow aspirate and on trephine biopsy, if biopsy is performed, maintained for a minimum of 6 weeks. 25–49% reduction in the size of soft tissue plasmacytomas (by radiography or clinical examination). No increase in the size or number of lytic bone lesions (development of a compression fracture does not exclude response).
SD/NC	Not meeting criteria for CR, VGPR, PR or progressive disease (not recommended for use as an indicator of response; stability of disease is best described by providing the time to progression estimates)	Not meeting the criteria of either minimal response or progressive disease.

^aAll response categories require two consecutive assessments made at anytime before the institution of any new therapy; complete and PR and SD categories also require no known evidence of progressive or new bone lesions if radiographic studies were performed. Radiographic studies are not required to satisfy these response requirements.

^bConfirmation with repeat bone marrow biopsy not needed.

^cPresence/absence of clonal cells is based upon the k/λ ratio. An abnormal k/λ ratio by immunohistochemistry and/or immunofluorescence requires a minimum of 100 plasma cells for analysis. An abnormal ratio reflecting presence of an abnormal clone is k/λ of $>4:1$ or $<1:2$. Alternatively, the absence of clonal plasma cells can be defined based on the investigation of phenotypically aberrant PC. The sensitivity level is 10–3 (less than one phenotypically aberrant PC within a total of 1000 Pc). Examples of aberrant phenotypes include (1) CD38 +dim and CD56+ strong and CD19– and CD45–; (2) CD38+dim and CD138+ and CD56++ and CD28+; (3) CD138+, CD19– CD56++, CD117+. CR, complete response; EBMT, European Group for Bone Marrow Transplant; FLC, free light chain; IMWG, International Myeloma Working Group; MR, minimal response; NC, no change; PR, partial response; SD, stable disease; sCR, stringent complete response; VGPR, very good partial response.

Source: Durie *et al*, 2006¹² and Bladé *et al*, 1998.¹³

References

1. McCarthy PL, Holstein SA, Petrucci MT *et al.* Lenalidomide maintenance after autologous stem-cell transplantation in newly diagnosed multiple myeloma: a meta-analysis. *J Clin Oncol* 2017;35:3279–89.
2. McCarthy PL, Owzar K, Hofmeister CC *et al.* Lenalidomide after stem-cell transplantation for multiple myeloma. *N Engl J Med* 2012;366:1770–81.
3. Alliance for Clinical Trials in Oncology. Revlimid (lenalidomide). Clinical study report CALGB 100104. [Data on file].
4. European Medicines Agency. Summary of opinion (post authorisation). Revlimid (lenalidomide), March 2019. Available from: https://www.ema.europa.eu/en/documents/smop/chmp-post-authorisation-summary-positive-opinion-revlimid-ii-102-g_en.pdf. (Accessed 27 February 2020)
5. Palumbo A, Cavallo F, Gay F *et al.* Autologous transplantation and maintenance therapy in multiple myeloma. *N Engl J Med* 2014;371:895–905.
6. Jackson GH, Davies FE, Pawlyn C *et al.* Lenalidomide maintenance versus observation for patients with newly diagnosed multiple myeloma (Myeloma XI): a multicentre, open-label, randomised, phase 3 trial. *Lancet Oncol* 2019;20:57–73.
7. European Medicines Agency. Revlimid. Summary of product characteristics. June 2020. Available from: https://www.ema.europa.eu/en/documents/product-information/revlimid-epar-product-information_en.pdf. (Accessed 8 October 2020).
8. Holstein SA, Jung SH, Richardson PG *et al.* Updated analysis of CALGB (Alliance) 100104 assessing lenalidomide versus placebo maintenance after single autologous stem-cell transplantation for multiple myeloma: a randomised, double-blind, phase 3 trial. *Lancet Haematol* 2017;4:e431–e442.
9. Rajkumar SV, Harousseau JL, Durie B *et al.* Consensus recommendations for the uniform reporting of clinical trials: report of the International Myeloma Workshop Consensus Panel 1. *Blood* 2011;117:4691–5.
10. Food and Drug Administration. Clinical Trial Endpoints for the Approval of Cancer Drugs and Biologics Guidance for Industry. Available at, <https://www.fda.gov/media/71195/download> (accessed, 6 December 2019). 2018.
11. Celgene. Myeloma XI trial patients-level data analysis write-up. 2019 (data on file).
12. Durie BG, Harousseau JL, Miguel JS *et al.* International uniform response criteria for multiple myeloma. *Leukemia* 2006;20:1467–73.
13. Bladé J, Samson D, Reece D *et al.* Criteria for evaluating disease response and progression in patients with multiple myeloma treated by high-dose therapy and haemopoietic stem cell transplantation. Myeloma Subcommittee of the EBMT.

European Group for Blood and Marrow Transplant. *Br J Haematol* 1998;102:1115-23.

Issue 2. Survival estimates in the economic model based on Myeloma XI data, with CALGB 100104 data (adjusted to reflect the Myeloma XI population as closely as possible, and conditional on the underlying survival of patients in Myeloma XI) used to help longer-term extrapolation, with all methods explained in detail

1 Methods

Based on the qualitative comparison of CALGB's and Myeloma XI (Issue 1, above), potential differences were found between the two studies' populations, with regards to prognostic factors such as age and gender, ISS scores at diagnosis and in the rates and distribution of response after ASCT. In addition to trial population, Myeloma XI and CALGB differed by the lenalidomide dose used in each, lenalidomide 21/28 days and lenalidomide 28/28 days respectively.

Analyses were conducted to explore and adjust for the impact of potential population differences between the CALGB study and Myeloma XI. (DSU Report 18. Phillippo et al, 2016), adjusting the estimates of treatment effect for lenalidomide maintenance from CALGB to Myeloma XI. The impact of potential heterogeneity and treatment effect modifiers, covariates and imbalances in prognostic factors between CALGB and Myeloma XI was assessed to ensure that adjustments would have a material impact on the efficacy estimated for CALGB.

CALGB and Myeloma XI include a placebo arm and an 'observation' arm, respectively. Although it could be argued that the two control modalities may have an impact of outcomes in the respective studies, the outcome of interest in this analysis, overall survival (OS) is unlikely to be affected by placebo or to be prone to observation biases. For this reason, the two control arms should be considered equivalent all other factors considered. Therefore, this analysis was conducted using the common comparator (placebo / observation).

In addition, because the CALGB study hit the superiority boundary in the early follow up (2 years), all patients initially randomised to placebo that had not progressed after ASCT (thus still fulfilling eligibility for maintenance) were offered to switch to active

maintenance therapy. For methods relating to the adjustment for treatment switching, please See Issue 7 below. In the remainder of this Section, all analyses were conducted using the rank preserving structural failure model (RPSFT) adjustment for CALGB.

Adjustments were conducted using two statistical approaches as detailed here:

- Propensity score weights (PW) adjusted analysis. Indirect comparison methods were used to generalise estimates from CALGB to the potentially different population in Myeloma XI. This analysis utilised patient level data from both trials to estimate the probability of each patient in CALGB to be in Myeloma XI. The propensity scores were then used to reweight the patients in CALGB to match the Myeloma XI population. The approach has been widely used in medicine to remove potential biases from treatment comparisons (Stuart et al 2001);
- Matched adjusted indirect comparison (MAIC) analysis. This method utilised patient level data from CALGB, matched to aggregate data from Myeloma XI. The use of MAIC in indirect comparisons has been described in the NICE Methods Guidance (Phillippo et al, DSU Report 18). MAICs have been extensively used in HTA submissions, and specifically in the synthesis of aggregate data, to account for potential variations in treatment effect modifiers between studies.

The propensity score adjusted analysis was chosen as the base case. The motivation for this preference was justified based on some of the necessary assumptions for MAIC being potentially violated by the data. MAIC assumes shared common treatment effect modifier(s). During the assessment of trial population characteristics, prognostic factors and treatment effect modifiers in the two studies, some characteristics were found to be treatment effect modifiers in one of the two studies only (Section 1.3 below). In addition, prognostic factors were also found to differ between studies, whilst some prognostic factors were potentially imbalanced between arms in CALGB. Compared with PW, MAIC may be prone to amplify potential biases originating from this heterogeneity.

Imbalances within CALGB and between CALGB and Myeloma XI may have originated from true differences, from aspects of the study design and randomisation procedures or from differences in the data collection or classification, however, at this point it is not possible to ascertain which reason is the most plausible.

To address this uncertainty, scenarios for both PW and MAIC were generated based on different sets of covariates used in the adjustment models. Response after ASCT, in particular, was imbalanced between arms in CALGB. Likewise, uncertainty was present because the two studies differed by the methods used to measure response. However, owing to depth of response being potentially correlated with prior exposure to lenalidomide, and prior exposure to lenalidomide also being an adjustment factor to match the randomised prior treatments in Myeloma XI, the models were estimated both without and with adjustment for response after ASCT.

Therefore, the following analyses were conducted:

- 1 Propensity scores weight-adjusted analyses using ISS scores, age, gender and prior use of lenalidomide (no adjustment for response to ASCT)
- 2 MAIC-adjusted analyses using ISS scores, age, gender and prior use of lenalidomide (no adjustment for response to ASCT)

Two alternative scenarios were also conducted, including response after ASCT in the set of covariates:

- 3 Propensity scores weight-adjusted analyses using ISS scores, age, gender and prior use of lenalidomide and response to ASCT
- 4 MAIC-adjusted analyses using ISS scores, age, gender and prior use of lenalidomide and response to ASCT.

The rationale for developing scenarios adjusted for ASCT response is provided in Section 1.3.

It is important to note that both MAIC and propensity scores adjusted analysis assume the exchangeability of the two doses of lenalidomide, as this characteristic cannot be matched with either PW or MAIC. This is because Myeloma XI did not

include patients treated with the 28/28 dose and CALGB, conversely, had no patients treated with the 21/28 dose.

Therefore, a third scenario was developed, using an indirect comparison between Myeloma XI and CALGB, including adjustments for covariates, a term for lenalidomide and an interaction term for 'treatment x trial', as this approach is the only model that allows the consideration of the difference between the two doses. A weighted pooled analysis using patient level data from Myeloma XI combined with the propensity weighted analysis in CALGB was conducted, including covariates for trial and treatment. An interaction term was included to account for any difference in treatment effect between trials.

All propensity scores analyses were conducted using the 'WeightIt' package in the statistical program R.(Greifer et al, 2020; R Core team)

Methods used in adjusted comparison of lenalidomide maintenance compared with placebo are described in detail in the Sections below.

1.1 Propensity scores weighting

The CALGB trial was reanalysed using propensity score weights assigned to each patient in the study, to adjust for differences in the distribution of prognostic factors and treatment effect modifiers between CALGB and Myeloma XI.

Propensity scores were the probability of treatment assignment (or in this case trial assignment) to Myeloma XI as a function of a set of observable covariates for each patient in CALGB. (Rosenbaum et al, 1987) A model-based approach was taken to estimate propensity scores, using logistic regression to predict the conditional odds of being enrolled into CALGB or Myeloma XI, given patient covariates.

Myeloma XI was used as the target population for the matching. The population average treatment effect for the treated (PATE) estimand was used for the weighting, which assigns a weight of $w=1$ to all patients in Myeloma XI and a weight w to all CALGB patients, proportional to the probability that they are under/overrepresentation in Myeloma XI, as in the equation

$$w_{(MXI,CALGB)} = \left(1, \frac{PS}{1 - PS}\right)$$

Where PS = *propensity score*.

Covariates used in the analysis are described in Section 1.3 below; in brief, they were age, gender, ISS score at baseline and prior lenalidomide exposure, but not prior thalidomide exposure and β 2-microglobulin; in addition, response after ASCT was also used in one scenarios.

1.2 Matched adjusted indirect comparison methods

An alternative approach was used to derive propensity score weighting, based on matched adjusted indirect comparisons (MAIC) methods. (Philippo et al, 2016; Signorovitch et al, 2012) The MAIC used individual level patient data from CALGB to match baseline summary statistics of Myeloma XI. Treatment outcomes were then compared after matching.

A propensity score logistic regression model was used to estimate the odds of being enrolled into Myeloma XI or CALGB. The method of moments was used to consolidate data from individual patients from CALGB and aggregate data from Myeloma XI. Similarly to propensity scores weights in Section 1.1 above, the MAIC propensity scores were used to up or down-weight CALGB patients in proportion to their representativeness of those in Myeloma XI. The weights were derived as the inverse odds of being in Myeloma XI.

1.3 Choice of matching variables

Drawing on the qualitative comparison of the two studies made in Issue 1. and summarised in Table 15 below, the following prognostic factors and treatment effect (TE) modifiers were studied:

- Age / Gender
- ISS stage at diagnosis (CALGB); at trial entry (Myeloma XI)
- Use of prior lenalidomide / thalidomide (Resulting from randomisation (Myeloma XI) vs historical data (CALGB))
- Type of response after ASCT (complete response (CR) or very good partial response (VGPR), vs no CR/VGPR).

Prognostic factors and treatment effect modifiers were assessed using both statistical analysis and clinical opinion.

From the literature review, imbalances between CALGB and Myeloma XI were identified in age (Myeloma XI had a slightly higher proportion of younger patients) and gender (with Myeloma XI having a slightly higher proportion of males).

The distribution of ISS scores at diagnosis differed for CALGB and Myeloma XI. Values for CALGB were retrospectively collected and therefore about a third of patients had missing values; it was also unclear whether the scores may have been recalculated from tested values obtained from medical records.

The studies differed by the proportion of people who received prior thalidomide or prior lenalidomide. This was a result of participants in Myeloma XI being randomised to one of three induction therapies (CRD, CTD or KCRD) whilst induction in CALGB preceded the start of the trial and was at the discretion of the treating physician (Also see Issue 1). Based on clinical opinion, the type of induction used before ASCT was deemed unlikely to have an impact on the efficacy of maintenance; newer inductions would also favour a deeper response; therefore the efficacy results for maintenance would be generalisable across populations who received different induction regimens. The adjustment for prior therapies would therefore be useful to improve the comparability between CALGB and Myeloma XI.

β 2-microglobulin at baseline also differed by studies. CALGB was stratified by β 2-microglobulin at randomisation, although according to clinical opinion, this clinical parameter should be assessed at diagnosis. Similarly to the ISS score, a large proportion of β 2-microglobulin values were missing. Perhaps for this reason, the distribution at baseline in CALGB are different to that in Myeloma XI and are imbalanced by arm in CALGB. Nevertheless, as β 2-microglobulin is a component of the ISS score, the value for β 2-microglobulin was not included in the set of covariates in the analyses.

Finally, the two studies differed by type of response after ASCT, with Myeloma XI having a much higher proportion of people who had complete response (CR) or very good partial response (VGPR). This may be explained by differences in the way ASCT response was assessed (bone marrow in CALGB and a mix of bone marrow and clinical assessment in Myeloma XI), or by differences in induction therapies used in the two studies.

Table 15 Descriptive characteristics of CALGB and Myeloma XI populations

		CALGB		Myeloma XI	
		Maintenance	Placebo	Maintenance	Observation
Age	< 60 years	56.7%	58.1%	████████	████████
Gender	M	52.4%	56.3%	████████	████████
ISS Scores	1	27.3% [39.1%]	38.4% [51.2%]	████████	████████
	2	25.1% [36.0%]	21.0% [27.9%]	████████	████████
	3	17.3% [24.8%]	15.7% [20.9%]	████████	████████
	Missing	30.3%	24.9%	████████	████████
Prior therapies	Prior len	34.6%	34.1%	████████	████████
	Prior Thal	44.2%	45.4%	████████	████████
B2-microglobulin	< 2.5mg/dL	24.2% [31.8%]	3.3% [41.1%]	████████	████████
	>=2.5mg/dL	51.9% [68.2%]	46.3% [58.9%]	████████	████████
	Missing	23.8%	21.4%	████████	████████
Response category post ASCT	CR or VGPR	55.4%	66.8%	████████	████████
	Not CR or VGPR	44.6%	33.2%	████████	████████
Key: CR, complete response; ISS, international staging system; Len, lenalidomide; VGPR, very good partial response; Thal, thalidomide; [% in brackets: % in non missing]					

The next step in the assessment was to assess which factors were found to be prognostic and which factors were likely to be treatment effect modifiers. Factors were assessed for each study independently. Forest plots were used to derive the hazard ratios by prognostic factors and treatment effect modifiers and are summarised in the Tables below (prognostic factors (Table 16) and treatment effect modifiers (Table 17 and 18)).

Table 16 Hazard ratios by group, prognostic factors for CALGB and Myeloma XI

Group (reference)	Prognostic factors, PFS (HR, pVal)		Prognostic factors, OS (HR, pVal)	
	CALGB		CALGB	
Age 60+ (<60)	0.89 (NS)		0.99 (NS)	
Sex: M (Female)	1.28 (NS)		1.36 (0.028)	
ISS 2 (ISS 1)	0.93 (NS)		0.88 (NS)	
ISS 3 (ISS 1)	1.25 (NS)		1.21 (NS)	
ISS missing (ISS1)	1.01 (NS)		1.1 (NS)	
Prior Len (No prior Len)	0.85 (NS)		0.86 (NS)	
Prior Thal (no prior Thal)	1.05 (NS)		1.08 (NS)	
β 2-microglobulin \geq 2.5 (<2.5)	1.16 (NS)		1.14 (NS)	
Post-ASCT response Not CR / VGPR (CR/VGPR)	1.28 (0.028)		1.15 (NS)	

Key: CR, complete response; ISS, international staging system; Len, lenalidomide; VGPR, very good partial response; Thal, thalidomide; [% in brackets: % in non-missing]

Table 17 Hazard ratios by group, treatment effect modifiers, CALGB and Myeloma XI, Progression Free Survival

Groups	TE modifiers, CALGB Overall HR = 0.53 (0.43-0.67)				TE modifiers, Myeloma XI Overall HR = 0.45 (0.36-0.56)			
	1	2	3	Interaction	1	2	3	Interaction
Age (60+ /<60)	0.43	0.69		P=0.022				
Gender (M/F)								
ISS (1,2,3)	0.54	0.33	0.88	P=0.031				
Prior Len (Y/N)	0.49	0.55		NS				
Prior Thal (Y/N)	0.53	0.47		NS				
β 2-microglobulin (\geq / <2.5)	0.53	0.49		NS				
Post-ASCT resp.	0.47	0.59		NS				

(CR/VGPR vs noCR/VGPR)								
Key: CR, complete response; ISS, international staging system; Len, lenalidomide; VGPR, very good partial response; Thal, thalidomide; Y, yes; N,no [% in brackets: % in non-missing]								

Table 18 Hazard ratios by group, treatment effect modifiers, CALGB and Myeloma XI, Overall Survival

Groups	TE modifiers, CALGB Overall HR = 0.52 (0.39-0.68)				TE modifiers, Myeloma XI = 0.61 (0.43-0.87)			Interaction
	1	2	3	Interaction	1	2	3	
Age (60+ /<60)	0.39	0.73		0.018				
Gender (M/F)	0.47	0.60		NS				
ISS (1,2,3)	0.52	0.39	0.64	NS				
Prior Len (Y/N)	0.34	0.62		NS				
Prior Thal (Y/N)	0.69	0.39		NS				
β2-microglobulin (>= / <2.5)	0.45	0.50	0.61	NS				
Post-ASCT resp. (CR/VGPR vs noCR/VGPR)	0.44	0.60		NS				
Key: CR, complete response; ISS, international staging system; Len, lenalidomide; VGPR, very good partial response; Thal, thalidomide; Y, yes; N,no [% in brackets: % in non-missing]								

The variables assessed that resulted in statistical difference in prognosis were gender for overall survival in CALGB and β2-microglobulin in Myeloma XI for both progression free survival and overall survival. (Table 2)

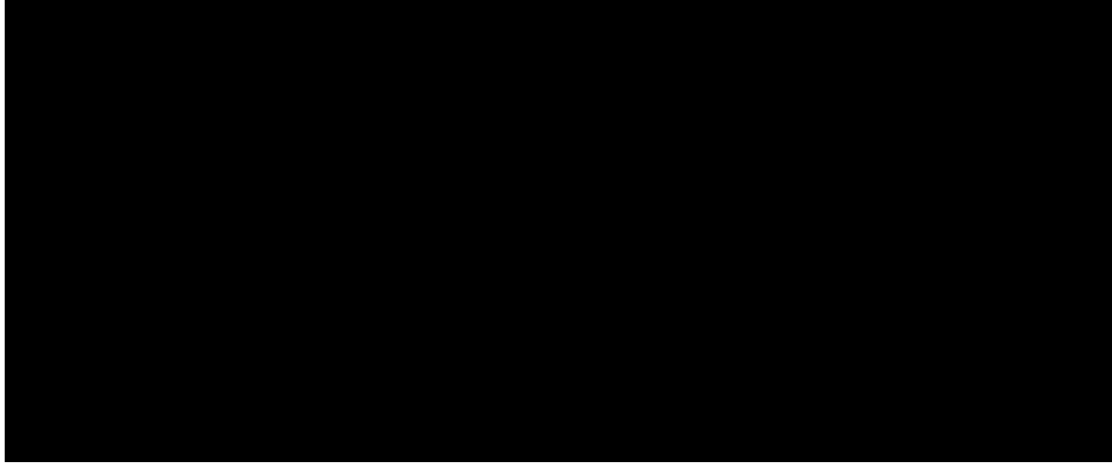
It was deemed clinically relevant to adjust for ISS score, as this is an important prognostic factor in multiple myeloma. This consideration was driven by the observed imbalance of ISS by arm in CALGB, although the imbalance may be the

result of the retrospective data collection for the score. Clinical opinion suggested that the use of β 2-microglobulin cut-off of 2.5 in CALGB was not appropriate as the clinically relevant cut-off for the parameter is 3.5. In addition, as β 2-microglobulin is a component of the ISS score (See Issue 1), an adjustment was not considered for β 2-microglobulin due to correlation with ISS stage.

Use of prior lenalidomide / prior thalidomide was matched for in the MAIC but not in the propensity scores weighted analysis, as patients in Myeloma XI received either prior thalidomide or lenalidomide by design, not both or neither. Use of prior lenalidomide and prior thalidomide are therefore collinear in Myeloma XI.

Differences were found in the distribution of CR and VGPR between studies. Importantly, ASCT response was also found to be a prognostic factor in CALGB for PFS, but not a prognostic factor or a treatment effect modifier in either CALGB or Myeloma XI.(Table 2-4). Based on clinical advice, response after ASCT was potentially a relevant variable for adjustment.

Considering the important difference between Myeloma XI and CALGB in the way response was assessed (bone marrow or clinical assessment) (see Issue 1), it is unknown whether such differences were of substance or due to misclassification. However, to explore the impact of ASCT response on OS, a more detailed assessment was conducted for this variable. The Kaplan-Meier curves were plotted separately for each subgroup in each study (Figure 1). The comparison of overall survival by CR / VGPR suggests that in CALGB these two groups have a similar overall survival over the duration of follow-up. In Myeloma XI, there may be a slight difference between people with VGPR and people with CR (expected to show the best response), with the CR subgroup performing slightly worse than both similar people in CALGB and people with VGPR in Myeloma XI. Although small numbers at risk remained in the two subgroups towards the end of follow-up for Myeloma XI, the comparison suggested that the two response groups ultimately had similar prognosis in CALGB, and given the counterintuitive OS in best responders in Myeloma XI, misclassification may be present in the latter.



For this reason, the impact of matching for response after ASCT may bias the comparison. Therefore, the base case scenarios were developed using matched CALGB data, excluding ASCT response from the set of matching covariates, to minimise the biases due to unknown confounding.

Two scenarios analyses were also developed including the adjustment for ASCT response, to describe the uncertainty deriving from the differences in the data collection and study design particularities in the two studies.

1.4 Assessment of effective sample sizes and weights distributions

To ascertain the impact of weighting on the estimation of the treatment effect, the estimated weights for CALGB were rescaled in the analysis to match the effective sample size (ESS) to prevent underestimating the uncertainty of parameter estimates,

$$rescaled\ w_i = \left(\frac{w_i}{\sum_{i=1}^n w_i} ESS \right)$$

where the ESS = number of independent non-weighted individuals that would be required to give an estimate with the same precision as the weighted sample estimate.(Greifer et al)

The distribution of weights was also explored using histograms to determine whether specific patients or groups of patients, based on covariate values, were over- or

under-represented in the analysis. Histogram-rescaled weights were calculated¹ to assess the proportion of cases which were under-weighted (<1) and that that were over-weighted (>1),

$$\text{Histogram - rescaled } w_i = \left(\frac{w_i}{\sum_{i=1}^n w_i} N \right)$$

2 Results

2.1 Propensity scores weights, matched-adjusted analysis

Table 19 below shows the baseline characteristics for the rebalanced CALGB population, in comparison with the original distribution (CALGB, non weighted) and with the same characteristics in Myeloma XI.

Table 19 Matched populations, baseline characteristics, Propensity scores weights (No response);

	CALGB, Non weighted		CALGB, Weighted		Myeloma XI	
	Lenalidomide	Placebo	Lenalidomide	Placebo	Lenalidomide	Placebo
N / ESS	231	229	95.07	105.07	621	411
Age < 60 years	56.71%	58.08%	41.51%	49.95%	48.95%	46.47%
Male	52.38%	56.33%	63.97%	63.38%	62.00%	62.29%
ISS I	27.27%	38.43%	39.76%	35.55%	36.88%	34.79%
ISS II	25.11%	20.96%	39.88%	43.00%	39.77%	44.04%
ISS III	17.32%	15.72%	19.56%	21.20%	22.54%	20.92%
Prior lenalidomide usage	34.63%	34.06%	71.69%	65.65%	71.01%	66.67%
Key: CR, complete response; ISS, international staging system; Len, lenalidomide; VGPR, very good partial response; Thal, thalidomide; Y, yes; N, no [% in brackets: % in non-missing]						

¹ These weights were not those used in the adjustments but were calculated only as support to the interpretation.

After matching, the characteristics are well balanced between treatment arms and between studies, although a small difference in the proportion of patients with age <60 years remains in CALGB.

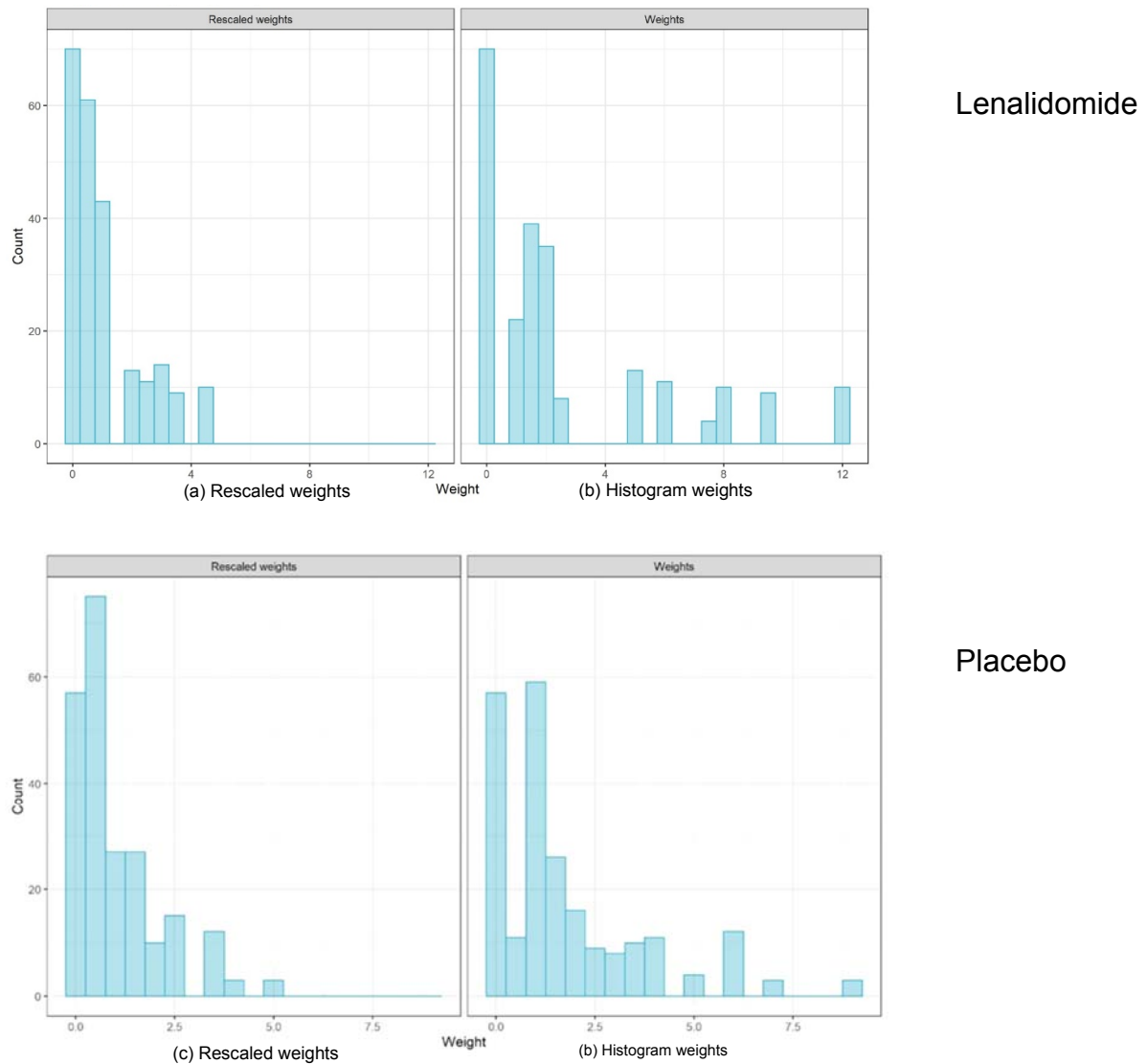
The ESS for the weighted CALGB analysis were approximately 95 for lenalidomide and 105 for placebo. This suggests that there is limited population overlap, as many patients in CALGB were not comparable to the Myeloma XI population.

The distribution of weights is presented in Figure 5 below. The figures show a relatively high proportion of participant were assigned a rescaled weight near 0; these were the patients in CALGB that were not represented in Myeloma XI. The proportion of people assigned a rescaled weight of 0 in placebo seems lower.

The PW match-adjusted Kaplan-Meier for overall survival is illustrated in Figure 5. The placebo arms are very similar to and overlapping the original unadjusted survival curve for CALGB, as well as to the Myeloma XI OS Kaplan-Meier for observation.

The estimated adjusted hazard ratio for death in CALGB resulting from the propensity scores weights analysis is HR = 0.41 (95% CI 0.27-0.63), slightly improved compared with the unadjusted hazard ratio reported in the original analyses, HR = 0.52 (95% CI 0.39-0.68) (see Issue 1).

Figure 4. Distribution of matching weights (propensity scores) used for CALGB (lenalidomide (a) and placebo (c)) and distribution of weights frequencies (lenalidomide (b) and placebo (d))



Key: ESS, effective sample size. **Notes:** For the rescaled weights, weight is equivalent to number of patients (e.g. a patient with a weight of 4 means that patient is equivalent to 4 patients).

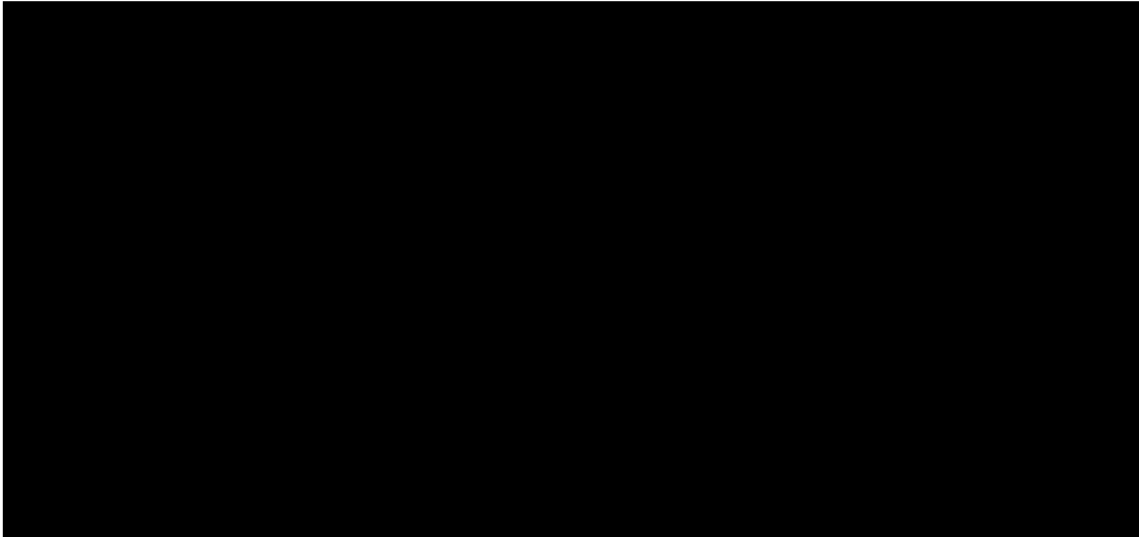


Table 20. OS Hazard ratio, lenalidomide maintenance compared with placebo, matched-adjusted analysis

Treatment (study)	ESS	Events	Median (95% CI)	HR (95% CI)
Placebo weighted (CALGB)	105	54	69.4 (59.4, 84.2)	Reference
Lenalidomide weighted (CALGB)	95	32	105.4 (101.8, NA)	0.41 (0.27, 0.63)
Key CI, confidence interval; ESS, effective sample size; HR, hazard ratio; ITT, intention-to-treat; LEN, lenalidomide; NA, not applicable; OBS, observation; PBO, placebo.				

2.2 MAIC

Table 7 below shows the baseline characteristics for the match-adjusted CALGB population using MAIC, in comparison with the original distribution and with the same baseline characteristics in Myeloma XI. After matching, the characteristics are well balanced between treatment arms and between studies.

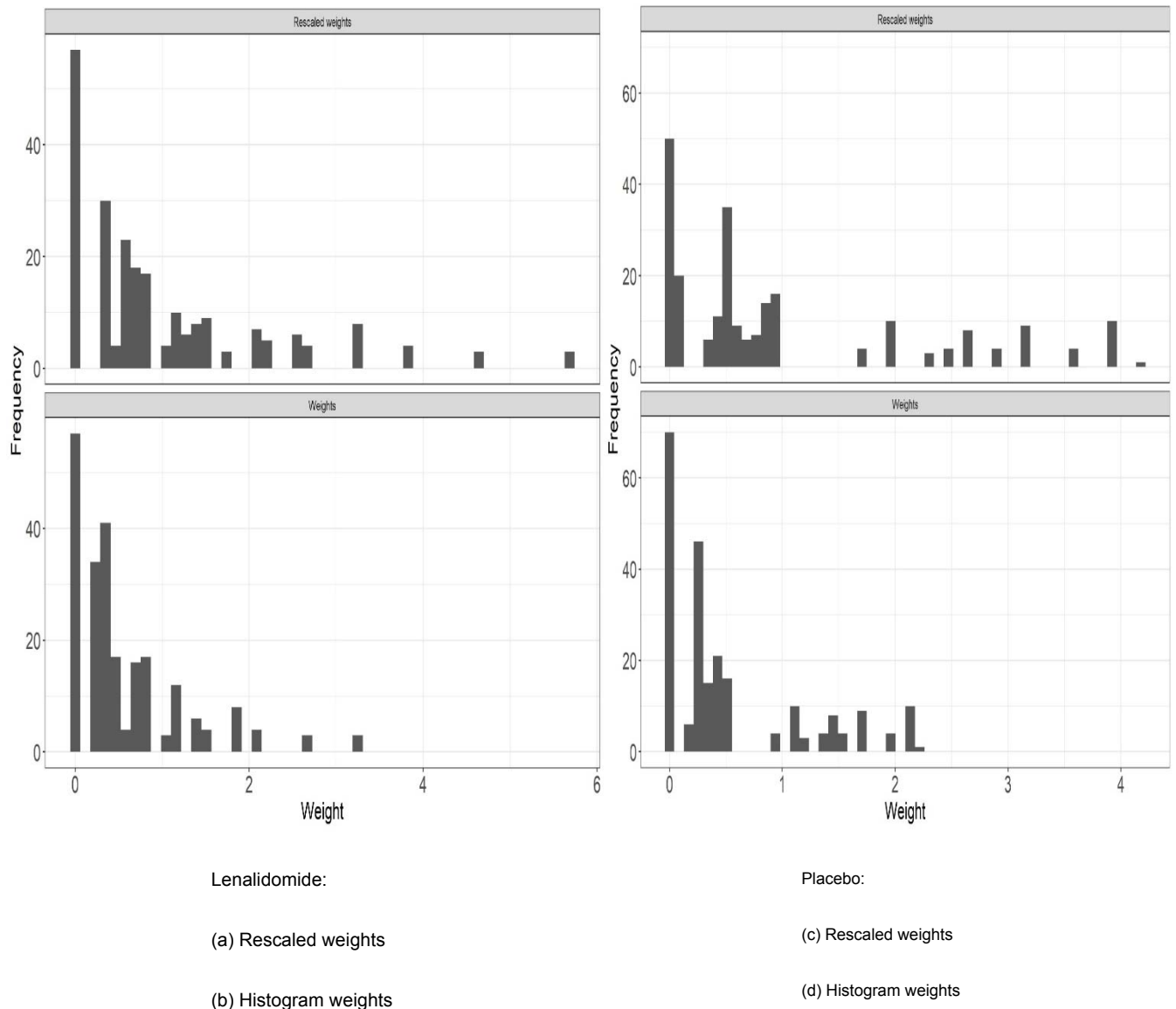
Table 21 Matched-adjusted baseline characteristics, MAIC (No response);

	CALGB		CALGB, matched adjusted		Myeloma XI	
	Lenalidomide	Placebo	Lenalidomide	Placebo	Lenalidomide	Placebo
N / ESS	231	229	98.5	98.8	621	411
Age < 60 years	56.71%	58.08%	49%	46%	48.95%	46.47%
Male	52.38%	56.33%	62%	62%	62.00%	62.29%
ISS I	27.27%	38.43%	37%	35%	36.88%	34.79%
ISS II	25.11%	20.96%	40%	44%	39.77%	44.04%
ISS III	17.32%	15.72%	23%	21%	22.54%	20.92%
Prior lenalidomide usage	34.63%	34.06%	71%	67%	71.01%	66.67%

Philippo et al (2018) report that an average of 80% reduction in effective sample size was reported in a number of other MAIC-adjusted analyses used in HTA. The CALGB ESS reduction are 58% approximately for both arms and are slightly less extensive than the reference 80%.

The distribution of weights is presented in Figure 7 below. The figures show a high proportion of participants were assigned a rescaled weight near 0; these were the patients in CALGB that were not represented in Myeloma XI. The proportion of people assigned a rescaled weight of 0 in placebo seems slightly lower.

Figure 5. Distribution of matching weights (MAIC) used for CALGB (lenalidomide (a) and placebo (c)) and distribution of weights frequencies (lenalidomide (b) and placebo (d))



The MAIC-adjusted Kaplan-Meier for overall survival is illustrated in

. The placebo arms are similar and overlapping the original unadjusted survival curve for CALGB, as well overlapping the Myeloma XI OS Kaplan-Meier for observation. The robustness of the placebo Kaplan-Meier to adjustments is similar to the results obtained with propensity scores.

The estimated adjusted hazard ratio for death in CALGB resulting from the MAIC analysis is HR = 0.40 (95% CI 0.26-0.62) (Table 22) , an improvement compared

with the unadjusted hazard ratio reported in the original analyses, HR = 0.52 (95% CI 0.39-0.68).

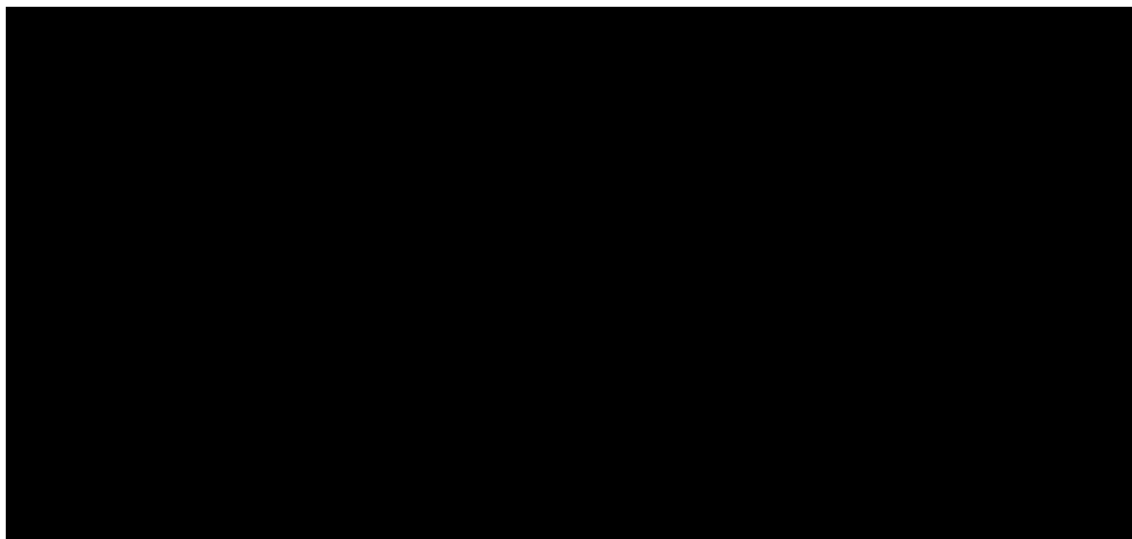


Table 22 OS Hazard ratio, lenalidomide maintenance compared with placebo, MAIC matched-adjusted analysis

Treatment (study)	ESS	Events	Median (95% CI)	HR (95% CI)
Placebo weighted (CALGB)	99	66	69.4 (59.4, 84.2)	Reference
Lenalidomide weighted (CALGB)	99	40	109.6 (101.8, NA)	0.40 (0.26, 0.62)
Key CI, confidence interval; ESS, effective sample size; HR, hazard ratio; ITT, intention-to-treat; LEN, lenalidomide; NA, not applicable; OBS, observation; PBO, placebo.				

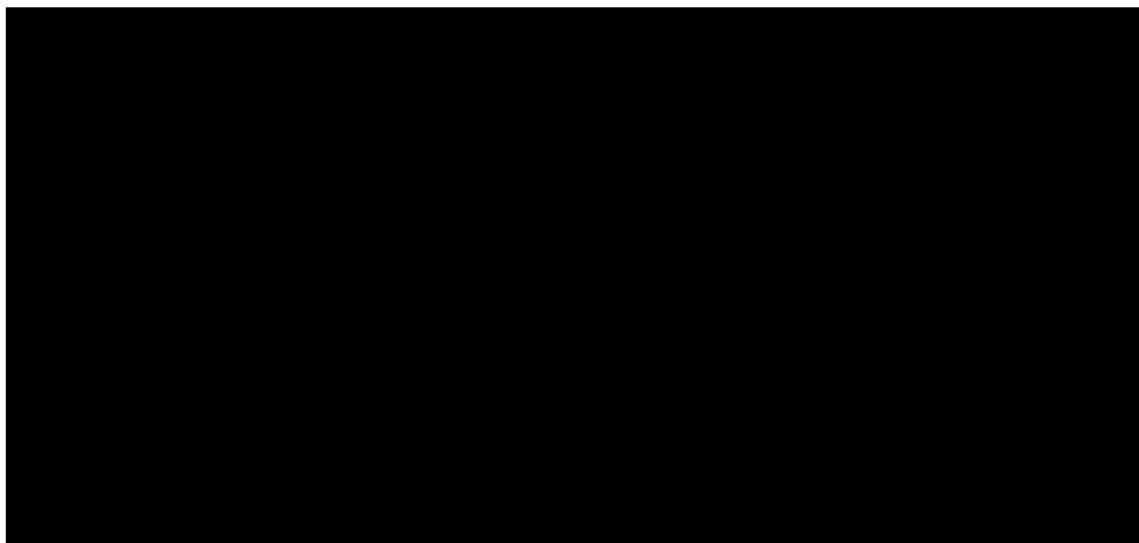
The same analyses were replicated for both propensity scores and MAIC-adjusted comparisons, including response after ASCT in the set of matching covariates. The results so obtained were similar, with very small differences between PW and MAIC and with respect to the matched-adjusted analyses with no ASCT response.

3 Propensity score match-adjusted pooled analysis

To address the comparison of CALGB and Myeloma, taking into account the difference in treatment dose (21/28 days compared with 28/28 days), a pooled model was also developed.

A Cox regression was used, based on data from both trials, to also adjust for the difference in the dose for lenalidomide maintenance 21/28 days dose vs 28/28 dose, compared with placebo. All factors were matched adjusted as in the propensity score weighted models and MAIS-adjusted models, therefore the methods used to identify matching variables will not be replicated here.

In addition to the population covariates, the model also included a term for treatment, a term for trial and a term for the interaction between treatment and trial.



The model was used to assess the impact of the difference between the two study doses for lenalidomide maintenance and to derive a matched-adjusted hazard ratio for Myeloma XI, to inform a scenario in the cost-effectiveness analysis.

3.1 Results

The pooled model confirmed the estimate of treatment effect generated with the PW-adjusted and MAIC-adjusted analyses (0.40, CI 0.26-0.62). A non statistically significant difference between the two doses was found (treatment x trial interaction 1.48, CI 0.85-2.59). (Table 23 and Figure 9)

Table 23. Results from the matched-adjusted, pooled model

Covariate	HR (95% CI)
Treatment (lenalidomide)	0.40 (0.26-0.62)
Trial (Myeloma XI)	0.97 (0.65-1.47)
Interaction treatment (lenalidomide) x trial (Myeloma XI)	1.48 (0.85-2.59)

4 Discussion

A reanalysis of the CALGB study was conducted to match the trial population to that of Myeloma XI to inform the comparison of the two studies and to obtain revised estimates of cost-effectiveness of lenalidomide used in maintenance after ASCT compared with observation.

The matched adjusted analyses show lenalidomide used in maintenance achieved a reduction in the hazard of death of 0.41 (0.27, 0.63) compared with placebo, when the CALGB population is matched to that of Myeloma XI. With respect to the unmatched analysis, matching CALGB data to Myeloma XI data shows that the efficacy of lenalidomide in maintenance is slightly improved.

Four adjusted analyses were conducted, using propensity score weighting and MAIC; and with or without adjustment for type of response after ASCT. These analyses were conducted to explore the impact of different types of matching on the results and also to assess the importance of adjustment biases potentially introduced by unknown confounders, derived from differences in data collection methods, clinical definitions and assessment of some important baseline characteristics in the two studies and study design features.

The analyses show that the comparison remains robust in all scenarios, with little differences in the estimated matched hazard ratio for death with lenalidomide maintenance compared with placebo, across all types of methods used for the adjustments.

The additional contribution of the adjusted analyses in the strengthening of the case for the two studies being conducted in comparable populations. Although some differences were identified in the distribution of some potential prognostic factors, the analysis showed that the matched-adjusted OS Kaplan-Meier for placebo were stable in all analyses and showed a high degree of concordance with both the unadjusted CALGB Kaplan-Meier and the Kaplan-Meier for the observation arm in

Myeloma XI. It is therefore unlikely that the prognosis with placebo may be impacted by factors other than those accounted for in the matching models. This would suggest that the populations in the two studies were not largely different with respect to baseline characteristics compared with each other, and that the untreated population in the two studies has a similar prognosis. The robustness of the KM curve with respect to matching adjustment also constitutes indirect support to the difference between the lenalidomide arm in CALGB and that in Myeloma XI could be largely attributed in lenalidomide dosing.

In terms of methodological uncertainty, the MAIC approach used in the analyses uses aggregate data from Myeloma XI. The limitations of MAIC are well understood: MAIC can be subject to residual confounding resulting from utilising summary data. Despite this, MAIC is often used in HTA and is one of the reference methods in the NICE DSU TSD 18.(Philippo et al 2018)

Propensity scores weighting, on the other hand, uses individual level patient data for both studies; however, propensity scores may also suffer from biases derived from potential unknown confounders.

Both MAIC and PW methods rely on there being no difference in treatment effects other than in the balancing factors; therefore, both adjustment methods share the limitation that the 21/28 vs 28/28 days difference in treatment dosing between the two studies cannot be matched.

Both analyses are associated with a reduction in the effective sample size in the range of 55%-60%. Small effective sample sizes are an indication that the weights are highly variable due to limited population overlap, and that the estimate may be unstable. Effective sample sizes were similar in the two approaches. Although this may introduce uncertainty, it is worth noting that the effective sample size reduction in the two trials is in line with that shown in similar analyses used in HTA (Phillippo et al, 2016).

In conclusion, for the reasons above, the comparison between Myeloma XI and CALGB was shown to be robust to heterogeneity across studies.

References

- Greifer, N., 2020. WeightIt: Weighting for Covariate Balance in Observational Studies (R package version 0.10.2). <https://CRAN.R-project.org/package=WeightIt>
- Li, F., Morgan, K.L. and Zaslavsky, A.M., 2018. Balancing covariates via propensity score weighting. *Journal of the American Statistical Association*, 113(521), pp.390-400.
- Phillippo, D.M., Ades, A.E., Dias, S., Palmer, S., Abrams, K.R., Welton, N.J. NICE DSU Technical Support Document 18: Methods for population-adjusted indirect comparisons in submission to NICE. 2016. Available from <http://www.nicedsu.org.uk>
- R core team, 2020, R: A Language and Environment for Statistical Computing. <https://www.R-project.org/>
- Rosenbaum, P.R., 1987. Model-based direct adjustment. *Journal of the American Statistical Association*, 82(398), pp.387-394.
- Signorovitch JE, Sikirica V, Erder MH, et al. Matching-adjusted indirect comparisons: a new tool for timely comparative effectiveness research. *Value in Health*. 2012; 15(6):940-7.
- Stuart EA, Cole SR, Bradshaw CP, Leaf PJ. The use of propensity scores to assess the generalizability of results from randomized trials. *J R Stat Soc Ser A Stat Soc*. 2001 Apr 1;174(2):369-386

Issue 3. Modelling the cost-effectiveness of lenalidomide in maintenance compared with observation using the adjusted CALGB curves

1 Introduction

The committee expressed a preference to see analysis based on extrapolation with CALGB, “*survival estimates in the economic model based on Myeloma XI data, with CALGB 100104 data (adjusted to reflect the Myeloma XI population as closely as possible, and conditional on the underlying survival of patients in Myeloma XI) used to help longer-term extrapolation, with all methods explained in detail (see section 3.9)*”. (ID475, ACD)

Multiple analyses of CALGB were performed to generate OS and PFS estimates matching the Myeloma XI population. These alternative analyses are detailed in Issue 2. above.

The matched-adjusted analyses were used to develop cost-effectiveness estimates. Specifically, the following scenarios were included in the cost-effectiveness analysis:

1. Myeloma XI until month 60, followed by OS and PFS estimates from the matched-adjusted CALGB analysis, using propensity scores (PW), to predict outcomes in a population with characteristics matched to those of the Myeloma XI population. All matched analyses were incorporated into the model, not including and including matching with response after ASCT.
2. Myeloma XI until month 60, followed by OS and PFS estimates from CALGB adjusted using the MAIC approach to predict outcomes in a population with characteristics matching those of the Myeloma XI population. Similarly, two scenarios with and without ASCT response were used.

3. Unadjusted estimates of OS and PFS from CALGB, to facilitate comparisons between the new analysis based on Myeloma XI data until month 60 and (non-matched) CALB data from month 60 onwards.

All the scenarios developed during the analysis were incorporated in the cost-effectiveness model and presented. This was to allow the transparent and complete comparison across scenarios and to characterize the impact of uncertainties in the cost-effectiveness analyses.

A most plausible scenario is also identified although the consideration of the most plausible scenario together with all scenarios jointly considered and the assessment of relevant and residual uncertainty is preferred.

2 Treatment effects

To use the CALGB outcomes beyond month 60 in both lenalidomide and observation arms would implicitly assume that the hazard ratio becomes that of CALGB at this point. In order to allow for alternative assumptions, the model was extended to include scenarios based on an alternative treatment effect, applied to the observation arm to predict outcomes for lenalidomide.

The following scenarios were generated:

- Use of treatment effects from Myeloma XI until month 60, and thereafter using the treatment effects for CALGB, following the Committee's request;
- Use of treatment effects from Myeloma XI at all time points;
- Use of treatment effects from the revised pooled analysis of CALGB and Myeloma XI at all time points, incorporating covariates to control for study, treatment, and study-by-treatment interaction.

Revised estimates of clinical effectiveness are reported in Issue 2 above.

Scenario analyses have also been added incorporating treatment waning from 10 and 15 years, with the former representing the follow-up from CALGB and for which evidence of constant treatment effect is available.

2.1 Choice of distributions for extrapolations in the cost-effectiveness model

CALGB-adjusted parametric models were derived from the matched-adjusted CALGB analyses to extrapolate overall survival and progression-free survival in the cost-effectiveness model.

The remainder of the analysis focusses on the impact of matched adjusted OS data. Although matched-adjusted PFS was included in the new version of the model, PFS remains a relatively uninfluential parameter in the cost-effectiveness. The adjusted PFS curves are not presented here; in addition, the same parametric distributions for PFS used in the previous versions of the model (Technical Engagement) were maintained, the gamma and the Weibull, as per original base case and ERG preferred distributions, respectively. The results of the PFS analyses are presented in the Appendices for completeness.

Step 1. The assumption of proportional hazards in the matched-adjusted CALGB analyses was reassessed based on the matched models. As in previous versions of the submission, the log -cumulative hazard plot and QQ plot for all CALGB adjusted models were obtained, for all adjusted models developed and presented in Issue 2 above.

Step 2. Measures of statistical fit (AIC, BIC) were used to assess the goodness of fit of the extrapolation to the data. We used this step as an aid to curve selection. As in previous scenarios submitted for Appraisal for Committee 1, measures of statistical fit provided useful information to identify the best ranking distributions but did not provide clear-cut indications to choose the best fitting distribution. For this reason, the statistical fit was used to *exclude* distributions that were clearly the worst performers in the rankings.

Step 3. Using both visual fit and clinical assessment, plausible distributions were identified in each scenario (MAIC and PW). For the purpose of visual inspection, the plot of the hazard ratio derived from the distributions was also considered. For all distributions which did not fail statistical fit, the impact on the ICER was assessed, together with predicted rates of OS over time for both placebo and lenalidomide maintenance.

As the MAIC and PW matched analyses showed that the placebo arm remained stable for all types of adjustments used, and the placebo arm in CALGB and observation arm in Myeloma XI remained consistently similar and overlapping regardless of which adjustments were conducted and which model was chosen (Issue 2), the best fitting curves to the placebo arm were chosen and used to generate a range of plausible scenarios. Therefore, a range of ICERs were generated and presented, to allow the Committee to consider the impact of distributional assumptions for OS on the ICER.

2.2 Statistical assumptions and statistical fit

The log-cumulative hazard plots for both PW adjusted and MAIC adjusted models are plotted in Figure 10 and Figure 11 respectively. The plots reconfirmed the appropriateness of using the proportional hazard / accelerate failure time (AFT) assumption for overall survival, as in the prior phases of the assessment, and regardless of the newly developed adjustment methods. For this reason, joint models only were estimated for extrapolating OS.

Figure 6 CALGB, overall survival adjusted with propensity scores weights, Log-cumulative hazard plot and QQ plot

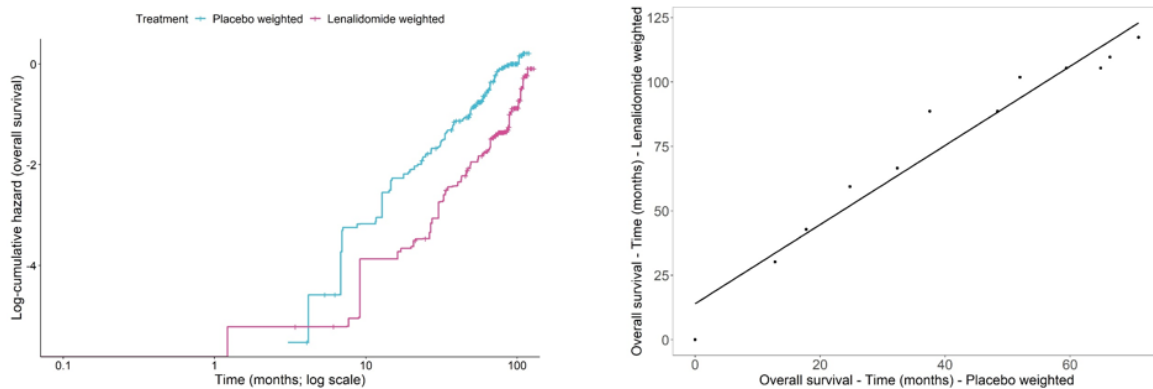
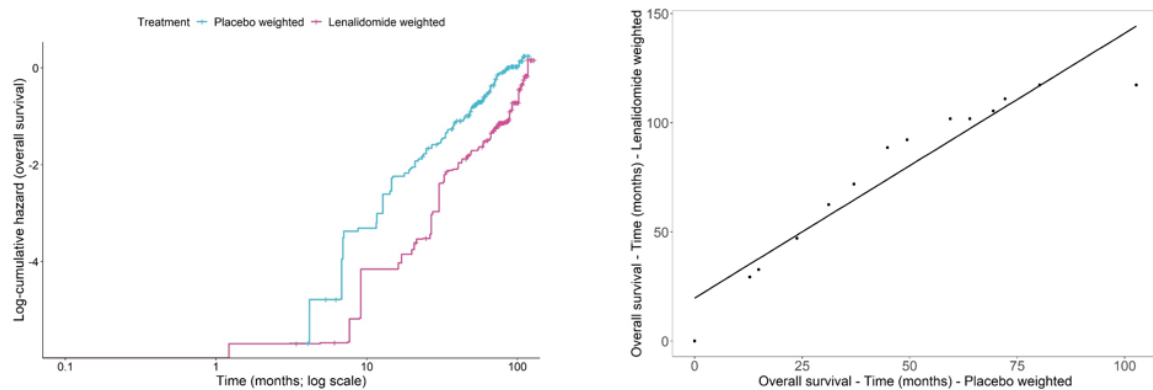


Figure 7 CALGB, overall survival MAIC- adjusted, Log-cumulative hazard plot and QQ plot



Statistical fit for all distributions was assessed using AIC and BIC (Figure 12 and Figure 13). The AIC / BIC rankings showed that the exponential distribution and the Gompertz distributions were consistently the worst fitting distributions, across all scenarios and all adjustment models for overall survival. The generalized gamma was consistently poorly ranked when using BIC but not when using AIC. Because the BIC penalizes distributions with more parameters compared with the AIC, this was not surprising. Therefore, the generalized gamma was retained for visual assessment.

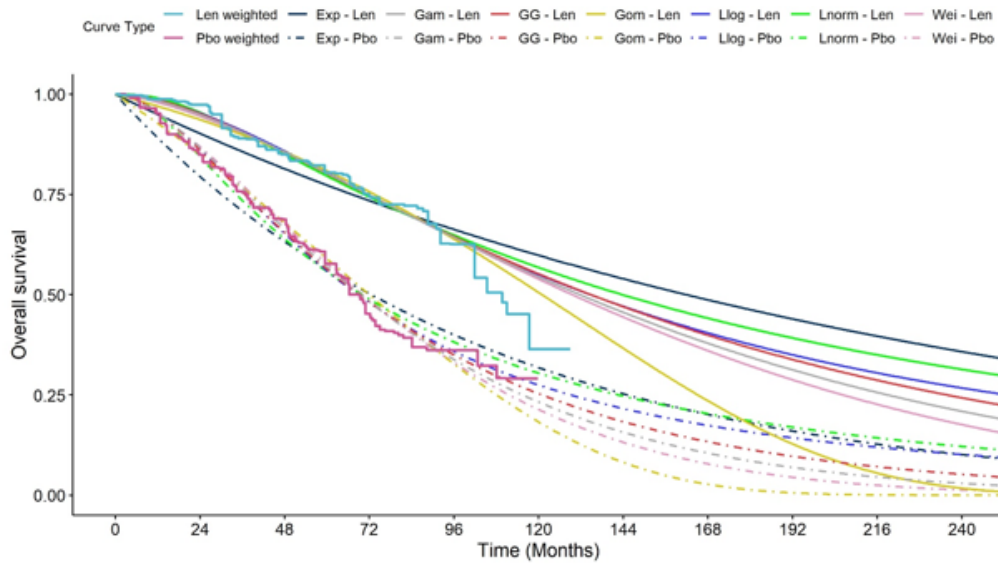
The gamma and log-logistic curves provided the best statistical fit in both MAIC-adjusted and PW adjusted models.

2.3 Visual fit

When considering visual assessment of the parametric distributions to CALGB data, the best fitting curves were the gamma distribution. This method was deemed most robust as the placebo arm showed stability in both MAIC-adjusted and PW-adjusted analyses and with respect to the Myeloma XI observation Kaplan-Meier. In addition, from visual assessment, joint models did not fit the lenalidomide arm particularly well. Visual inspection also showed the log-normal distribution not a suitable choice as the rates of overall survival projected for both lenalidomide and placebo seemed overly optimistic. (Figure 12 and Figure 13)

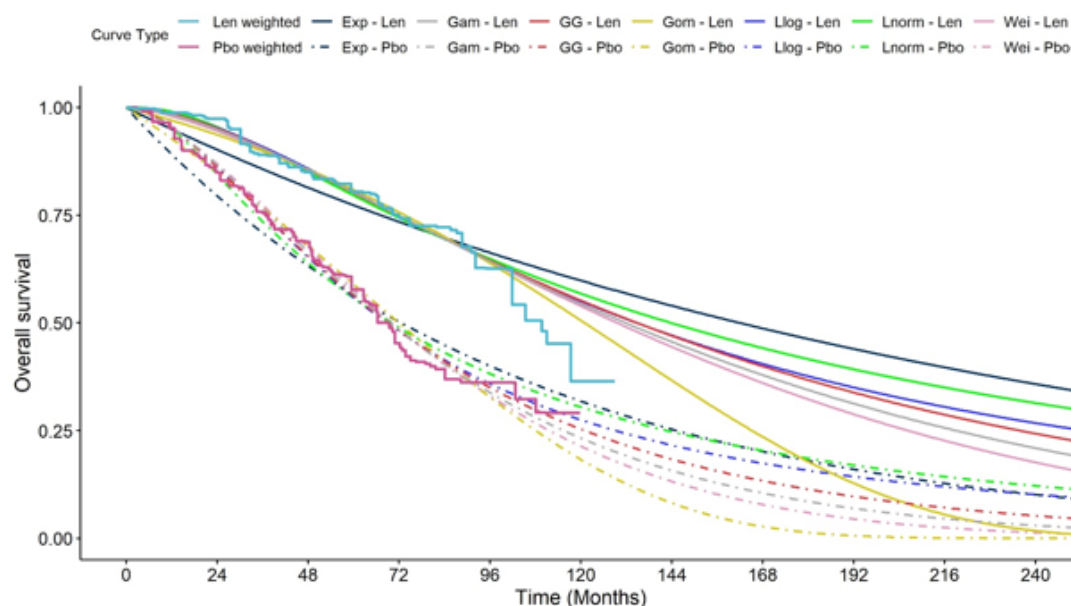
Figure 9)

Figure 8 Parametric distributions, PW weighted CALGB: joint models



Model	AIC	AIC rank	BIC	BIC rank
Exponential	2003.2	7	2011.5	7
Gamma	1985.9	1	1998.3	1
Generalized Gamma	1987.5	4	2004.1	5
Gompertz	1992.2	6	2004.6	6
Log-logistic	1985.9	2	1998.3	2
Log-normal	1988.0	5	2000.4	4
Weibull	1986.9	3	1999.3	3

Figure 9 Parametric distributions, MAIC-adjusted CALGB: joint models



Model	AIC	AIC rank	BIC	BIC rank
Exponential	2003.2	7	2011.5	7
Gamma	1985.9	1	1998.3	1
Generalized Gamma	1987.5	4	2004.1	5
Gompertz	1992.2	6	2004.6	6
Log-logistic	1985.9	2	1998.3	2
Log-normal	1988.0	5	2000.4	4
Weibull	1986.9	3	1999.3	3

The plot of the implicit hazard ratio derived from the parametric extrapolations was also used to assess the appropriateness of parametric curves. The HR plot shows the evolution of treatment effect with respect to follow-up duration (Figure 14). As follow-up increases, the relative efficacy of lenalidomide compared with placebo can stay constant, increase (as the hazard ratio approaches 1, the rate of deaths observed with lenalidomide approaches that observed with placebo, implying decreasing efficacy over time) or decrease (approaching 0). As expected, the plot showed constant hazard ratio for proportional hazard distributions (exponential, Weibull and Gompertz) and increasing hazard ratios (i.e. approaching 1) for generalised gamma, gamma, log-logistic and log-normal.

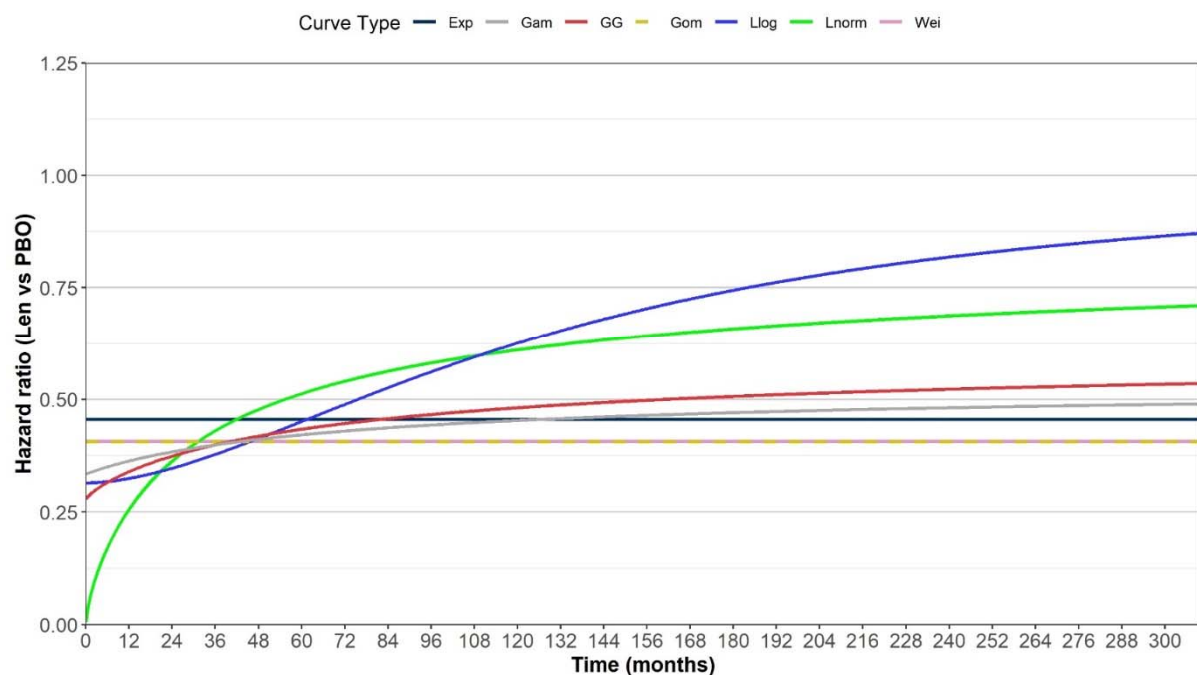
As the exponential distribution was excluded based on BIC and AIC, the discussion of the HR will not be taken forward.

The log-normal distributions provided overly optimistic predictions; however, they were retained in the analyses for completeness and also because the BIC and AIC ranks were fair.

Of the four remaining distributions, the Gompertz and Weibull distributions showed a similar (constant) implicit HR (both HR = 0.406). The HR in the log-logistic distribution showed the largest decrease in efficacy over the longer term, whilst the HR for the gamma remained in the range of that of the Weibull and Gompertz over the longer term, although slightly decreasing over time.

In conclusion, the HR with the Gompertz, Weibull, and gamma remained relatively unchanged over time, whilst the efficacy of maintenance implicitly assumed in the log-logistic distribution decreased over the course of the model.

Figure 10 Parametric distributions, plot of Hazard ratio from joint distributions, OS (PW matched analysis)



3 Revised cost-effectiveness results

The updated cost-effectiveness results are plotted in Table and Table, for all matched-adjusted CALGB models and using all possible options with regard to incorporation of the HR in the model.

For all scenarios, two alternative PFS distributions were used: Gamma distribution (company prior base case) and Weibull distribution (ERG preferred distribution). The

PFS distributions incorporated in the model are the MAIC-adjusted and PW-adjusted PFS distributions.

The analysis shows that for all scenarios using the Myeloma XI data until month 60 and adjusted CALGB data matching Myeloma XI from month 60, lenalidomide maintenance was cost-effective for most plausible OS distributions. (Table 23)

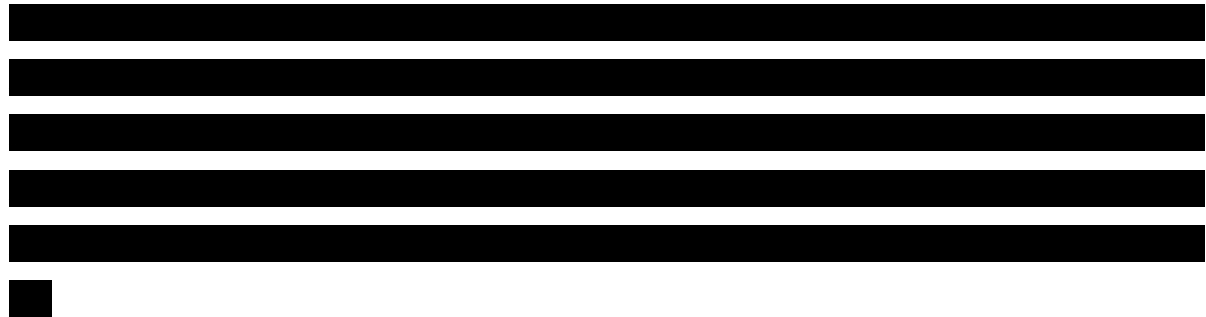


Table 24 below shows that for the scenarios requested by the Committee, and based on any of the adjusted CALGB data, lenalidomide maintenance remained robustly cost-effective.

Table 24: ICERs, lenalidomide maintenance compared with observation: Revised results and scenario analysis, Implicit HR (Myeloma XI m0-60; CALGB m60+): PFS distribution: Gamma (prior company base case) and ERG preferred PFS distribution (Weibull)

	OS distributions				
Extrapolation scenario	Weibull	Gamma	Log-normal	Log-logistic	Gompertz
PFS distribution: Gamma					
Myeloma XI+CALGB: PW without response	██████	██████	██████	██████	██████
Myeloma XI+CALGB: PW with response	██████	██████	██████	██████	██████
Myeloma XI+CALGB: MAIC without response	██████	██████	██████	██████	██████
Myeloma XI+CALGB: not matched	██████	██████	██████	██████	██████
PFS Distribution: Weibull (ERG preferred)					
Myeloma XI+CALGB: PW without response	██████	██████	██████	██████	██████
Myeloma XI+CALGB: PW with response	██████	██████	██████	██████	██████
Myeloma XI+CALGB: MAIC without response	██████	██████	██████	██████	██████
Myeloma XI+CALGB: not matched	██████	██████	██████	██████	██████

The ICER remained robust when using the hazard ratio for lenalidomide derived from the matched-adjusted pooled model with interactions (Table 25)

Table 25: ICERs, lenalidomide maintenance compared with observation: Revised scenario analysis, hazard ratio from `pooled model with treatment by trial interactions, PFS distribution: Gamma (prior company base case), Weibull ERG preferred)

Extrapolation scenario	OS distribution				
	Weibull	Gamma	Log-normal	Log-logistic	Gompertz
PFS distribution: Gamma					
Myeloma XI+CALGB: PW without response					n/c
Myeloma XI+CALGB: PW with response					n/c
Myeloma XI+CALGB: MAIC without response					n/c
Myeloma XI+CALGB: not matched					n/c
PFS Distribution: Weibull (ERG preferred)					
Myeloma XI+CALGB: PW without response					n/c
Myeloma XI+CALGB: PW with response					n/c
Myeloma XI+CALGB: MAIC without response					n/c
Myeloma XI+CALGB: unadjusted					n/c

n/c :did not converge

When considering the hazard ratio from Myeloma XI only and using the gamma or Weibull distributions, the ICER remained below the acceptability threshold for all types of adjusted models. When using the log-logistic or log-normal distributions, the ICER gave estimates slightly above the cost-effectiveness threshold, at around £32,000. (Table 26)

The only case when the ICER was above accepted threshold for cost-effectiveness was when using the Myeloma XI estimate for treatment effect and the Gompertz distribution, with an ICER in the range of £35,000 to £36,000 approximately. The model based on the Gompertz distribution failed to converge in the pooled model with interaction terms. (Table 26)

Very similar conclusions held when using the Weibull distribution for PFS (ERG preferred scenario).(Tables 24-25) The ICER remained largely under the cost-effectiveness threshold of £30,000 in all scenarios using the Myeloma XI data until 60 months and the CALGB extrapolations thereafter, regardless of the type of methods used for matching and adjustment used. For scenarios based on the Myeloma XI only hazard ratio, the ICER was under the cost-effectiveness threshold when using the Weibull and Gamma distributions, and slightly above when using the log-logistic and log-normal distributions. (Table 26)

Table 26: ICERs, lenalidomide maintenance compared with observation: Revised scenario analysis, Myeloma XI only hazard ratio. PFS distribution: Gamma (prior company base case), Weibull ERG preferred)

Extrapolation scenario	OS distribution				
	Weibull	Gamma	Log-normal	Log-logistic	Gompertz
PFS distribution: Gamma					
Myeloma XI+CALGB: PW without response	Green	Green	Yellow	Yellow	Orange
Myeloma XI+CALGB: PW with response	Green	Green	Yellow	Yellow	Orange
Myeloma XI+CALGB: MAIC without response	Green	Green	Yellow	Yellow	Orange
Myeloma XI+CALGB: unadjusted	Green	Green	Yellow	Yellow	Yellow
PFS distribution: Weibull					
Myeloma XI+CALGB: PW without response	Green	Green	Yellow	Yellow	Orange
Myeloma XI+CALGB: PW with response	Green	Green	Yellow	Yellow	Orange
Myeloma XI+CALGB: MAIC without response	Green	Green	Yellow	Yellow	Orange
Myeloma XI+CALGB: unadjusted	Green	Green	Yellow	Yellow	Yellow

E. Model Version used in response to ACD

All model changes were made to electronic model version Lenalidomide_CEA_ERG version_for_stakeholders (ACIC).xlsm sent by the ERG prior to Technical Engagement.

The new model version is aligned with the model file sent in in response to Technical Engagement.

1. Version ID475 Lenalidomide maintenance CEA_TE 20 07 2020 SUBMITTED.xlsm sent by Celgene in response to Technical Engagement (these changes were detailed in the submitted addendum)
2. Further changes subsequently made to align with minor issues identified in the Technical Report prior to the first committee meeting
3. Additional changes to align with the committee's preferred base-case, as detailed in the ACD.

This document details these changes and their impact on cost-effectiveness.

3.1 Alignment with model sent following Technical Engagement

Changes done to the version and corresponding Addendum submitted in response to Technical Engagement were:

- The re-weighting undertaken to omit CDF regimens from the UK Clinician Survey results was re-performed to exclude patients not receiving treatment and to only include the proportion of patients that were expected to receive any subsequent treatment before and after the removal of CDF drugs (as suggested on page 79 of the ERG Report, introduction of this difference was unintentional)
- Set cost of "other" treatments equivalent to CTD regimen
- Set MRU costs post-relapse same as pre-relapse and halve pre-relapse outpatient visits for observation
- An option has been incorporated in the economic model to choose extrapolation of outcomes based on the pooled CALGB and Myeloma XI analysis described in Addendum A in the response to Technical Engagement.

- New subsequent therapy scenario, from interviews with UK Clinicians
- Updated RDI as described in our response to the Technical Engagement.

3.2 Alignment with model used in the first committee meeting

Further changes in response to corrections made by the Evidence Review Group (ERG) identified in the Technical Report:

- Adding a scenario to capture the effect of 28/28 day dosing.
- Alternative scenarios with different distributions of subsequent therapies were included in the model to ascertain the impact of different assumptions.
- A correction was made to increase the number of oncologist appointments in the pre-progression observation arm. The ERG also assumed that there was no difference in medical resource use after progression and so a correction was made to align with the ERG assumptions.
- The cost of bortezomib was updated to be sourced from eMIT, and the PAS of 15% was removed. The cost of the 'other' therapies used 2nd line and beyond was set to be equivalent to the cost of the cyclophosphamide, thalidomide, and dexamethasone regimen.
- The Relative dose intensity (RDI) from Myeloma XI was included

Issue 4. Treatment waning effect

When considering the extrapolation of treatment effect beyond the observed follow-up in a clinical trial, a treatment effect 'waning' factor is often applied.

This is justified in some circumstances if it is reasonable to assume that the observed treatment efficacy will not be carried over into the future.

Common reasons for a decrease in treatment effect are:

1. Treatment interruptions or treatment capping: when a treatment is given for a finite period in time, for example because of stopping rules in reimbursement
2. Patients become non-compliant
3. Patients report progression or adverse events
4. Treatment effect is shown to decrease over time within the period for which data are available.

In some circumstances, the length of trial follow up is simply too short to lend itself to confident prediction of treatment efficacy over the longer term.

In the case of lenalidomide, most common reasons to suspect waning of treatment effect are not verified.

1. Maintenance with lenalidomide continues until disease progression or unacceptable toxicity. There is no capping rule for the duration of maintenance;
2. Non-compliance has been shown in Myeloma XI to be unlikely. Maintenance with lenalidomide is spaced with 'off treatment' intervals such that tolerability is actively pursued.
3. In the case of prolonged non-compliance, it is unlikely that a patient would remain in a prolonged state of pre-progression and therefore a progression would be observed in the not too distant future. The occurrence of progression would be captured in PFS.
4. The Myeloma XI and CALGB trials have a long follow-up; PFS data are mature in the case of CALGB; therefore, the prediction over the longer term is required only after the majority of PFS events have been accrued, with reduced uncertainty.

5. From the analysis of CALGB and Myeloma XI, there is robust evidence that the proportional hazard holds over the course of the trials for OS. The assumption of a decreasing treatment effect is in contradiction with the information obtained from the data.

For these reasons, and based on clinical opinion received, it is unlikely that a waning treatment effect would be plausible for maintenance with lenalidomide. Nevertheless, the results of applying a waning treatment effect in the model have been assessed and are reported here below.

Treatment waning was implemented by the ERG in electronic model version `Lenalidomide_CEA_ERG_version_for_stakeholders (ACIC).xism`. the same functionality has been used to generate the results reported here.

When the 'ERG_haz' option is enabled (ERG!R23), the user is able to specify a point at which the treatment effect for OS becomes 1.00; i.e. from this point the efficacy of lenalidomide and observation is assumed to be equivalent. This is achieved by setting the hazard in the lenalidomide arm to be the same as the observation arm.

The scenario tested the assumption that maintenance with lenalidomide loses all residual efficacy at 10 years, i.e. just after the end of follow-up in CALGB. This is the most extreme scenario as scenarios when treatment efficacy is lost after 10 years would be more favorable; therefore, it is the only scenario reported here.

3.3 Results

Table 27 shows that most plausible scenarios from the model are robust to the highest possible waning of treatment effect. The ICER is increased in all scenarios as a result of assuming that treatment effect wanes at 10 years.

Table 27 ICERs, lenalidomide maintenance compared with observation: Waning of treatment effect, PW matched-adjusted analysis, no ASCT matching, implicit hazard ratio until 10 years; hazard ratio=1 after 10 years

Extrapolation scenario	OS distribution				
	Weibull	Gamma	Log-normal	Log-logistic	Gompertz
PFS distribution: Gamma					
Myeloma XI+CALGB: PW without response	Green	Green	Green	Green	Yellow
Myeloma XI+CALGB: PW with response	Green	Green	Green	Green	Yellow
Myeloma XI+CALGB: MAIC without response	Green	Green	Green	Green	Yellow
Myeloma XI+CALGB: not matched	Yellow	Green	Yellow	Yellow	Orange
PFS Distribution: Weibull (ERG preferred)					
Myeloma XI+CALGB: PW without response	Green	Green	Green	Green	Yellow
Myeloma XI+CALGB: PW with response	Green	Green	Green	Green	Yellow
Myeloma XI+CALGB: MAIC without response	Green	Green	Green	Green	Yellow
Myeloma XI+CALGB: unadjusted	Yellow	Green	Yellow	Yellow	Orange

Issue 5. Distributions of subsequent therapies

The impact of subsequent therapies was explored in the model to assess the impact on cost-effectiveness of lenalidomide used in maintenance.

For these analyses, the following steps were undertaken:

- Subsequent therapies in Myeloma XI and CALGB were assessed and tabulated for comparison, to verify the material similarity between subsequent treatment used in CALGB and Myeloma XI. The aim of this comparison was to assess possible differences in study results because of subsequent therapies (performance bias)
- The preferred distributions of subsequent therapies in the cost-effectiveness model will be described, together with the rationale underpinning the proportions chosen
- Sensitivity analyses will be used to assess the impact of subsequent therapies on the cost-effectiveness.

3.4 Subsequent therapies: second-line anti-myeloma therapy (AMT) in CALGB and Myeloma XI

Table 27 illustrates the proportions of second line therapies received after first progression in CALGB and Myeloma XI.

Both studies had a substantial proportion of bortezomib-based second line therapy, 20% to 50% approximately, with Myeloma XI having a higher proportion. Lenalidomide was used in second line in both studies; the proportions were higher in CALGB (27% to 42%), with Myeloma XI presenting a proportion of approximately 10% in both lenalidomide maintenance and observation. This is likely a reflection of different reimbursement arrangements as in the UK, lenalidomide is mainly funded in third line.

At the time when CALGB was conducted, thalidomide, carfilzomib and pomalidomide were considered novel drugs. Overall, they were given to 20-30% of progressed

patients in CALGB, with a higher proportion for people who received placebo. In Myeloma XI, the overall proportions were 8% and 15%, with a higher rate in people who did not receive maintenance. The use of thalidomide in second line was similar in CALGB and Myeloma XI; carfilzomib and pomalidomide were more frequent in CALGB than in Myeloma XI. Similarly, pomalidomide in the UK is funded in later treatment lines than after first progression.

Importantly, the two studies also included second ASCT. The rates of second transplant were between 3-8% in CALGB and 1-3% in Myeloma XI. These rates are representative of a scenario where none of the (current) CDF drugs were available to clinicians. Second transplants rates were not different for maintenance and placebo; in CALGB, people who were switched to maintenance as part of study unblinding procedures had a second transplant rate of 8%.

The proportion of participants who did not receive post-progression anti-myeloma treatment was higher in Myeloma XI than in CALGB (15-17% compared with 4-8%).

Table 28 Second line anti-myeloma therapies in CALGB and Myeloma XI

	CALGB			Myeloma XI	
	Lenalidomide maintenance	Placebo	Placebo Switchers ^a	Lenalidomide maintenance	Observation
	N=233	N=229	N=76	N=621	N=411
Any 2 nd line AMT	████████	████████	████████	████████	████████
Bortezomib (+/- dex)	████████	████████	████████	████████	████████
Lenalidomide (+/- dex)	████████	████████	████████	████████	████████
Other novel drugs ^b or combinations, including:	████████	████████	████████	████████	████████
Thalidomide	████████	████████	████████	████████	████████
Pomalidomide	████████	████████	████████	████████	████████
Carfilzomib	████████	████████	████████	████████	████████
No novel drug	████████	████████	████████	████████	████████
Transplantation	████████	████████	████████	████████	████████
No second-line AMT	████████	████████	████████	████████	████████
Not progressed	████████	████████	████████	████████	████████
Died before 2 nd line	████████	████████	████████	████████	████████
Other	████████	████████	████████	████████	████████

Source: CALGB: CSR, 2015 cut-off; Myeloma XI: Data on file, 2017 cut-off

^a Participants switched to lenalidomide maintenance before progression, as part of study unblinding procedures

^b including thalidomide, carfilzomib, pomalidomide

^c Excluding lenalidomide received by subjects who had not progressed and were switched to maintenance after study unblinding

[percentages in bracket as a proportion of total number of patients treated with 2nd line therapies]

Key: dex=dexamethasone

As mentioned in Issue 1 above, CALGB and Myeloma XI were conducted at a time when new therapies were being launched. Specifically, within the period for the specific cut-offs for trial data used in this Appraisal, (CALGB: 2016, Myeloma XI: 2017), daratumumab had not received marketing authorisation. Some patients

received monoclonal antibodies in very small proportions as subsequent therapies (CALGB: 1.7-3% Myeloma XI: 1-2%), perhaps as a result of participation into clinical trials, however these proportions were very low and are unlikely to have a material effect on the overall survival of patients included in the data cuts that support this Appraisal.²

For this reason, subsequent therapies in the two studies should be considered a realistic depiction of subsequent therapies in a scenario without therapies that are currently on the Cancer Drug Fund.

The two studies were somehow different with regards to subsequent therapies however, the clinical opinion received was that these differences would not translate into material differences in overall survival between CALGB and Myeloma XI as neither study had any second line monoclonal antibodies use.

In summary:

- Both studies show that most participants who progressed received subsequent therapies
- The mix of subsequent therapies in the two studies is comparable, owing to the range of drugs available during study conduction
- Both studies included subsequent therapies based on real clinical choices in a context where monoclonal antibodies' availability was limited or not an option (pre-registration)

² For completion, a Table is provided with regulatory approval dates for currently approved antimyeloma drugs

Drug	US approval date	EU approval date
Bortezomib	2003	2004
Thalidomide	2006	2008
Lenalidomide	2005	2007
Pomalidomide	2013	2013
Carfilzomib	2012	2015
Panobinostat	2015	2015
Daratumumab	2015	2016
Ixazomib	2015	2016

- Both studies provide clear, concordant, real world estimates of subsequent rates of second ASCT, which are low and similar by arm.

3.5 Subsequent therapies scenarios in the cost-effectiveness model

Scenarios with subsequent therapies were run in the model to assess the impact of subsequent therapy lines on the ICER.

These scenarios closely reflect the data observed in Myeloma XI and in CALGB, and therefore are in line with clinical efficacy reflected in the data from these two studies. Furthermore, they closely reflect real clinical choices in a context before monoclonal antibodies had become available.

The scenarios also have been extensively validated by means of clinical opinion, and in addition, closely reflect the scenarios that, based on clinical opinion received, would still be observed in clinical practice in the absence of CDF drugs.

Table 29 shows the range of subsequent therapies that were tested.

The scenarios obtained modifying the proportion of second ASCT show that lenalidomide maintenance remains cost-effective at both 5% and 10% for this parameter. (Table 30) The scenario analysis in Table 30 use the highest of these values and shows that the ICER remains relatively unchanged for all distributional assumptions.

We also tested a scenario where lenalidomide in second line is given to 10% of people who would not receive maintenance (Table 31). This proportion was chosen to be closely representing the proportions seen in Myeloma XI and CALGB and in alignment with the observed clinical efficacy in the studies; this scenario faithfully reflects the current situation with respect to the use of lenalidomide in second line.

An increase in the rates of lenalidomide second line in observation improves the ICER substantially. This improvement may be substantial if, in the future, the proportion of lenalidomide in second line, should maintenance not be approved, may become much higher than rates from Myeloma XI and CALGB. To test this scenario, however, the associated change in the efficacy should be incorporated in the model; therefore the results should be viewed with caution.

Finally, the main scenarios tested include 5% of patients treated with carfilzomib. This number closely reflects real life clinical choices and the efficacy observed in CALGB and Myeloma XI.

In conclusion:

- The cost-effectiveness of lenalidomide in maintenance is robust to uncertainty regarding the most plausible rates of subsequent therapies after first progression; although the therapeutic pathway in MM is rapidly evolving, the most commonly used therapies at the time of CALGB and Myeloma XI have remained actual and therefore the distributions reflected in the model are valid and closely concordant with the efficacy data;
- All scenarios used are highly concordant with the exclusion of therapies currently available in the UK via the Cancer Drug fund;
- The cost-effectiveness is robust to variations in current, plausible values for second ASCT;
- The cost-effectiveness improves with higher rates of lenalidomide and carfilzomib used in second line; the values used in the model are highly concordant with clinical data and with the expected use of these therapies in clinical reality

Table 29 Subsequent therapies tested in the model as scenario analyses

	Treatment	Celgene's revised estimates		Scenario with higher second ASCT rates		Scenario with higher lenalidomide second line rates	
		Lenalidomide	Observation	Lenalidomide	Observation	Lenalidomide	Observation
After first relapse	Lenalidomide + dexamethasone	0%	0%	0%	0%	0%	10%
	Bortezomib + dexamethasone	60%	60%	60%	60%	60%	60%
	Carfilzomib + dexamethasone	0%	5%	0%	5%	0%	5%
	ASCT	5%	5%	10%	10%	5%	5%
	Other treatments	30%	25%	25%	20%	30%	15%
	No treatment	5%	5%	5%	5%	5%	5%
After second relapse	Lenalidomide + dexamethasone	0%	65%	0%	65%	0%	65%
	Bortezomib + dexamethasone	20%	10%	20%	10%	20%	10%
	Panobinostat + bortezomib + dexamethasone	20%	15%	20%	15%	20%	15%
	Pomalidomide	0%	0%	0%	0%	0%	0%
	Other treatments	50%	5%	50%	5%	50%	5%
	No treatment	10%	5%	10%	5%	10%	5%

Table 30 ICERs, lenalidomide maintenance compared with observation:
 subsequent therapies = Celgene’s revised estimates, second ASCT rates (10%)

Extrapolation scenario	OS distributions				
	Weibull	Gamma	Log-normal	Log-logistic	Gompertz
PFS distribution: Gamma					
Myeloma XI+CALGB: PW without response	████████	████████	████████	████████	████████
Myeloma XI+CALGB: PW with response	████████	████████	████████	████████	████████
Myeloma XI+CALGB: MAIC without response	████████	████████	████████	████████	████████
Myeloma XI+CALGB: not matched	████████	████████	████████	████████	████████
PFS Distribution: Weibull (ERG preferred)					
Myeloma XI+CALGB: PW without response	████████	████████	████████	████████	████████
Myeloma XI+CALGB: PW with response	████████	████████	████████	████████	████████
Myeloma XI+CALGB: MAIC without response	████████	████████	████████	████████	████████
Myeloma XI+CALGB: not matched	████████	████████	████████	████████	████████

Table 31 ICERs, subsequent therapies = Celgene’s revised estimates, higher
 second ASCT rates (10%) and lenalidomide 2nd line (10%)

Extrapolation scenario	OS distributions				
	Weibull	Gamma	Log-normal	Log-logistic	Gompertz
PFS distribution: Gamma					
Myeloma XI+CALGB: PW without response	████████	████████	████████	████████	████████
Myeloma XI+CALGB: PW with response	████████	████████	████████	████████	████████
Myeloma XI+CALGB: MAIC without response	████████	████████	████████	████████	████████
Myeloma XI+CALGB: not matched	████████	████████	████████	████████	████████
PFS Distribution: Weibull (ERG preferred)					
Myeloma XI+CALGB: PW without response	████████	████████	████████	████████	████████
Myeloma XI+CALGB: PW with response	████████	████████	████████	████████	████████
Myeloma XI+CALGB: MAIC without response	████████	████████	████████	████████	████████
Myeloma XI+CALGB: not matched	████████	████████	████████	████████	████████

Issue 6. Model scenario reflecting a 1-day to 28-day lenalidomide treatment regimen

A scenario analysis was conducted using the matched-adjusted CALGB data and the corresponding matched-adjusted time on treatment curve.

The analysis (Table 32) shows that for all plausible distributions and for all matched-adjusted scenarios, lenalidomide maintenance is cost-effective when given according to marketing authorisation, for 28/28 days administration cycles. The Gompertz distributions is associated with a slightly increased ICER.

The cost-effectiveness in the non matched-adjusted scenarios is less favorable, at ████████, although these scenarios are provided for reference only.

Table 32 ICERs, Lenalidomide maintenance compared with observation, using 28-days therapy (CALGB data); by matched-adjusted scenarios and extrapolated OS distributions

Extrapolation scenario	OS distributions				
	Weibull	Gamma	Log-normal	Log-logistic	Gompertz
PFS distribution: Gamma					
Myeloma XI+CALGB: PW without response	██████	██████	██████	██████	██████
Myeloma XI+CALGB: PW with response	██████	██████	██████	██████	██████
Myeloma XI+CALGB: MAIC without response	██████	██████	██████	██████	██████
Myeloma XI+CALGB: not matched	██████	██████	██████	██████	██████
PFS Distribution: Weibull (ERG preferred)					
Myeloma XI+CALGB: PW without response	██████	██████	██████	██████	██████
Myeloma XI+CALGB: PW with response	██████	██████	██████	██████	██████
Myeloma XI+CALGB: MAIC without response	██████	██████	██████	██████	██████
Myeloma XI+CALGB: not matched	██████	██████	██████	██████	██████

Issue 7. The company's justification for using the rank-preserving structural-failure time model over other methods to adjust for treatment switching in CALGB 100104

1 Introduction

The CALGB study was designed as a parallel, randomised double blind comparison of lenalidomide maintenance compared with placebo, in people with a diagnosis of multiple myeloma who received ASCT. At regular intervals, the Data Safety Monitoring Board (DSMB) reviewed study data as part of the protocol; subsequent to the review at 2 years, maintenance was found to be associated with significantly longer progression-free survival (PFS) with lenalidomide compared to placebo, and as a result, unblind treatment assignment and offer maintenance with lenalidomide to patients on placebo to patients who had not reached progression and had therefore remained eligible for maintenance. (Bertagnolli and Brigham and Women's Hospital, 2013).

1.1 Switched population

At the time of un-blinding on 17 December 2009, 110 (48%) subjects were ongoing in the placebo group; when switching was offered, 76 (69%) subjects switched over to lenalidomide maintenance prior to progressive disease (PD).

1.2 Feasibility assessment

A feasibility assessment was conducted using the 1st March 2015 data-cut to assess the relative merits of applying published methods to the data collected within the CALGB trial to adjust for the potential diluting effects introduced by patients switching over to lenalidomide from the placebo arm. The National Institute for Health and Care Excellence (NICE) guideline Adjusting survival time estimates in the presence of treatment switching (Latimer and Abrams, 2014) and the review paper of Watkins et al. (2013) identifies some methods that were considered in this assessment. The following methods were considered:

- Per protocol approach

CALGB CSR analyses utilised a per protocol approach. Although per protocol methods are not generally recommended because subject to informative censoring (Latimer and Abrams, 2014; Watkins et al., 2013), they may have merit when applied to the data of CALGB as the switching in CALGB was prompted by study design and not by patient's prognosis, therefore unrelated to survival outcome. As such, the potential for informative censoring being induced by patients being censored at switching may be limited.

The exploration of patient characteristics for people switched over to lenalidomide suggests that switchers had similar or improved disease characteristics at the time of diagnosis, pre-ASCT, and post-ASCT, compared to the overall placebo arm. The cohort was slightly younger cohort (median age, 55.0; 65.8% are < 60 years old) compared with the overall placebo arm (median age, 58.0; 58.1% are < 60 years old), with similar proportions of patients with ISS Stage III at diagnosis (15.8% versus 15.3%, respectively).

- Inverse probability of censoring weights (IPCW)

IPCW method artificially censors patients who switching, and outcome data after the point of switching are excluded from the analysis. In contrast to per-protocol analyses, the observations for the remaining patients are weighted to represent both themselves as well as the deleted data, while the data from the experimental arm are not weighted. The IPCW method was not recommended due to insufficient patient numbers; of those with the opportunity to switching (n=110), 76 did so. Therefore, data from patients who did not switching (N=34) would have to represent both themselves as well as the data of the patients who are censored due to switching.

- Rank-preserving structural failure time model (RPSFTM)/iterative parameter estimation algorithm (IPE)

It was recommended to further explore the RPSFT (IPE) method and the validity of the 'common treatment effect' assumption for PFS and OS; the method assumes that the experimental treatment effect is the same regardless of whether it is administered at randomisation or a later point.

- Two-stage methods

Use of the two-stage method was also not recommended. The method requires a common secondary baseline (a time-point at which all patients are at a similar stage of disease). The trigger for switching is not prognosis driven, and hence, there was no suitable secondary baseline.

1.3 Investigation of the common treatment effect assumption

The key assumption of the RPSFTM (IPE) methodology is the 'common treatment effect' assumption; this method assumes that the experimental treatment effect is the same (relative to first dose of treatment) regardless of whether it is administered at randomisation or at a later time point. Although this assumption is largely untestable, a number of analyses were performed to make an assessment of the suitability of the RPSFTM (IPE) methodology. If the common treatment effect assumption was to hold, it would be expected that the patients who receive lenalidomide as switching treatment receive sufficient treatment to benefit. Table 33 presents a summary of treatment duration for patients randomised directly to lenalidomide and for patients randomised to placebo, including the duration of lenalidomide treatment received following switching to lenalidomide. The median (min, max) treatment duration of those who received lenalidomide as switching treatment (N=76) is 20.6 months (-0.1, 61.3). Although this is numerically lower than those patients randomised directly to lenalidomide (median duration 25.4 months [0.3, 107.5]), this approximate 5-month difference is considered small in comparison to the variability inherent in this measure.

These analyses were conducted using the original data with cut off 1 March 2015, and later updated with the most recent data cut (cut-off 19 Oct 2016), leading to similar conclusions to those drawn from the original analyses. This report present the original feasibility analysis as this approach was retained; however, all analyses conducted in for the cost-effectiveness model and the Appraisal are based on the most recent datacut (October 2016); the updated results for the new datacut is illustrated here where pertinent.

Table 33: Summary of treatment duration (safety population) (CALGB, cut-off 1 March 2015)

	Lenalidomide (N = 224)	Placebo			
		Up to switching (N = 221)	Prior to switching (N = 76)	Lenalidomide after switching (N = 76)	Overall (N = 221)
Treatment duration (months)					
N	224	221	76	76	221
Mean	30.3	13.2	13.5	24.6	21.7
SD	25.43	9.64	8.93	21.17	19.22
Median	25.4	10.9	11.3	20.6	14.8
Min, max	0.3, 107.5	0.4, 50.7	2.9, 50.7	-0.1, 61.3	0.4, 85.8
Key: N, number; SD, standard deviation.					

As described above, as measured at randomisation, patients who switched over to lenalidomide were slightly younger than those randomised to the trial as a whole. However, when age at the time of initiating lenalidomide therapy was compared, this difference was reduced (Table 34). At the time of initiating lenalidomide therapy, the median age of the 76 subjects in the placebo arm who switched over to lenalidomide without PD was 56, and 60.5% were < 60 years old, compared to 58 (56.7%) in the group randomised directly to lenalidomide. Again, it is hard to determine what magnitude of difference in age is significant in this setting and in light of the variability present.

Table 34: Age at initiation of lenalidomide therapy - (CALGB, cut-off 1 March 2015)

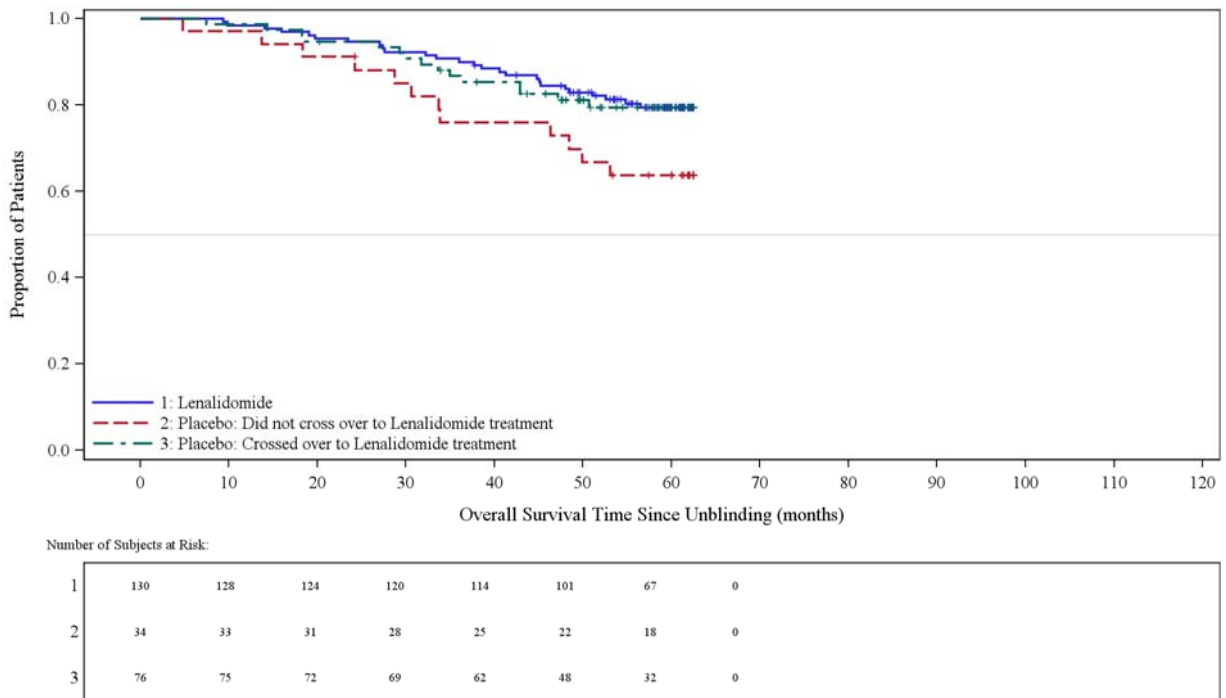
	Baseline characteristic	Lenalidomide N = 231	Placebo switching N = 76
Age at start of treatment	<60 yr	131 (56.7%)	46 (60.5%)
	>=60 y	100 (43.3%)	30 (39.5%)
Age at start of treatment	N	231	76
	Mean	57.3	56.6
	Std Dev	8.1	7.7
	0% Min	29.0	43.0
	25% Q1	51.0	51.0
	50% Median	58.0	56.0
	75% Q3	63.0	62.0
	100% Max	71.0	71.0

Figure 15 shows the Kaplan Meier curves of a 'landmark' analysis of OS; a comparison of overall survival, from the date of un-blinding, for the ITT population who were ongoing in the study at the time of un-blinding (17 December 2009). Survival, using the date of un-blinding as the origin, of patients originally randomised to lenalidomide is not dissimilar to that of placebo patients who switched over at the DSMB decision. Furthermore, a comparison of survival between placebo patients who did, and did not switching was performed to assess the plausibility of the common treatment effect assumption. The HR resulting from an analysis that compared the survival outcomes of all patients ongoing as of the 17 December 2009 was 0.53 [95% CI; 0.25, 1.13; N = 34 and N=76]. The HR that resulted from an analysis that compared the survival outcomes of all progression free patients ongoing as of the 17 December 2009 was 0.66 [95% CI; 0.29, 1.50; N = 30 and N = 76]. The treatment effect (HR) at the point of un-blinding is not of a different magnitude to that of the ITT analysis measured from the point of randomisation; 1st March 2015 cut-off, HR = 0.565 (95% CI:0.419, 0.761] and 17th December 2009 cut-off HR = 0.52 (95% CI: 0.26, 1.01).

Based upon these exploratory analyses, there is no data to suggest that the common treatment effect assumption of the RPSFT(IPE) methodology is severely violated. As

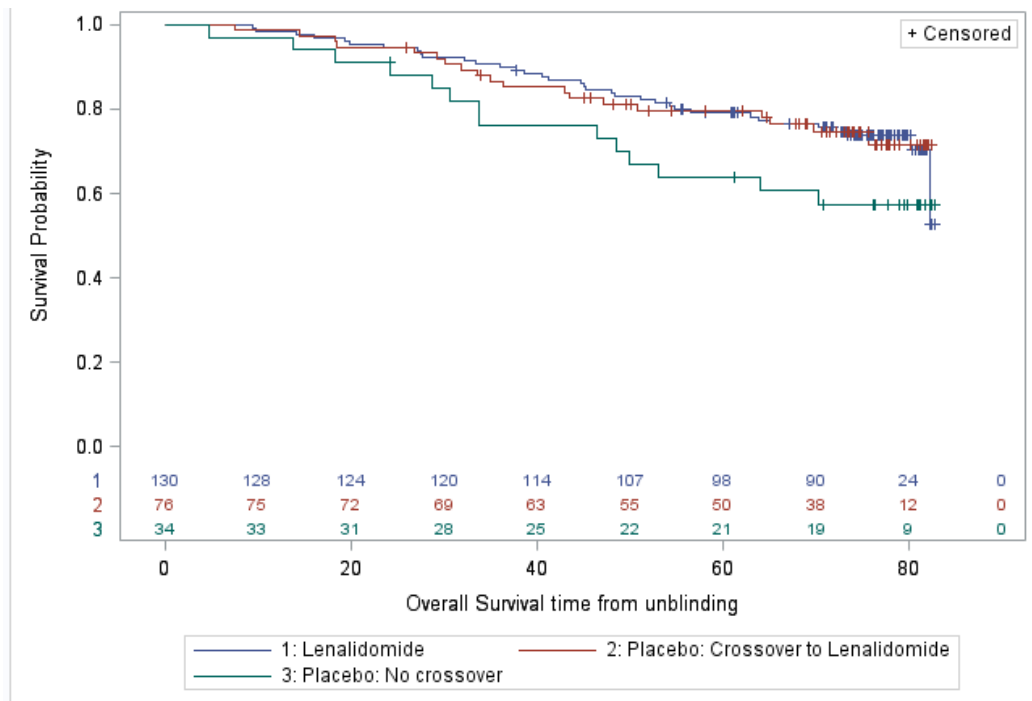
a result, the RPSFT method of adjusting for switching was applied to the data from the CALGB trial.

Figure 11: Kaplan Meier of Overall Survival since data of un-blinding – ITT population who were ongoing in the study at time of un-blinding (17 December 2009)



For the 19th October 2016 data cut off (Figure 16), the HR resulting from an analysis that compared the survival outcomes of all patients ongoing as of the 17 December 2009 was 0.57 [95% CI; 0.29,1.15: N = 34 and N=76] (Figure 12). The treatment effect (HR) at the point of un-blinding is not of a different magnitude to that of the ITT analysis measured from the point of randomisation at this cut-off date; HR = 0.61 (95% CI:0.47, 0.81] which is consistent with the conclusions of the original feasibility assessment.

Figure 12: Kaplan Meier of Overall Survival since data of un-blinding – ITT population who were ongoing in the study at time of un-blinding (19th October 2016 data cut-off)



1.4 Analysis populations

The analysis population was all randomised patients in the study, independent of whether or not they received study treatment (lenalidomide or placebo).

The RPSFTM/IPE methods assume that the experimental treatment effect is the same regardless of when it is administered; that is, the treatment effect of lenalidomide is the same for patients randomised directly to lenalidomide as for those who switching from placebo to lenalidomide. As such, only patients who received lenalidomide prior to investigator assessed progression (N=76) have had their survival times adjusted using the methodology described in Section 1.5.

Patients who receive lenalidomide post PD and/or as combination therapy do not have their survival times adjusted.

1.5 Analysis methods

Statistical analyses were conducted in R (R Core Team, 2016). A number of R packages were utilised, including the `boot` package (Canty and Ripley, 2016), the `survival` package (Therneau, 2015) and the `rpsftm` package (Bond, 2015).

2 Rank-preserving structural failure time model

The RPSFTM method was initially developed by Robins and Tsiatis (1991) to adjust for non-compliance in randomised trials. NICE Technical Support Document (TSD) 16 (Latimer and Abrams, 2014) and Watkins et al. (2013) describe the application of the RPSFTM method to estimate relative treatment effects in OS for randomised controlled trials where patients originally randomised to the control arm switch to experimental treatment.

2.1 Estimation of counterfactual survival times ('on-treatment' approach)

The RPSFTM uses a causal model for counterfactual survival. The counterfactual survival times, U_i , were defined as those that would have been observed if no treatment had been given.

The model splits the survival time, T_i , for each patient (i) into two: the time when the patient was 'on' lenalidomide treatment was assigned to T_{L_i} , and all other time was assigned to T_{P_i} .

Patients randomised to placebo who subsequently switching to lenalidomide prior to PD had the time that they receive lenalidomide assigned to T_L ³. All other time prior to and post lenalidomide administration was assigned to T_P . Patients randomised to placebo who did not switching had all their survival time assigned to T_P . Patients randomised to lenalidomide had all the time they receive lenalidomide assigned to T_L . All other time prior to and post lenalidomide administration was assigned to T_P . This approach is described as the 'on-treatment' group approach. An additional analysis was conducted that assumed the effect of lenalidomide is retained until death; this is labelled the 'treatment group' approach and is described in Section 2.4.

The counterfactual survival time (U_i) for each patient (i) was defined as:

³ Patients who receive lenalidomide post PD or as combination therapy will not have time assigned to T_L .

Equation 1

$$U_i = T_{P_i} + e^{\psi} T_{L_i}$$

The value of ψ was estimated using g-estimation. For each possible value of ψ , Equation 1 was used to estimate U_i , and the value of ψ was identified by determining which value of ψ results in counterfactual survival times that are equalised across randomised groups. To identify this value, a log-rank test (unadjusted) was used; this tests the hypothesis that the baseline survival curves are identical in location parameter in the two treatment groups. The selected value of ψ was the value for which the test statistic (z) equals zero.

To assess whether there is a unique solution, the test statistic (z) was plotted against the corresponding values of ψ . Kaplan–Meier plots of counterfactual survival times were plotted for each value of ψ to assess the suitability of each solution; for example, a value of ψ might have been considered unsuitable if these Kaplan–Meier curves crossed.

The assumption of non-informative censoring is required in time-to-event analysis. However, censored survival times on the counterfactual scale are likely to depend on the underlying prognosis through their dependence on the amount of treatment received. Unbiased estimation requires the dependence of censoring time on treatment to be broken. To attempt to remove some of the potential bias of this dependence, the re-censoring algorithm of White et al. (1999) was applied to the calculation of ψ . The re-censoring algorithm utilises the maximum follow-up time for a patient (data cut-off – randomisation), denoted as C_i . For each value of ψ , a patient's counterfactual survival time, $U_i(\psi)$, was compared to $D_i(\psi) = \min(C_i, C_i \exp(\psi))$. If $D_i(\psi) < U_i(\psi)$, $U_i(\psi)$ was replaced by $D_i(\psi)$, and the event time was censored. Re-censoring for the calculation of ψ was applied in both randomised arms.

2.2 Adjusted survival analysis

The estimated value of ψ was used to derive counterfactual survival times for patients who switched over from placebo to lenalidomide. Survival times for the remaining patients remained unaltered. Following estimation of adjusted survival

times, the statistical analysis techniques utilised for primary analysis of OS were performed to obtain a point estimate of the HR, an un-stratified log-rank test and an unadjusted Cox-proportional hazards model, and standard summaries of survival data (KM plots, median survival, etc.).

To obtain a CI for the HR, the observed dataset was re-sampled (with replacement), and for each sample, the RPSFTM was applied, ψ was estimated, and an HR was derived. Using this distribution of HRs, the lower and upper, 2.5th and 97.5th, quantiles were derived and used to represent the 95% CI. White et al. (1999) states that, as the RPSFTM method is randomisation-respecting, this preserves the intention to treat (ITT) p-value, and therefore, no p-value was generated for this procedure.

2.3 Covariate adjustment

As described in above, there were some differences in the baseline characteristics of patients who switched over from placebo to lenalidomide compared to those who did not and those originally randomised to lenalidomide. Therefore, a form of covariate adjusted RPSFTM was implemented in order to explore the effect these differences may have on the comparative efficacy of lenalidomide versus placebo. The covariates of prior lenalidomide treatment, prior thalidomide treatment and β 2-microglobulin were used to adjust the model.

2.4 Estimation of counterfactual survival times ('treatment group' approach)

As for the 'on treatment' approach, the 'treatment group' approach uses the same causal model for counterfactual survival (Equation 1). That is, the survival time, T_i , for each patient (i) was partitioned into two: time assigned to lenalidomide (T_{L_i}) and time assigned to placebo (T_{P_i}). However, for this model, it was assumed that the residual effect of treatment with lenalidomide is retained until death. As such, patients randomised directly to lenalidomide had all time until death (or censoring) assigned to T_L , irrespective of how long they remained on lenalidomide or whether they switched treatment. Patients randomised to placebo who did not switching had all their survival time assigned to T_P . Patients randomised to placebo who subsequently switching to lenalidomide prior to PD had all time until death (or

censoring) following first dose of lenalidomide assigned to T_L . All other time prior to lenalidomide administration was assigned to T_P .

Following construction of counterfactual survival (U_i) for each patient, the procedures outlined as part of the methods were performed. However, re-censoring for the calculation of ψ was only applied to the placebo arm. The 'treatment group' model assumes that all lenalidomide patients have received treatment up until event or censoring. Therefore, on the counterfactual scale, censored survival times are no longer likely to depend on the underlying prognosis through their dependence on the amount of treatment received.

2.5 Iterative parameter estimation algorithm

The IPE method was developed by Branson and Whitehead (2002). The IPE method belongs to the class of accelerated failure time (AFT) models, and it is an extension of the RPSFTM method using a parametric likelihood approach.

2.6 Estimation of counterfactual survival times

The same model for counterfactual survival time (U_i) was used as for RPSFTM (Equation 1).

Initially, ψ was estimated from a parametric accelerated failure time model fitted to the original, unadjusted, survival data. The parametric accelerated failure time model included a factor for treatment from which ψ was estimated. Using Equation 1, the initial estimate of ψ was used to estimate U_i for patients who switching. These were combined with the observed survival times for all other patients, and a parametric accelerated failure time model with a factor for treatment was fitted. The estimate of ψ was again used to estimate U_i for patients who switch treatment and combined with the observed survival times for all other patients and a parametric accelerated failure time model with a factor for treatment fitted. This was repeated until the value of ψ converges.

The choice of parametric accelerated failure time model chosen was based upon the lenalidomide arm. The following 3 parametric distributions were considered:

Weibull (Exponential⁴)

Log-normal

Log-logistic

The best fitting model was selected based upon Akaike information criterion (AIC) and Bayesian information criterion (BIC) statistics and visual inspection of the parametric curves to the Kaplan–Meier graph.

3 *Adjusted survival analysis*

As for RPSFTM (Section 2.2), the estimated value of ψ was used to derive counterfactual survival times for patients who switched from placebo to lenalidomide, combined with the survival times for the remaining patients and the analysis described in Section 2.2 performed.

If the value of ψ and resulting treatment effect estimators (HR, median survival, etc.) were not consistent between the RPSFTM and IPE methodology, this may have indicated that the value of ψ estimated from one or both of the methods was not plausible. An implausible solution may arise when a mathematical solution was found using G-estimation (RPSFTM) or through the iterative procedure (IPE) but one or more of the assumptions underlying the methods were not supported, and/or if adjusted treatment effect estimators (HR, median, etc.) were not supported by other data sources (e.g. magnitude and direction of unadjusted OS, PFS, response rate). This was explored using a number of techniques. For example, Kaplan–Meier plots of counterfactual survival times were plotted for each solution, ψ . The g-estimation technique utilised by the RPSFTM methodology aims to find a value of ψ that equalises counterfactual survival times across the randomised treatment arms. If survival curves were seen to cross despite the test statistic being zero, a mathematical solution to the problem of equalising counterfactual survival would have been found but the distribution of counterfactual survival times would not be equivalent. In such a case, the corresponding value of ψ would not be considered as a legitimate solution; the RPSFTM assumes that the only difference between

⁴ Exponential is a special case of the Weibull distribution where the shape parameter equals 1.

randomised groups is the treatment received. As such, the value of ψ would not be used to create adjusted treatment effect estimators (HR, median, etc.). Similarly, it would be expected that the value of the treatment effects estimator (HR, median, etc.) shows a treatment effect in the same direction as the primary analysis already conducted ($HR < 1$) and of the same order of magnitude. In this case, a mathematical solution for ψ may have been found, which is not plausible. In such cases, this report will provide a full description of all possible solutions and the merits and limitations of each.

4 Results

Given the immunological method of action of lenalidomide, and potential impact of prior treatment on subsequent therapy choices, the ‘on treatment’ method (where no residual effect of lenalidomide is retained once treatment is stopped) was considered to be an unlikely scenario, and the ‘treatment group’ method was selected as the base case method for these analyses. In this section, the results from RPSFTM and IPE are presented, firstly for the ‘treatment group’ method and secondly for the ‘on treatment’ method. Re-censoring for ψ is employed for all RPSFTM processes, but this was not employed for the IPE method, as it has been proposed that recensoring for IPE be limited to patients who switch on the control arm, and only if survival times are projected beyond the end of the study; that is, only if the experimental treatment has a detrimental effect compared to control (Branson and Whitehead, 2002), which is not the case in the CALGB trial.

In the following analyses, in line with the CALGB CSR, the per protocol method of censoring at switching for PFS was defined using the date of last adequate assessment before switching, and for OS, the exact date of switching was used.

5 Base case: ‘Treatment group’ approach

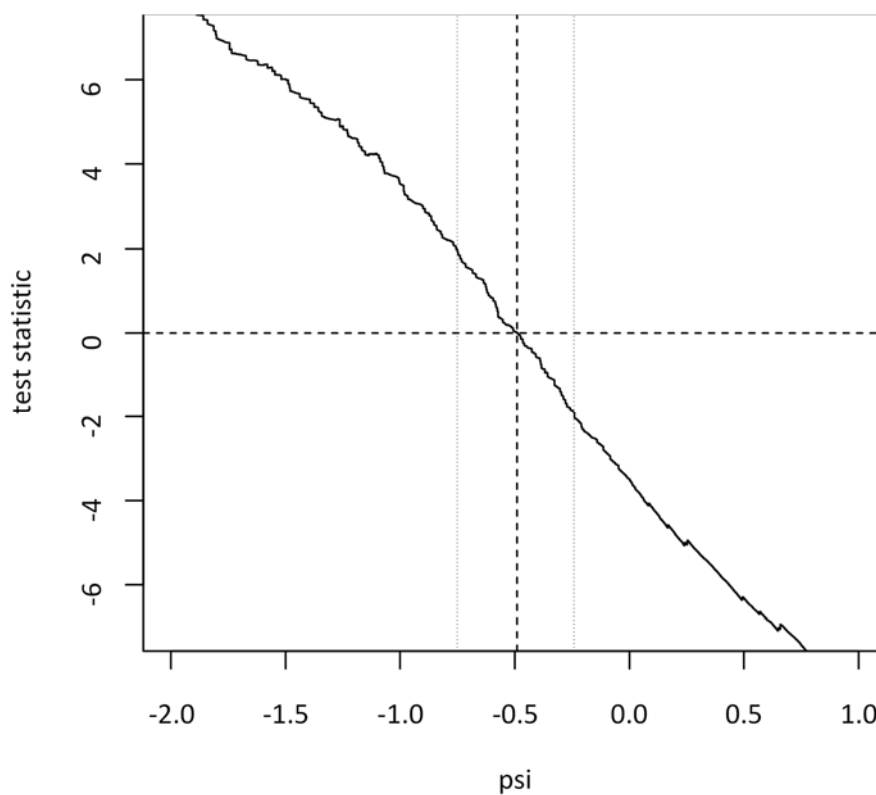
4.1 Overall survival

RPSFTM

Figure 13 shows ψ plotted against the corresponding log-rank test statistic; the plot is approximately monotonic and has only one solution for ψ : -0.490.

Figure 14 shows a Kaplan–Meier plot and a quantile-quantile (QQ) plot of the counterfactual survival times U_i defined in Equation 1 and calculated for the chosen value of ψ . These plots can be used to assess the similarity in distribution of two arms, which is a requirement for the RPSFTM method. The two arms in the KM plot look similar, and the QQ plot follows the identity line $y = x$ moderately well, as required. Based on these plots, the requirements of the RPSFTM method appears to be satisfied.

Figure 13: Log-rank test statistics plotted against the corresponding values of ψ – ‘treatment group’ method – overall survival (CALGB, cut-off 19 Oct 2016)

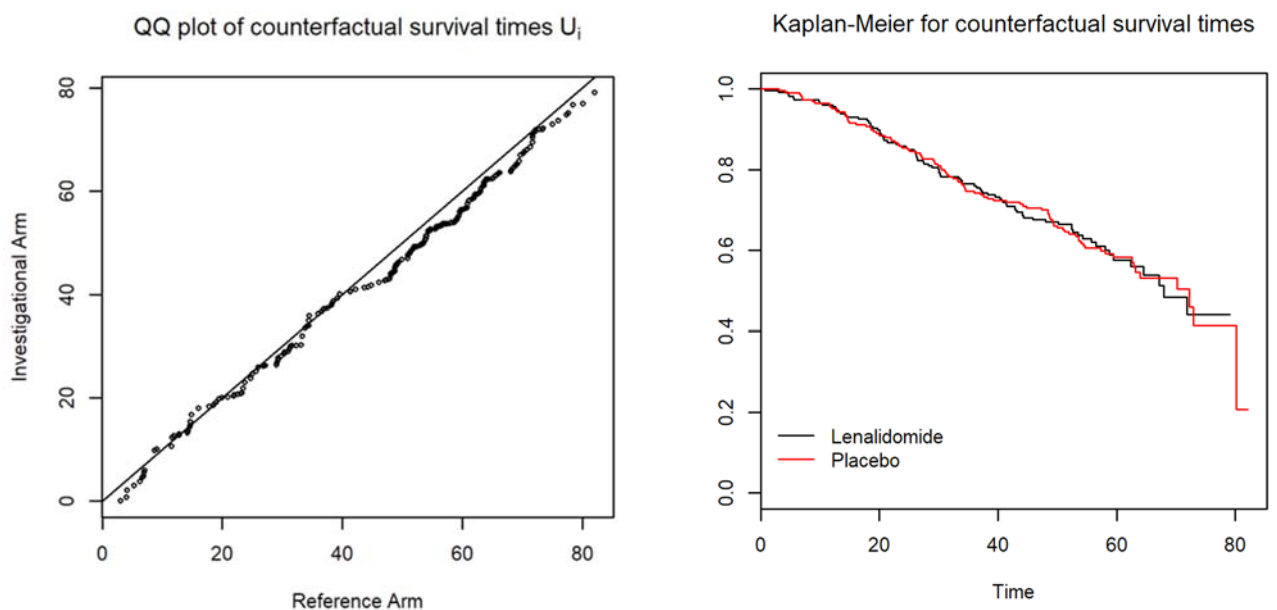


Once a solution for ψ was identified, this was used to adjust the survival times of patients who switched over from placebo to lenalidomide using Equation 1; no other patients’ survival times were adjusted.

Figure 15 presents KM curves of RPSFTM adjusted OS for placebo in addition to unadjusted OS and OS censored at switching. The RPSFTM adjusted OS placebo arm lies between unadjusted OS and OS censored at switching. As described in

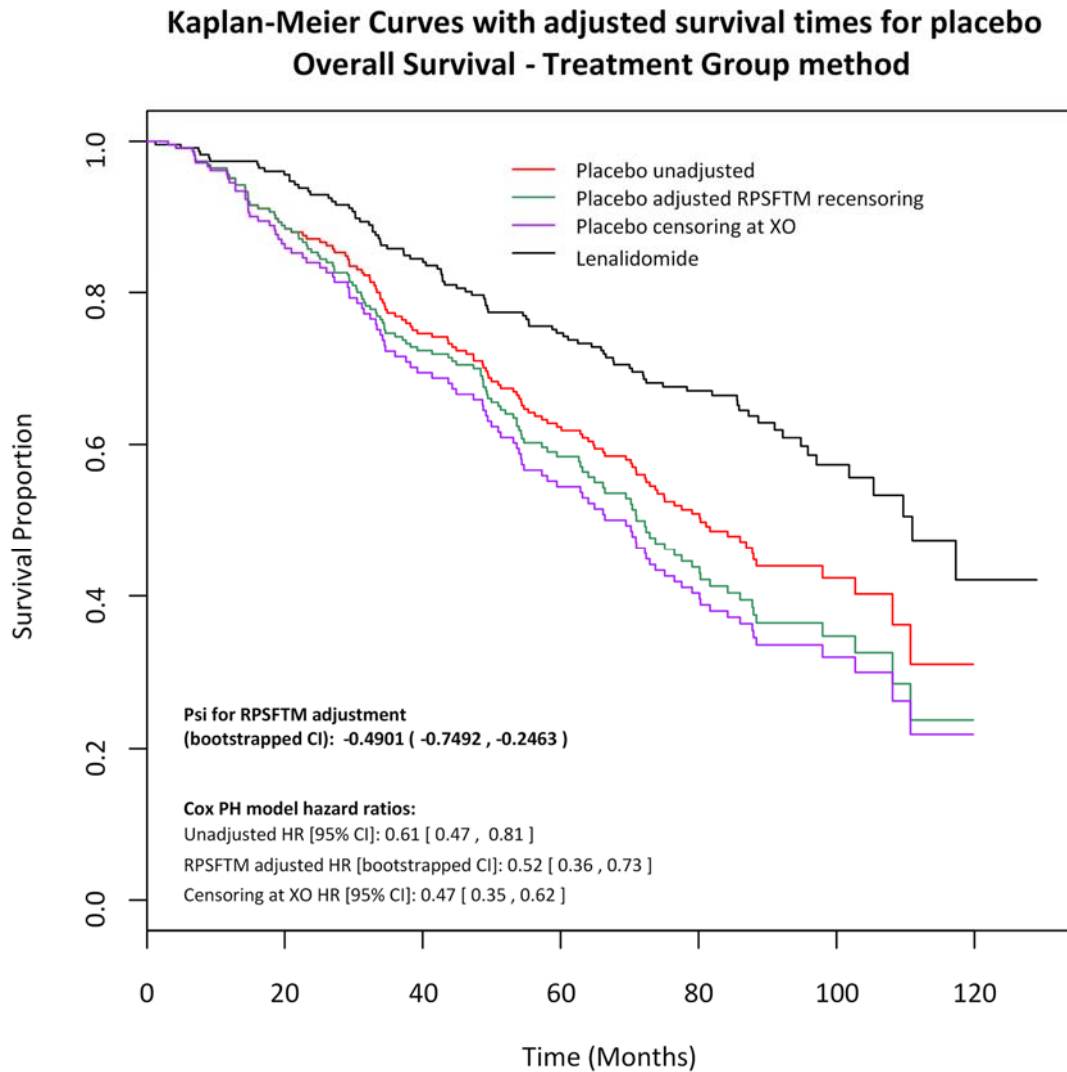
Section 3.1, the analysis that censors survival at switching may be biased in favour of lenalidomide. The RPSFTM HR is 0.52, with a bootstrapped 95% CI of 0.36 to 0.73. As this interval does not contain 1, the difference between lenalidomide and adjusted placebo is statistically significant; this significance is maintained from the ITT analysis, HR of 0.61 (0.47, 0.81), and there is a numerical improvement shown in the RPSFTM result.

Figure 14: Diagnostic plots of counterfactual survival times under the chosen value of ψ (ψ : -0.490) – ‘treatment group’ method – overall survival (CALGB, cut-off 19 Oct 2016)



Key: QQ, quantile-quantile.

Figure 15: KM plot for RPSFTM adjusted OS together with unadjusted OS and OS censoring at switching – ‘treatment group’ method – overall survival (CALGB, cut-off 19 Oct 2016)



Key: CI, confidence interval; HR, hazard ratio; KM, Kaplan–Meier; OS, overall survival; RPSFTM, rank-preserving structural failure time model; XO, switching.

Table 35: Median, lower and upper quartile OS (months) – ‘treatment group’ method – overall survival (CALGB, cut-off 19 Oct 2016)

Analysis	Treatment	25% Quantile	Median (50% Quantile)	75% Quantile
Unadjusted OS	Lenalidomide	59.40	111.01	NA
	Placebo	39.19	80.26	NA
RPSFTM Adjusted OS	Lenalidomide	59.40	111.01	NA
	Placebo	34.56	70.96	110.71
IPE Adjusted OS	Lenalidomide	59.40	111.01	NA
	Placebo	34.88	72.21	110.71
Censoring at XO OS	Lenalidomide	59.40	111.01	NA
	Placebo	33.84	69.45	110.71

Key: NA, not applicable; OS, overall survival; RPSFTM; rank preserving structural failure time model; XO; switching.

Rank-preserving structural failure time model (RPSFTM)/iterative parameter estimation algorithm (IPE)

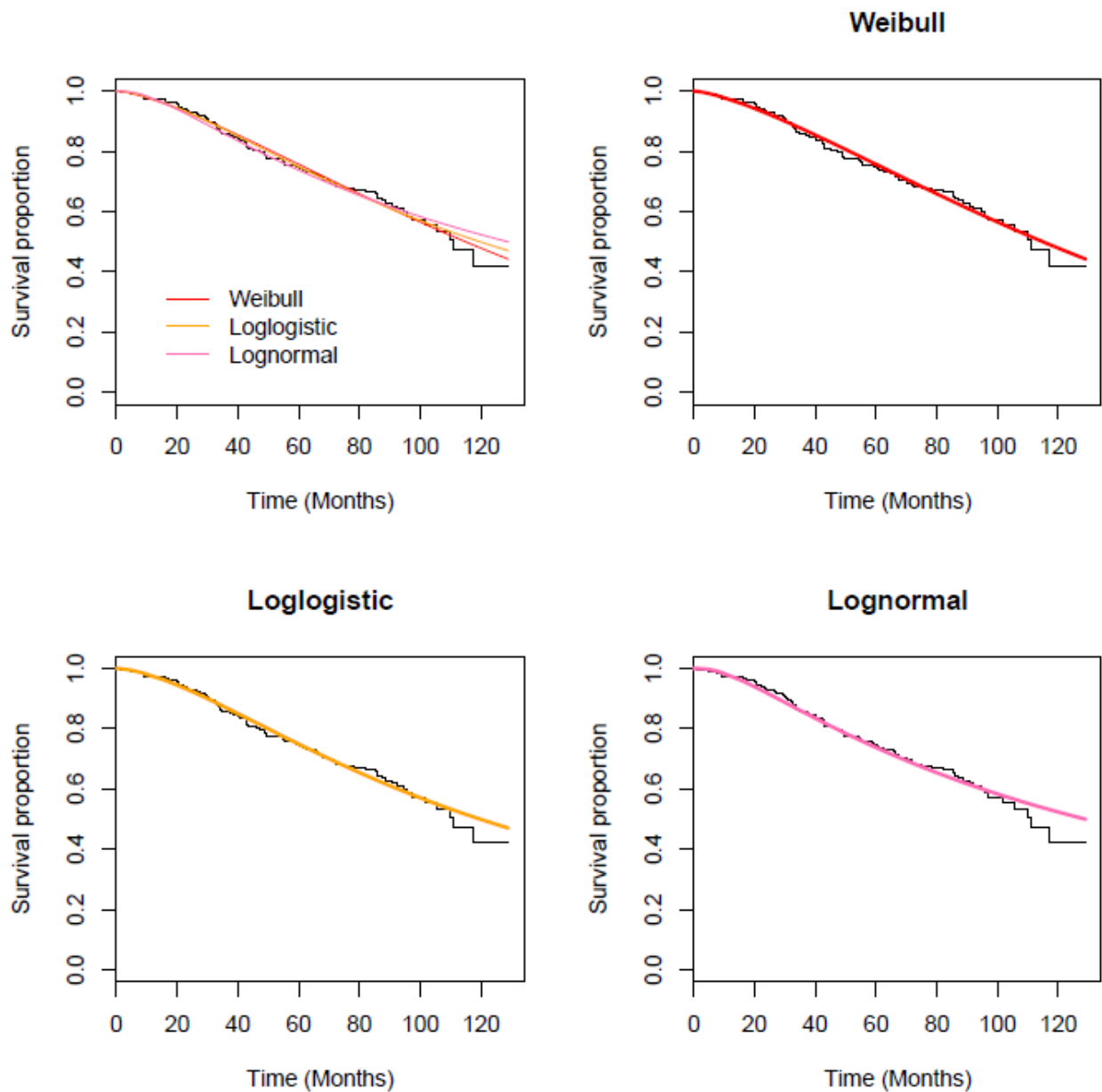
As a way of validating the RPSFTM result, the IPE method was performed. Each model was fitted to the lenalidomide arm in order to identify the best fitting curve and, consequently, the parameterisation to use within the IPE method. Table 36 presents the AIC and BIC values for each parameterisation, and Figure 20 presents the parametric curve fits overlaid onto the lenalidomide KM curve. Based on AIC, BIC and visual inspection of the curves, the Weibull curve was chosen for use in the IPE method.

Table 36: AIC/BIC values for each parameterisation

Model	AIC	BIC
Weibull	1104.546	1111.431
Log-normal	1108.549	1115.434
Log-logistic	1104.611	1111.496

Key: AIC, Akaike information criterion; BIC, Bayesian information criterion.

Figure 16: Lenalidomide KM plots overlaid with parametric curves

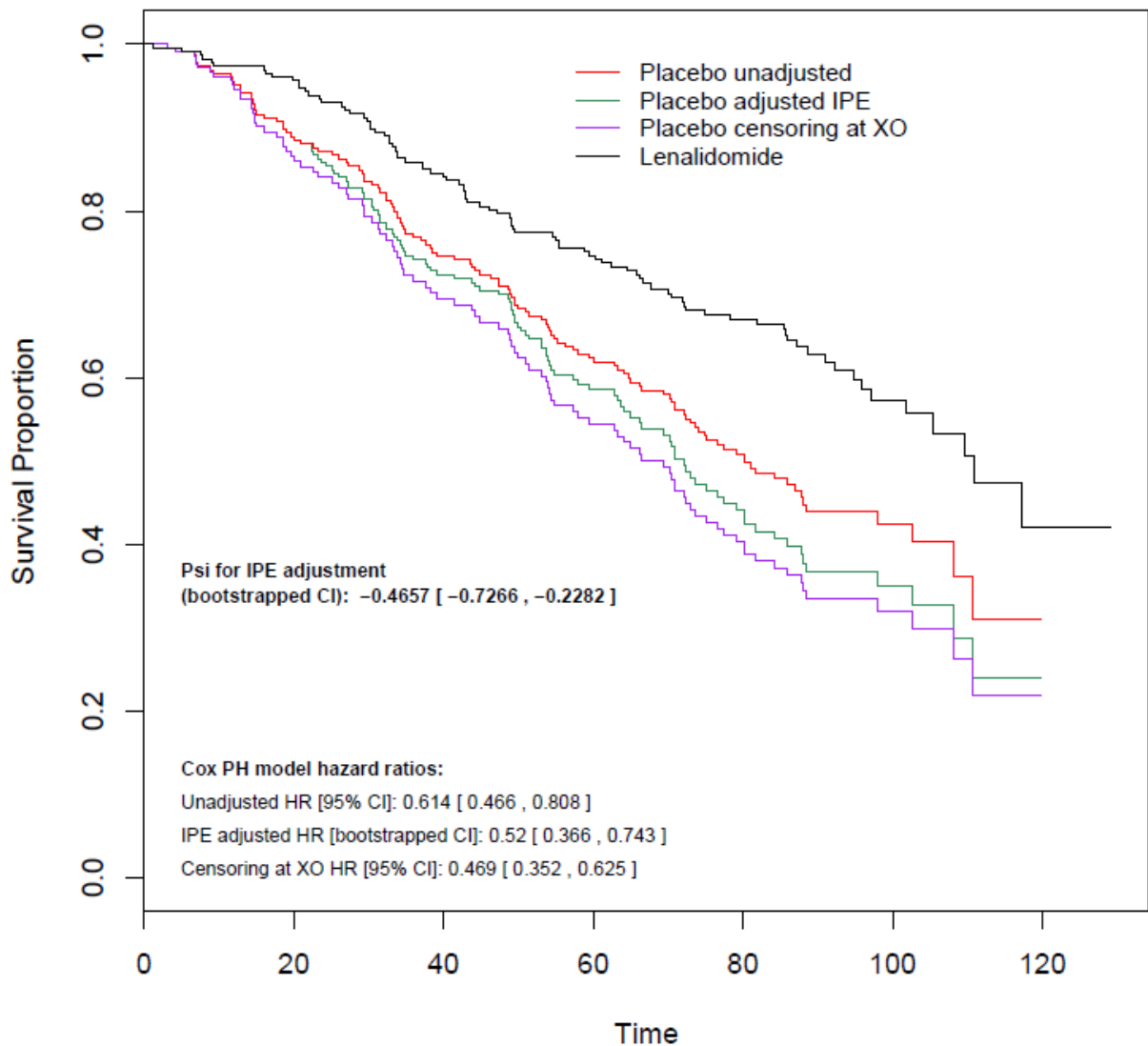


Key: KM, Kaplan–Meier.

An iterative procedure in which parametric models were fitted to the data with a covariate for treatment was implemented, as described in Section 2.5, until the difference in successive e^ψ s was smaller than 10^{-5} . This process resulted in a value of ψ at -0.4657 , and as before, the survival times of patients who switched over from placebo to lenalidomide were adjusted using Equation 1. Figure 17 presents the KM curves for IPE adjusted OS along with unadjusted OS and OS censored at switching.

The results from the IPE adjustment are similar to the RPSFTM adjustment presented above. The hazard ratio is now 0.52 (95% bootstrapped CI [0.37, 0.74]) compared to 0.52 (95% bootstrapped CI [0.36, 0.73]); this consistency provides some support to the RPSFTM results.

Figure 17: KM plot for IPE adjusted OS together with unadjusted OS and OS censoring at switching – ‘treatment group’ method (CALGB, cut-off 19 Oct 2016)



Key: CI, confidence interval; HR, hazard ratio; IPE, iterative parameter estimation; KM, Kaplan–Meier; OS, overall survival; PFS, progression-free survival; XO, switching.

The median, lower and upper quartile survival for IPE adjusted OS was equal (to 2 decimal places) are presented in Table.

4.2 Progression-free survival

As switching occurred prior to progression, PFS was also affected by patients switching over to lenalidomide; therefore, adjustment for switching was also made for this endpoint.

RPSFTM

Figure 18 presents ψ plotted against the corresponding log-rank test statistic. The plot is approximately monotonic and results in a unique solution for ψ at -0.547. The KM and QQ plots of the counterfactual survival calculated for this solution for ψ are presented in Figure 19. There is a slight tendency in the QQ plot to lie slightly to one side of the identity line. However, the KM plot suggests that the absolute magnitude of this discrepancy is small.

Figure 18: Log-rank test statistics plotted against the corresponding values of ψ – ‘treatment group’ method – progression-free survival (CALGB, cut-off 19 Oct 2016)

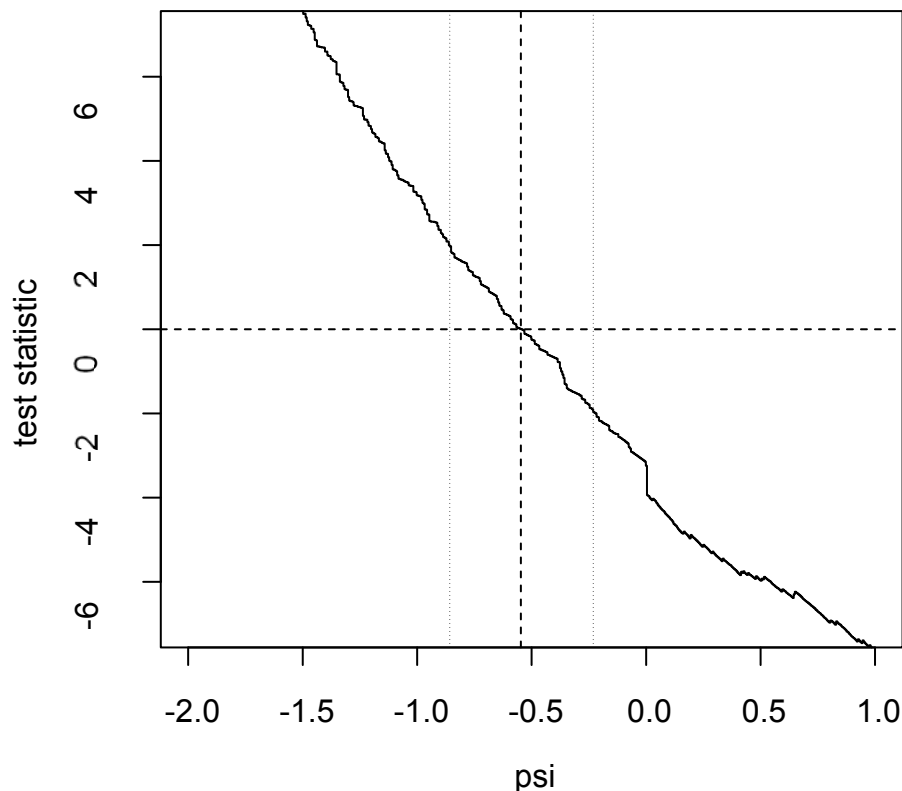
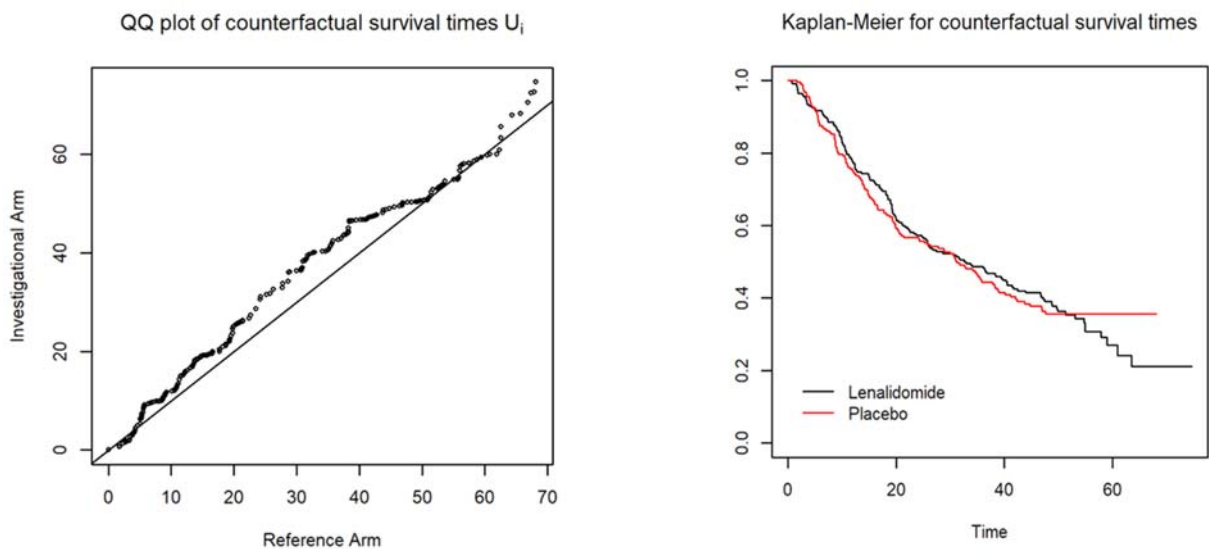


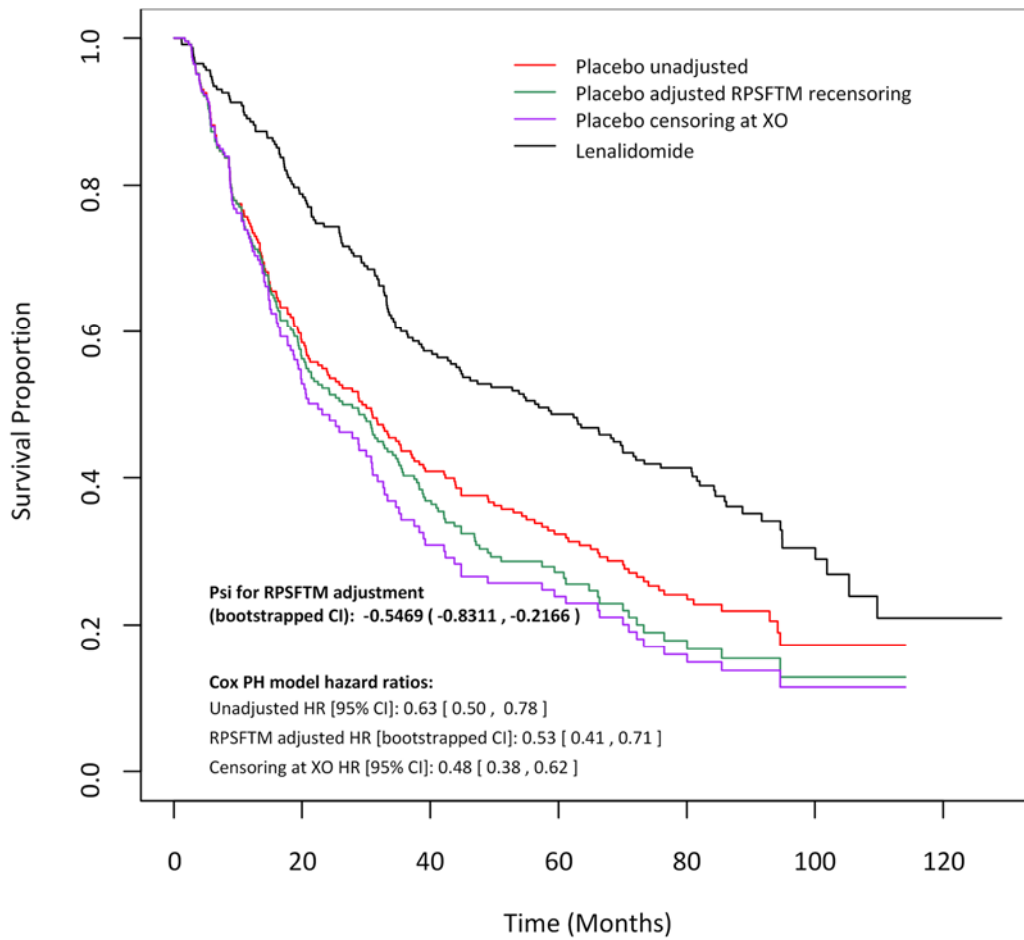
Figure 19: Diagnostic plots of the counterfactual survival time under the chosen value of ψ (ψ : -0.547) – ‘treatment group’ method – progression-free survival (CALGB, cut-off 19 Oct 2016)



Key: QQ, quantile-quantile.

The KM for RPSFTM adjusted PFS based on the selected value of ψ is shown in Figure 20, along with unadjusted PFS and PFS censored at switching. As with OS, the majority of the adjusted PFS KM lies between unadjusted PFS and PFS censored at switching and results in an HR of 0.53 (0.41, 0.71), maintaining the significant difference shown between lenalidomide and placebo in the ITT analysis (HR 0.63 [0.50, 0.78]) and showing a numerical improvement over the ITT.

Figure 20: KM plot for RPSFTM adjusted PFS together with unadjusted PFS and PFS censoring at switching – ‘treatment group’ method (CALGB, cut-off 19 Oct 2016)



Key: CI, confidence interval; HR, hazard ratio; KM, Kaplan–Meier; PFS, progression-free survival; RPSFTM, rank-preserving structural failure time model; XO, switching.

Table 37: Median, lower and upper quartile PFS (months) – ‘treatment group’ method – progression-free survival (CALGB, cut-off 19 Oct 2016)

Analysis	Treatment	25% Quantile	Median (50% Quantile)	75% Quantile
Unadjusted PFS	Lenalidomide	22.08	56.87	105.29
	Placebo	11.70	29.37	75.66
RPSFTM Adjusted PFS	Lenalidomide	22.08	56.87	105.29
	Placebo	10.91	26.31	64.77
IPE Adjusted PFS*	Lenalidomide	22.08	56.87	105.29
	Placebo	10.55	25.79	59.36
Censoring at XO PFS	Lenalidomide	22.08	56.87	105.29
	Placebo	10.91	22.47	57.42

Key: IPE, iterative parameter estimation; NA, not applicable; PFS, Progression Free Survival; RPSFTM, rank preserving structural failure time model; XO, switching.

Notes: *Presented below.

IPE method

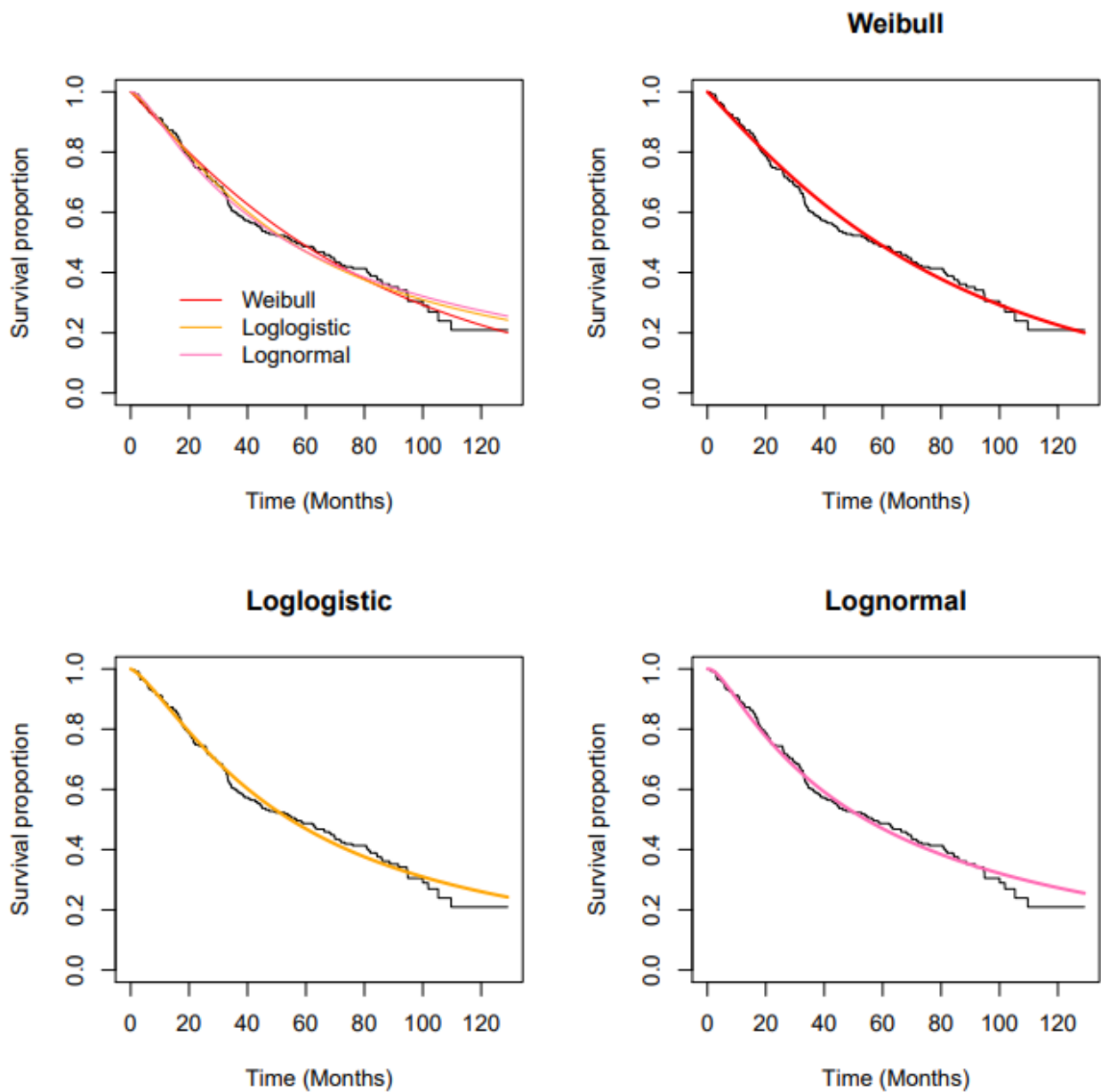
As above, as a validation the RPSFTM result, the IPE method was also conducted. Table 38 presents the AIC and BIC values for each parametrisation, and Figure 21 presents the parametric curve fits overlaid on the lenalidomide KM curve for PFS. Based on the AIC, BIC and visual inspection of the curves, the log-logistic curve was chosen for use in the IPE.

Table 38: AIC/BIC values for each parameterisation

Model	AIC	BIC
Weibull	1597.148	1604.033
Log-normal	1594.361	1601.246
Log-logistic	1594.103	1600.988

Key: AIC, Akaike information criterion; BIC, Bayesian information criterion.

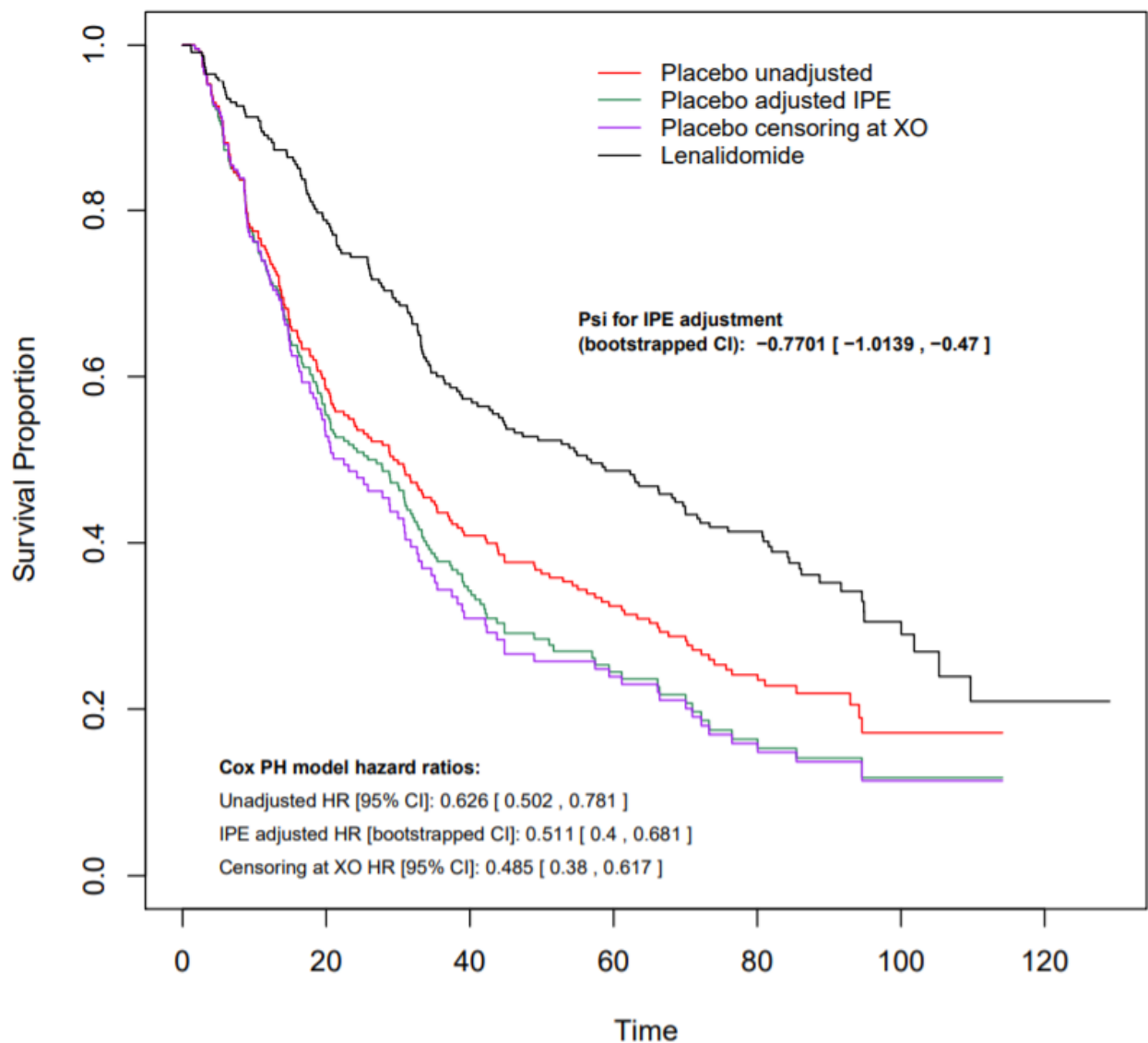
Figure 21: Lenalidomide KM plots for PFS overlaid with parametric curves



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

The IPE method was implemented with the log-logistic distribution to fit the parametric curves for the iterative procedure, resulting in a value of ψ as -0.770. The IPE adjusted KM and KMs for unadjusted PFS and PFS censored at switching are presented in Figure 22. The results of the IPE are similar to RPSFTM, producing a hazard ratio of 0.51 (0.40, 0.68) in comparison to 0.53 (0.41, 0.71).

Figure 22: KM plot for IPE adjusted PFS together with unadjusted PFS and PFS censoring at switching – ‘treatment group’ method – (CALGB, cut-off 19 Oct 2016)



Key: CI, confidence interval; HR, hazard ratio; IPE, iterative parameter estimation; KM, Kaplan–Meier; PFS, progression-free survival; XO, switching.

6 Sensitivity analysis: ‘On treatment’ approach

4.3 Overall survival

RPSFTM

Figure 23 shows ψ presents a plot of each value of ψ against their corresponding test statistics. The plot has some non-monotonic features but in general does tend to decrease for increasing ψ , with one solution for ψ : -1.229. The KM and QQ plots of

the counterfactual survival calculated for this solution for ψ are presented in Figure 24. The QQ plot follows the identity line reasonably well. The KM plot of counterfactual survival times shows that the distribution of survival times is approximately equal for the majority of time. Overall, the requirements of the RPSFTM method appear to be satisfied.

Figure 23: Log-rank test statistics plotted against the corresponding values of ψ – ‘on treatment’ method – (CALGB, cut-off 19 Oct 2016)

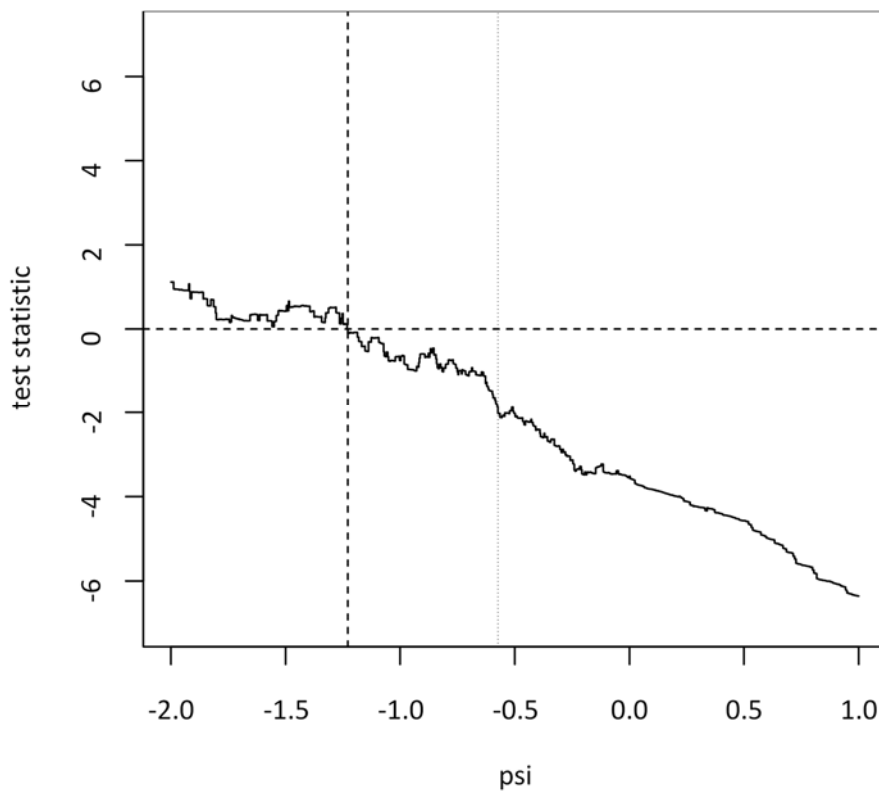
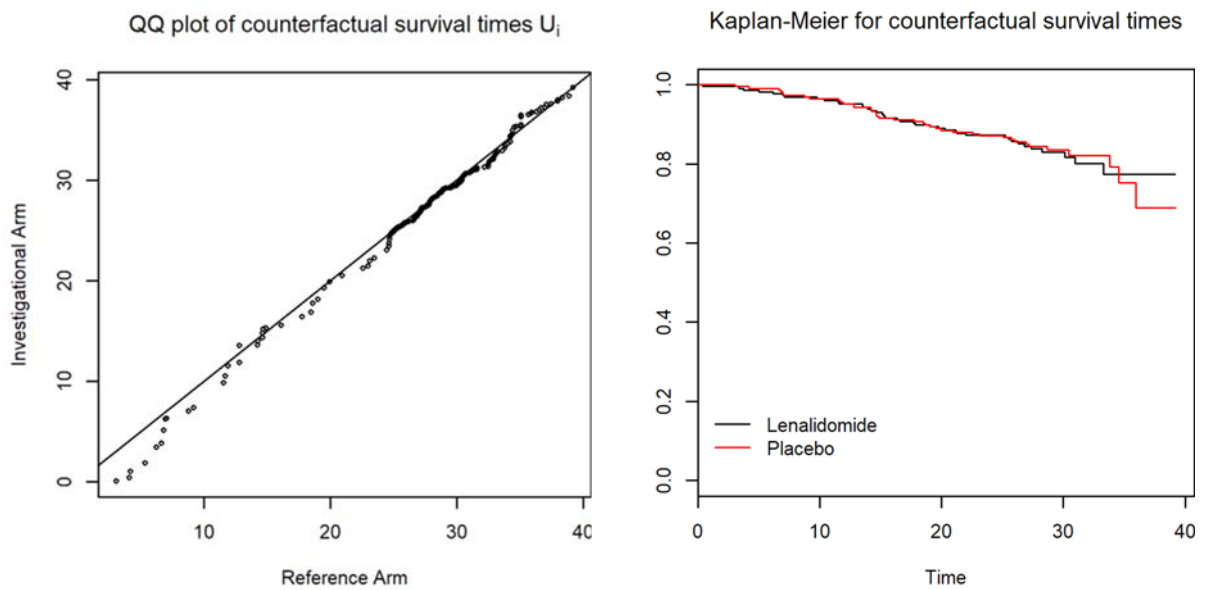
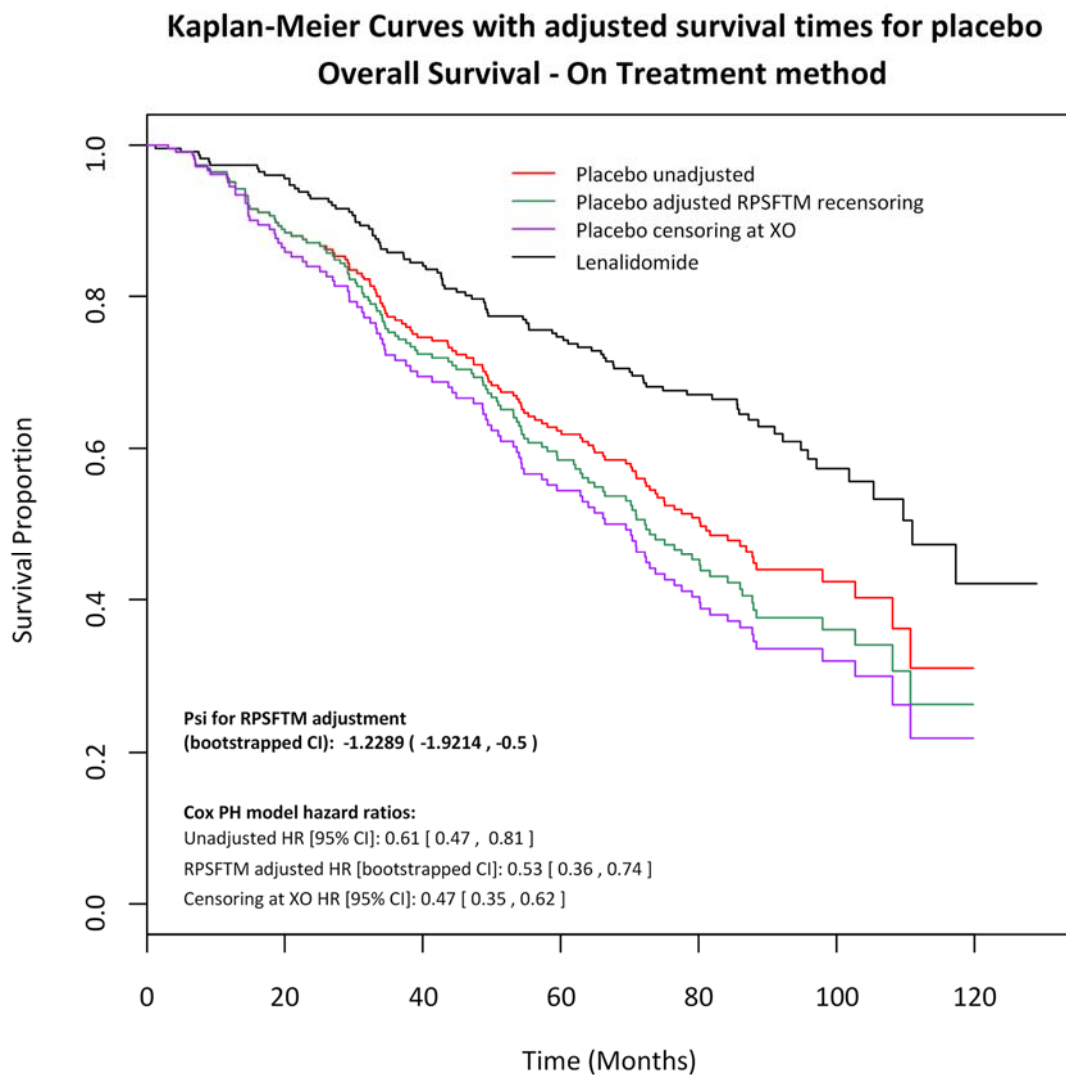


Figure 24: Diagnostic plots of the counterfactual survival time under one solution for ψ (ψ : -1.177) – ‘on treatment’ method – (CALGB, cut-off 19 Oct 2016)



Key: QQ, quantile-quantile.

Figure 25: KM plot for RPSFTM adjusted OS together with unadjusted OS and OS censoring at switching – ‘on treatment’ method – (CALGB, cut-off 19 Oct 2016)



Key: CI, confidence interval; HR, hazard ratio; KM, Kaplan–Meier; OS, overall survival; RPSFTM, rank-preserving structural failure time model; XO, switching.

Figure 29 presents the RPSFTM adjusted OS for the ‘on treatment’ approach as well as unadjusted OS and OS censored at switching. As with previous results, the RPSFTM adjusted OS lies between unadjusted OS and OS censored at switching. The KM results are also mostly consistent to the outcomes of the ‘treatment group’ approach leading to a HR (95% CI) of 0.53 (0.36, 0.74).

Table 39: Median, lower and upper quartile OS (months) – ‘on treatment’ method – (CALGB, cut-off 19 Oct 2016)

Analysis	Treatment	25% Quantile	Median (50% Quantile)	75% Quantile
Unadjusted OS	Lenalidomide	59.40	111.01	NA
	Placebo	39.19	80.26	NA
RPSFTM adjusted OS	Lenalidomide	59.40	111.01	NA
	Placebo	35.97	72.21	NA
IPE adjusted OS*	Lenalidomide	59.40	111.01	NA
	Placebo	39.19	76.48	NA
Censoring at XO OS	Lenalidomide	59.40	111.01	NA
	Placebo	33.84	69.45	110.71

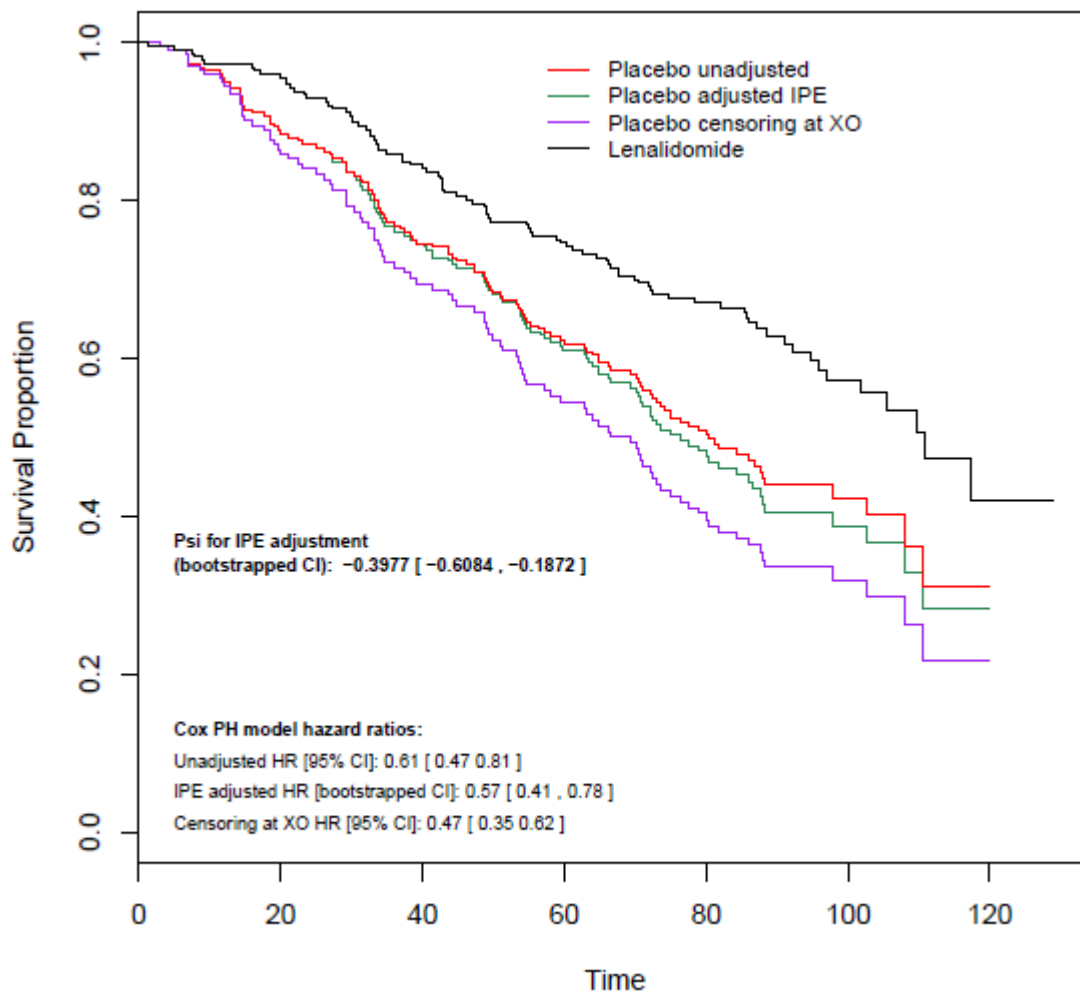
Key: IPE; iterative parameter estimation; NA; not applicable; OS; overall survival; RPSFTM, rank preserving structural failure time model; XO, switching.

Notes: *Presented below.

IPE

The IPE method was also conducted for the ‘on treatment’ approach. The Weibull parametrisation was again used as the most appropriate parametric model, as the choice of best fit was based on the lenalidomide arm only, which does not change between the two approaches. The IPE was conducted using the Weibull model and resulted in a value of ψ as -0.3977. The KM plots of IPE adjusted OS, unadjusted OS and OS censored at switching are shown in Figure 26. The IPE adjusted HR was 0.57 (95% CI; 0.41, 0.78), which is marginally but not substantially higher than the HR from the RPSFTM, and the same conclusions as the RPSFTM can be drawn here.

Figure 26: KM plot for IPE adjusted OS together with unadjusted OS and OS censoring at switching – ‘on treatment’ method – (CALGB, cut-off 19 Oct 2016)



Key: CI, confidence interval; HR, hazard ratio; IPE, iterative parameter estimation; KM, Kaplan–Meier; PFS, progression-free survival; XO, switching.

4.4 Progression-free survival

RPSFTM

Figure 27 presents a plot of each value of ψ against their corresponding test statistics. The plot is slightly non-monotonic; however, one solution for ψ is identified ($\psi = -1.1248$). The KM and QQ plots of the counterfactual survival calculated for this solution for ψ are presented in Figure 28. The QQ plot follows the identity line reasonably well. There is a slight tendency in the QQ plot to lie slightly to one side of the identity line. However, this difference is not reflected in the KM plots and

therefore it is concluded that the requirements of the RPSFTM method are not satisfied.

Figure 27: Log-rank test statistics plotted against the corresponding values of ψ – ‘on treatment’ method – (CALGB, cut-off 19 Oct 2016)

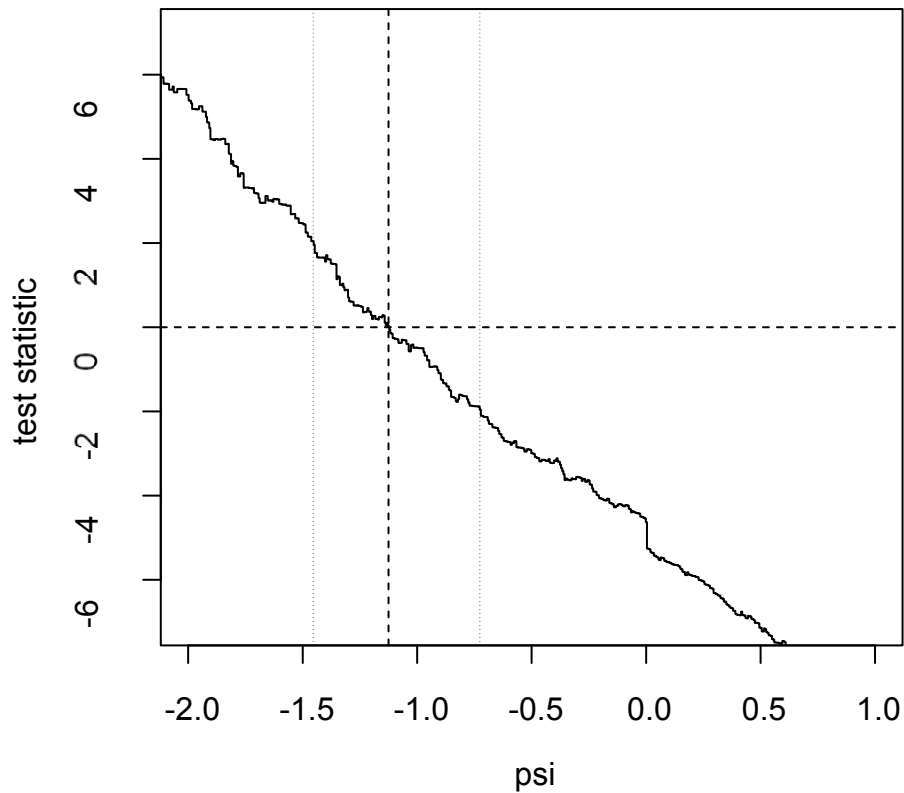
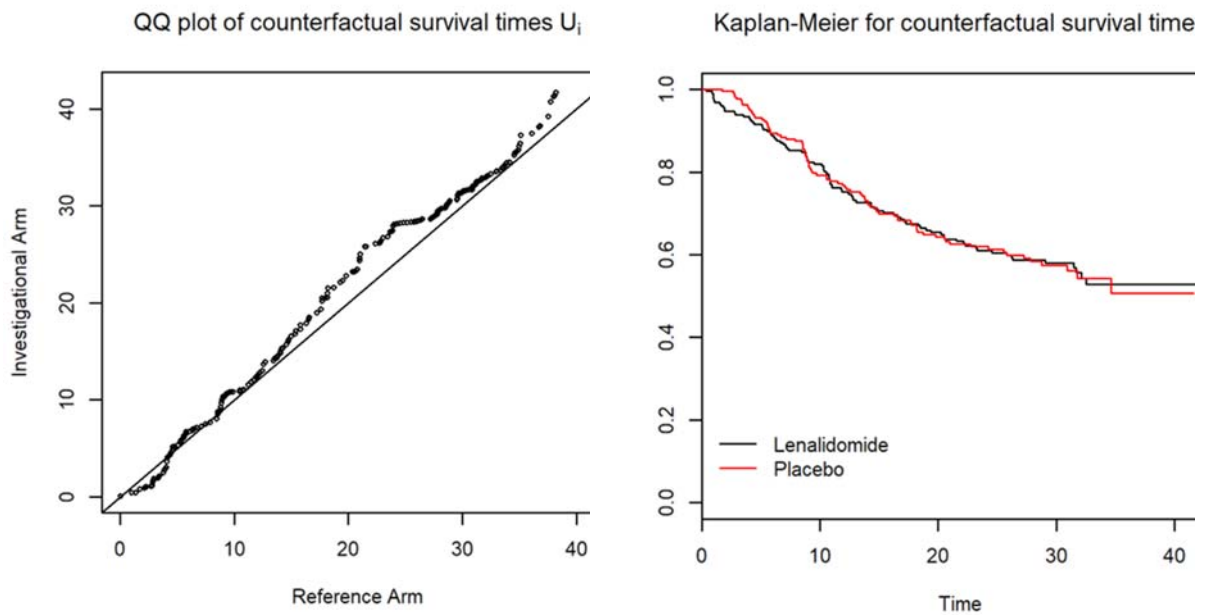


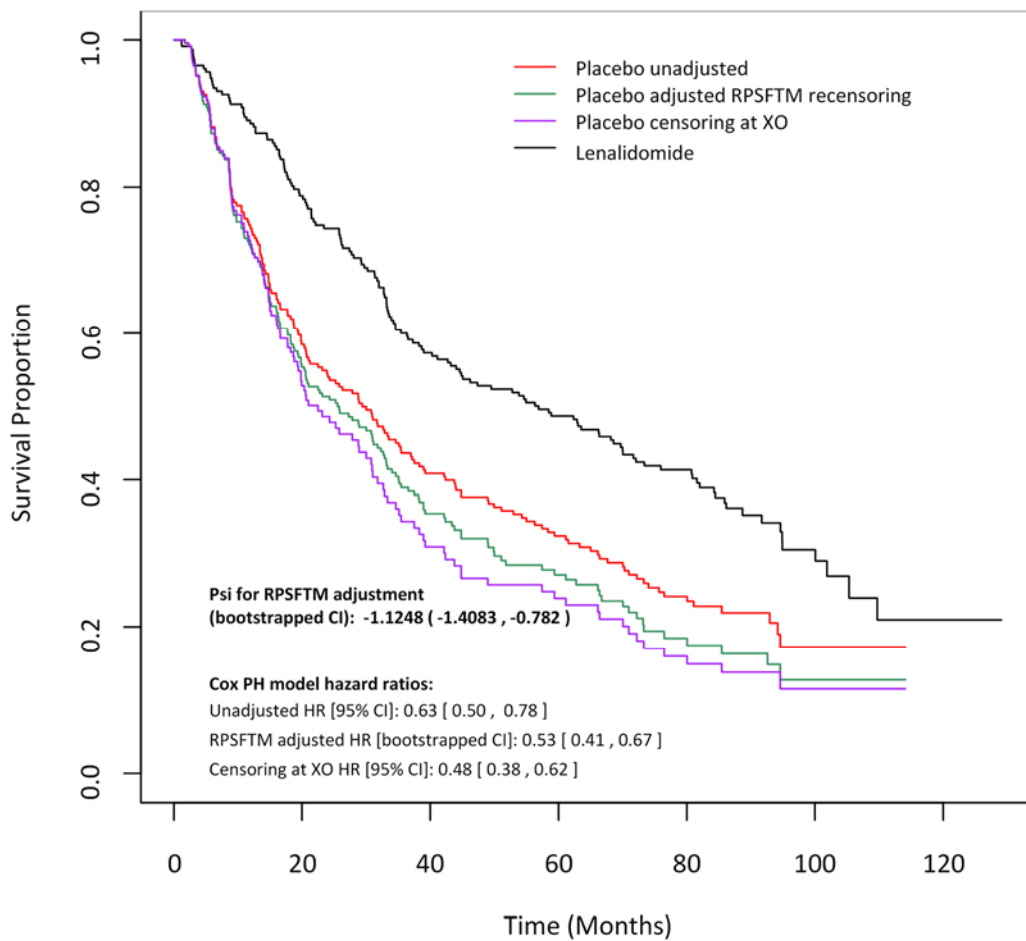
Figure 28: Exploratory plots of the counterfactual survival time under one solution for ψ (ψ : -1.491) – ‘on treatment’ method – (CALGB, cut-off 19 Oct 2016)



Key: QQ, quantile-quantile.

Figure 29 presents the RPSFTM adjusted PFS along with unadjusted PFS and PFS censored at switching, RPSFTM adjusted PFS lies mostly between the two and results in an HR of 0.53 (bootstrapped 95% CI [0.41, 0.67]), again maintaining the significance shown in the ITT analysis and showing a numerical improvement over this. This result is also consistent with the results of the ‘treatment group’ approach.

Figure 29: KM plot for RPSFTM adjusted PFS together with unadjusted PFS and PFS censoring at switching – ‘on treatment’ method – (CALGB, cut-off 19 Oct 2016)



Key: CI, confidence interval; HR, hazard ratio; KM, Kaplan–Meier; PFS, progression-free survival; RPSFTM, rank-preserving structural failure time model; XO, switching.

Table 40: Median, lower and upper quartile PFS (months) – ‘on treatment’ method – (CALGB, cut-off 19 Oct 2016)

Analysis	Treatment	25% Quantile	Median (50% Quantile)	75% Quantile
Unadjusted PFS	Lenalidomide	22.08	56.87	105.29
	Placebo	11.70	29.37	75.66
RPSFTM adjusted PFS	Lenalidomide	22.08	56.87	105.29
	Placebo	10.55	25.54	66.39
IPE adjusted PFS*	Lenalidomide	22.08	56.87	105.29
	Placebo	10.63	26.57	67.28
Censoring at XO PFS	Lenalidomide	22.08	56.87	105.29
	Placebo	10.91	22.47	57.42

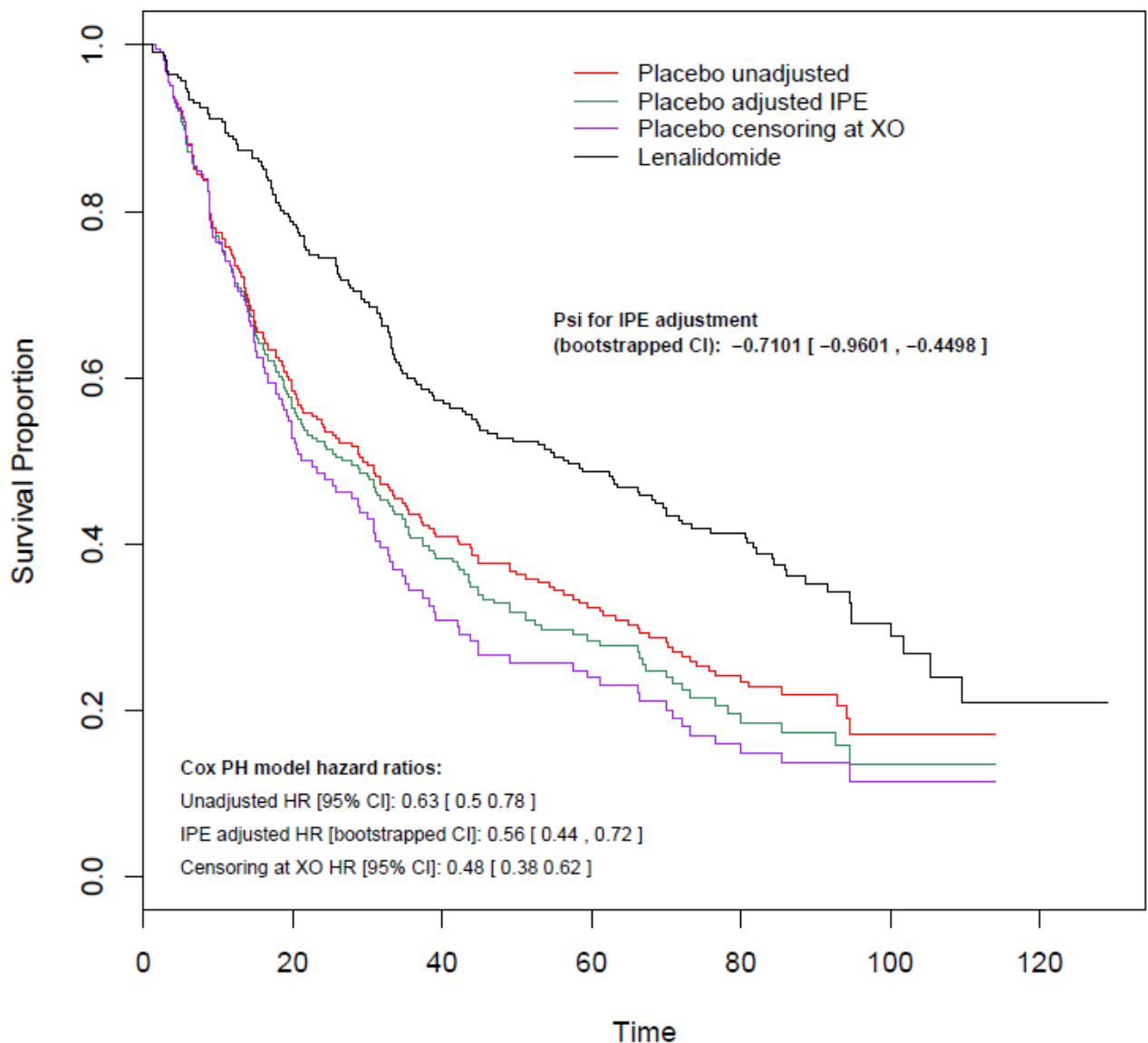
Key: IPE; iterative parameter estimation; NA; not applicable; PFS; Progression Free Survival; RPSFTM, rank preserving structural failure time model; XO, switching.

Notes: *Based on the most conservative solution for ψ

IPE

The IPE method was also conducted for the ‘on treatment’ approach. The log-logistic parameterisation was again used as the most appropriate parametric model, as the choice of best fit was based on the lenalidomide arm only, which does not change between the two approaches. The IPE was conducted using the log-logistic model, resulting in a value of ψ at -0.7101. The KM plots of adjusted PFS, unadjusted PFS and PFS censored at switching are shown in Figure 30. Again, the IPE adjusted PFS lies between the two, and has a hazard ratio of 0.56 (bootstrapped 95% CI [0.44, 0.72]), keeping the statistical significance shown in the ITT analysis and showing a numerical improvement.

Figure 30: KM plot for IPE adjusted PFS together with unadjusted PFS and PFS censoring at switching – ‘on treatment’ method – (CALGB, cut-off 19 Oct 2016)



Key: CI, confidence interval; HR, hazard ratio; IPE, iterative parameter estimation; KM, Kaplan–Meier; PFS, progression-free survival; XO, switching.

4.5 Covariate adjustment for ψ

Within each RPSFTM procedure, covariate adjustment for the calculation of ψ was explored for the covariates of prior lenalidomide treatment, prior thalidomide treatment and beta2-Microglobulin.

Table presents ψ estimated using both adjusted and un-adjusted models. Minimal differences were observed in the estimated values of ψ between the two models which therefore lead to approximately equivalent HRs in each case.

Table 41: Comparison of estimates around ψ and adjusted HR for adjustment and no adjustment for ψ

Analysis	Calculation for ψ unadjusted for covariates		Calculation for ψ adjusted for covariates of prior lenalidomide, prior thalidomide and beta2-Microglobulin	
	ψ (bootstrapped 95% CI)	HR (bootstrapped 95% CI)	ψ (bootstrapped 95% CI)	HR (bootstrapped 95% CI)
RPSFTM OS, 'treatment group'	-0.490 [-0.749, -0.246]	0.52 [0.36, 0.73]	-0.491 [-0.742, -0.242]	0.52 [0.37, 0.74]
RPSFTM PFS, 'treatment group'	-0.547 [-0.837, -0.231]	0.53 [0.42, 0.72]	-0.535 [-0.862, -0.231]	0.54 [0.41, 0.71]
RPSFTM OS, 'on treatment'	-1.229 [-1.921, -0.50]	0.53 [0.36, 0.74]	-1.233 [-1.99, -0.445]	0.53 [0.38, 0.76]
RPSFTM PFS, 'on treatment'	-1.124 [-1.449, -0.723]	0.53 [0.42, 0.66]	-1.124 [-1.452, -0.179]	0.53 [0.42, 0.69]

Key: CI, confidence interval; OS, overall survival; PFS, progression free survival; RPSFTM, rank preserving structural failure time model.

Notes: *Comparison made for one possible solution of ψ .

7 Conclusions

The purpose of these analyses was to estimate the relative treatment effect for lenalidomide compared to placebo for OS and PFS, adjusting for the potential diluting effects introduced by patients switching over to lenalidomide from the placebo arm.

The analyses presented in this report are based upon an updated 19 October 2016 data cut-off, following on from results of an earlier version of this data cut presented by Holstein et al. (2017).

A per protocol analysis of censoring at switching was presented in the CSR (Bertagnolli M and Brigham and Women's Hospital, 2013; 2016). Although censoring at switching is not generally recommended (Latimer, 2014; Watkins et al., 2013) because it assumes that censoring is not informative (the probability of censoring is

not related to prognosis), this method may have some merits when applied to the data of CALGB. The process by which a patient was permitted to switching was not explicitly based on a patient's prognosis; the DSMB recommended that the placebo therapy be stopped, and the patients be given the opportunity to initiate lenalidomide therapy.

With the aim to provide a more robust switching adjustment, a review of the available methods and their relative merits as applied to the data from CALGB was performed. Following this review, it was recommended to further explore the RPSFT and IPE methodologies and the validity of the 'common treatment effect' assumption which underpins them. The common treatment effect assumes that the experimental treatment effect is the same (relative to first dose of treatment) regardless of whether it is administered at randomisation or at a later timepoint. Although this assumption is largely untestable, analyses were performed to explore this assumption. Specifically, a landmark analysis was performed using this updated data cut-off. This analysis compared the survival outcomes of all patients ongoing as of the date of unblinding (17 December 2009). The treatment effect (HR) between those who did and did not switching (0.57 [95% CI; 0.29,1.15]) was not of a different magnitude to that of the ITT analysis measured from the point of randomisation (0.61, 95% confidence interval [CI] [0.47, 0.81]). As such, there was no data to suggest that the common treatment effect assumption which underpins the RPSFT and IPE methodologies was severely violated, and both methods were employed to adjust for switching, the results of which are summarised below.

Within the RPSFTM method, two approaches were taken, which partitioned the survival time for each patient differently: the 'treatment group' approach assumed that the residual effect of treatment with lenalidomide was retained until death, and the 'on treatment' approach assumed that there was no residual effect of lenalidomide after discontinuation. These two approaches were also the two extremes for partitioning survival time (i.e. total residual effect or no residual effect of lenalidomide). Based upon the mechanism of action of lenalidomide, the 'treatment group' approach was chosen as the base case as it was expected that some of the efficacy of lenalidomide will be retained past the last dose of treatment. The 'on treatment' approach was conducted as a sensitivity analysis.

For OS with the 'treatment group' approach, the statistically significant difference between lenalidomide and placebo shown in the ITT analysis was maintained using either RPSFTM or IPE, and numerical improvements on the HR were shown in both cases. The results were similar for both RPSFTM and IPE, with HRs and bootstrapped 95% CIs of 0.52 (0.36, 0.73) and 0.52 (0.37, 0.74), respectively. However, the results of the RPSFTM and IPE methods were more conservative than censoring at switching; HR and 95% CI of 0.47 (0.35, 0.62). This is as expected, the method of censoring at switching may be subject to some degree of informative censoring leading to biased treatment effect estimators; patients who switched treatment may have had more favourable prognostic disease characteristics compared with the overall placebo arm.

A similar improvement was observed when using RPSFTM and IPE to adjust PFS for switching using the 'treatment group' approach. The ITT PFS analyses showed a statistically significant difference between lenalidomide and placebo (HR 0.63, 95% CI [0.50, 0.78]), and this significance is maintained and improved numerically after adjusting for switching using the 'treatment group' approach, giving HRs and bootstrapped 95% CIs of 0.53 (0.41, 0.71) and 0.51 (0.40, 0.68) for RPSFTM and IPE, respectively. In line with the CALGB CSR, the per protocol method of censoring at switching for PFS was defined using the date of last adequate assessment before switching. The results of the RPSFTM and IPE methods were again more conservative than censoring at switching; HR and 95% CI of 0.48 (0.38, 0.62). Again, the method of censoring at switching may be subject to some degree of informative censoring.

For OS, the application of the RPSFTM method for the 'on-treatment' approach resulted in consistent treatment effects with the 'treatment group' approach result. The IPE method was also conducted for the 'on treatment' approach; resulting in an HR of 0.57, which, although slightly higher than the 'treatment group' approach still shows a greater benefit than the ITT analysis. For PFS, the adjusted HR was 0.53 (bootstrapped 95% CI [0.36, 0.74]) in comparison to an HR of 0.63 from the ITT analysis. This result is also similar to those of the 'treatment group' approach stated above and reasonably consistent with the result from the IPE analysis, which gave a HR of 0.56 (bootstrapped 95% CI [0.44, 0.72]).

In addition, with the accrual of 9 months of additional data since the last switching adjustment analysis was performed, the switching adjusted results for the 19 October 2016 data-cut have remained consistent. This is not unexpected as the ITT hazard ratios, especially for OS, have remained consistent in this time period.

References

Bertagnolli M and Brigham and Women's Hospital. (2013) CALGB 100104. Phase III Randomized, Double-Blind Study of Maintenance Therapy with Lenalidomide CC-5013 (NSC #703813, IND # 70116) or Placebo Following Autologous Stem Cell Transplantation for Multiple Myeloma (CALGB 100104) Intergroup Study of Lenalidomide versus Placebo Maintenance Therapy following Single Autologous Stem Cell Transplant (ASCT) for Multiple Myeloma. Clinical Study Report. Data on File.

Bertagnolli M and Brigham and Women's Hospital. (2016) CALGB 100104, A Phase III randomized, double-blind study of maintenance therapy with lenalidomide CC-5013 (NSC # 703813, IND # 70116) or placebo following autologous stem cell transplantation for multiple myeloma. Clinical study report. Data on file.

Bond S (2015). rpsftm: Rank Preserving Structural Failure Time Models. R package version 1.0.

Branson M and Whitehead J. (2002) Estimating a treatment effect in survival studies in which patients switch treatment. *Stat Med.*, 21(17):2449-63

BresMed (2016). Adjustment for treatment switching in the Cancer and Leukemia Group B (CALGB) trial in Multiple Myeloma: Feasibility assessment. Updated: 17 March 2016. Data on File.

BresMed (2017a). Adjustment for treatment switching in the CALGB trial in multiple myeloma: Statistical analysis report (Data cut-off 01 February 2016). Updated: 23 February 2017. Data on File.

BresMed (2017b). Adjustment for treatment switching in the CALGB trial in multiple myeloma: Statistical analysis report. Updated: 10 February 2017. Data on File.

Canty, A and Ripley, B (2016). boot: Bootstrap R (S-Plus) Functions. R package version 1.3-18.

Latimer NR and Abrams KR. NICE DSU Technical Support Document 16: Adjusting survival time estimates in the presence of treatment switching. (2014) Located at:

http://www.nicedsu.org.uk/TSD16_Treatment_Switching.pdf . Accessed: 26 May 2015.

Holstein SA, Jung SH, Richardson PG, Hofmeister CC, Hurd DD, Hassoun H, Giralt S, Stadtmauer EA, Weisdorf DJ, Vij R, Moreb JS. (2017). Updated analysis of CALGB (Alliance) 100104 assessing lenalidomide versus placebo maintenance after single autologous stem-cell transplantation for multiple myeloma: a randomised, double-blind, phase 3 trial. *Lancet Haematology* 4(9): e431-e442.

R Core Team (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.

Robins JM and Tsiatis AA. (1991) Correcting for noncompliance in randomized trials using rank preserving structural failure time models. *Commun Stat Theory Methods*, 20(8):2609-2631.

Therneau T (2015). *_A Package for Survival Analysis in S_*. version 2.38, URL <http://CRAN.R-project.org/package=survival> .

Watkins C, Huang X, Latimer N, et al. (2013) Adjusting overall survival for treatment switches: commonly used methods and practical application. *Pharmaceut Statist*, 12(6):348-57

White IR, Babiker AG, Walker S and Darbyshire JH. (1999) Randomization-based methods for correcting for treatment changes: examples from the Concorde trial. *Stat Med*, 18(19):2617-34.

Issue 8. Detailed methods for how the relative dose intensity was calculated using Myeloma XI data

1 Relative dose intensity (RDI) from Myeloma XI data

The objective of the Myeloma XI drug prescription analysis was to obtain the mean cost of lenalidomide prescribing for maintenance. Additional analyses of drug consumption data from Myeloma XI were performed, to obtain an estimate for relative dose intensity to apply to the economic model for lenalidomide in maintenance. The analysis had two objectives:

1. To obtain separate estimates for RDI for lenalidomide 10mg and 5g dosages, in a form suitable to incorporate the results from Myeloma XI data into the economic model for maintenance with lenalidomide and
2. To incorporate the effect of non-compliance and wastage into the costing of maintenance with lenalidomide.

As is customary in NICE submissions and has been the method of relevance in other settings⁵, it is appropriate for economic models that a relative dose intensity is applied to determine the total cost of drug, given consumption. The relative dose intensity was therefore calculated from the Myeloma XI trial and applied in the model to cost maintenance treatment.

In general, a relative dose intensity has been applied for indications when the dosage received by the patient is variable, based on patient weight. The volume of drug used in one cycle of therapy, for example, injectables, can be less than 100% of the dispensed dose unit (i.e. the vial) when posology is determined with flexible dosing, based on patient weight. In this context, it is appropriate to impute the cost of the proportion of drug discarded if the dispensed vial is larger than the total dose required.

⁵ <https://www.nice.org.uk/guidance/ta555/evidence/committee-papers-ta555-pdf-6654880909>, pp. 9

This is because the 'discarded' fraction cannot be redeployed. In other contexts, however, the 'wasted' fraction can be reused in the next patient, in which case the cost of treatment is calculated strictly that of the proportion of dispensing unit used for each patient (i.e. less than 100% of the vial).

The situation described above applies to flexible dose therapies.

Lenalidomide maintenance, on the contrary, is a fixed dose therapy. This means that patients are started on a 10mg pill per day as a standard, taken for 21 days in a cycle of 28, with a 7 day break. The dose is not personalised by patient weight.

In case the initial dose of 10mg is not well tolerated, and according to label, the daily dose can be reduced to 5mg for 21 days in 28 days, with a 7-day break. The justification of the 21 days in 28 day dosing has been discussed elsewhere before.

In addition, whilst in regulatory trials patients would be encouraged or prompted to strictly adhere to treatment protocol, in real world practice it is often the case that patients may be less than 100% compliant. In the case of maintenance, however, treatment intervals or reduction in frequency of intake is also a therapeutic strategy that physicians use to ensure tolerability, as well as to keep patients on maintenance treatment for as long as possible. This approach to therapy is specific to maintenance. According to clinical opinion, physicians can take other approaches to dose adjustments: either by spacing doses and cycles, or using mixtures of two doses, or a combination of the these approaches (**Error! Reference source not found.**).

For this reason, the interval between one treatment cycle and another may be longer than the 7-days break; consequently, treatment intensity is diluted with the effect of decreasing the drug cost of maintenance.

Finally, the Myeloma XI data showed that real-life dosing for some patients may be adapted by clinicians, reflecting less than 100% adherence to the 21 days regimen. Specific details on how clinicians may adapt dosing for some patients are illustrated in **Error! Reference source not found.**

Therefore, the appropriate calculation of the relative dose intensity for lenalidomide used in maintenance should include the following factors:

1. Drug regimen prescribed, either 10mg or 5mg dose;
2. Frequency of dose, adjusted by the physician on a patient basis;

- Spacing out of intervals between a cycle and another, in addition to the 7-days break, for treatment-related or treatment unrelated reasons.

Figure 31 Examples of dosing, per protocol cycles, dose reductions and spacing of cycle intervals



Figure 1 provides a graphical representation of doses administered in Myeloma XI.

Example 1 shows the pattern of a regular cycle, with 21 daily doses and 7 days interval. This pattern can include cycles at 10mg (1) or both 10mg and 5mg cycles (2).

Clinicians can alternate days on and off at either 10mg, 5mg or both (3,4,5), or in some cases can optimise consumption splitting one (per protocol) cycle over 2 cycles (4). Clinicians can use intervals either prolonging treatment breaks beyond 7 days (7) or shortening treatment days to less than 21 days (6, 10). Doses can also be mixed, for example using dose reduction within each cycle (8) or alternating 10mg and 5mg (9).

1.1 RDI Calculation

We used Myeloma XI treatment prescription data to calculate the RDI for lenalidomide used in maintenance in the real world.

The RDI for lenalidomide used in maintenance is the proportion of the number of packs used over the duration of therapy for a patient, and the number of packs that would be required to cover 100% compliance, per protocol treatment for the same patient for the same duration of treatment.

Using the examples in **Error! Reference source not found.** (1) and (6), the 100% per protocol dose over 3 cycles would be 210mg x3 (630mg); the corresponding dose taken by a patient on a 14 days on, 14 days off cycle would be 140mg x3 (420mg). The RDI is $RDI = 140mg / 210mg = 0.75$.

For example 7, the total follow-up of 84 days includes cycle 1(32 days), cycle 2(44 days) and a fraction of cycle 3 (8 days of treatment).

Overall, the theoretical 100% per protocol cumulative dose for this patient would be 630mg, as above; the actual cumulative dose received is 500mg. Therefore the RDI for these cycles for this patient is 80% despite this patient being on a per protocol dose, fully compliant, with no wastage.

The RDI was calculated per treatment cycle (28 days) and separately for cycles of 10mg/day and 5mg/day. The RDI for each was then incorporated into the model, weighted by the proportion of cycles that were of 10mg and those that were 5mg over the total number of treatment cycles received by the patients in the lenalidomide maintenance arm of the Myeloma XI trial. The detailed methods are provided in the remainder of this Section.

1.2 Lenalidomide prescription data collected during Myeloma XI

Patient-level data were collected for each maintenance treatment cycle in the Myeloma XI study.

The analysis that follows is restricted to data pertinent to patients in the model cohort.

As for clinical data, the model cohort data in Myeloma XI started to be recruited during Protocol 5 and continued throughout Protocol 6. Likewise, drug consumption data was collected under Protocol 5 for some patients, Protocol 6 for others and both protocol 5 and 6 for a third group.

At protocol 6, the methods used to collect drug consumption data were modified as part of the study protocol.

- Protocol 5

Consumption data collected in Myeloma XI, Protocol V5, were:

1. Start and end date for each treatment cycle, collected alongside Myeloma XI.
2. Total accumulated dose prescribed for the cycle (i.e. 210mg for the full per protocol dose, 105mg for a reduced dose, regular treatment, and a variety of other total cumulative doses between 5mg per cycle and 525mg per cycle.

- Protocol 6

Drug dosing data were subjected to a protocol amendment with Protocol 6. The variables collected were:

3. Start and end date for each treatment cycle, collected alongside Myeloma XI

4. Whether the cycle was a 'per protocol' cycle or not
5. Whether the treatment was reduced or delayed or omitted.

For both Protocol 5 and Protocol 6, the start and end of each treatment cycle was the time of the first and last doses taken for each cycle, and not the date on which the treatment was actually dispensed.

No cumulative dose was recorded for Protocol 6.

1.3 Lenalidomide drug consumption data analysis

For each patient, we calculated the following:

1. Total number of cycles of maintenance with lenalidomide prescribed, including all cycles between the start and end of maintenance (first and last prescriptions).
2. Time duration for each cycle: equal to the difference between the start date of one cycle and the start date of the following cycle. A cycle duration was 21 days (time of exposure to maintenance treatment) plus the time 'off treatment' i.e. the remaining time from day 22 to the start of next cycle. For a 'per protocol' cycle, this was 7 days. Therefore a 'per protocol' cycle lasted 28 days (21 days treatment + 7 days break).
3. Number of daily doses prescribed, per cycle, obtained from the total cumulative dose reported for each cycle. Specifically,
4. For Protocol 5, the number of daily doses per cycle was calculated from the total cumulative dose taken per cycle. Given that lenalidomide is packaged in boxes of 21 daily doses (pills), one (protocol compliant) cycle would require one box of lenalidomide; therefore, 21 daily doses were assigned to recorded total cumulative doses equal to 210mg or 105mg (21 days/28 @ 10mg and 21/28 days @ 5mg respectively)
5. For Protocol 6, as no cumulative dose was available, the number of daily doses per cycle was assigned based on whether a cycle was recorded as 'per protocol' (21 daily doses of 10mg), or as 'dose reduced' (21 daily doses of 5mg).
6. For Protocol 5 (but not for protocol 6) total dosages other than 210mg and 105mg were reported in the dataset. After consultation with the Myeloma XI

principal investigator, we concluded that these dosages corresponded to cycles where less than the 21 daily doses were used. We sought clinical opinion to interpret those cumulative doses, which were 'recognisable' as treatment approaches in use in clinical practice, when clinicians need to adapt maintenance on a patient basis. The treatment patterns and doses identified are reported in Table 42 below.

Additional assumptions were used when the total cycle dose was not interpretable. When a total dose per cycle was not recorded as a regular dose, or as any of those reported in Table 42, a per protocol cycle was assumed, and the default initial dosage of 210mg over 21/28 days was assigned.

Treatment-free interval was the duration for each cycle, less the number of daily doses consumed in the cycle. For a per protocol cycle, the treatment-free interval was of 7 days. Additional treatment free days were observed for cycles longer than 28 days.

Table 42. Maintenance with lenalidomide, lenalidomide reported dosages, from Myeloma XI

Treatment dosing	Total lenalidomide dose
██████████	██████████
██████████	██████████
██████████	██████████
██████████	██████████
██████████	██████████
██████████	██████████
██████████	██████████
██████████	██████████

1.4 From dose prescribed to packs dispensed

This step is required for the calculation of the RDI because Revlimid is not dispensed in single daily doses, given that it carries a warning in relation to HCPs handling the capsules.

As two different methods for data collection were used under Protocol 5 and Protocol 6, the calculation of boxes was slightly different for the two datasets. Whilst for Protocol 5, the number of packs was derived from total prescribed dose for each cycle, for Protocol 6 a much simpler approach was used, assuming that all cycles per protocol would be dispensed @ 210mg, all cycles 'reduced' would be dispensed @ 105mg and requiring no assumptions if a cycle was 'delayed' or 'postponed'.

In both cases, however, we applied general principles as follows:

- Once a prescription was given, it was assumed to be always filled and it was assumed that the patient would use all the daily doses. In other words, we considered 100% compliance between prescribed and used doses. This implies no wastage in the conventional sense (i.e. doses prescribed, and either not filled or filled but not used by the patient). From the NHS perspective, a cost is supported at the time when the prescriptions is filled, irrespective as to whether the patient effectively uses the prescribed and filled doses. The only exception to this rule was in the case of a (limited) number of treatment cycles at a reduced number of daily doses in Protocol 5 (i.e. less than 21 daily doses per cycle), where carry-over of daily doses between a cycle and another was plausible. A detailed illustration of how these cycles were handled is provided in Sections below. This only applies to Protocol 5 data, as in Protocol 6 no information on total cumulative dose was collected and all cycles were assumed at full 21 daily doses.
- Lenalidomide can only be prescribed to patients under the Pregnancy prevention programme. This consists of a pregnancy test (or reassessment of the risk of procreation), being conducted before issuing each prescription, and a negative result is required for the prescription to be filled. Therefore, stockpiling behaviours are not possible with lenalidomide. This has a clear implication on the calculation of RDI (and treatment costs) as the number of boxes dispensed cannot exceed the number of cycles prescribed. As in the

point above, there were a (limited) number of cycles where carryover of daily doses between a cycle and another was plausible.

- Patients on the full dose (100% compliant) can still have a treatment break beyond the 28 days cycle. This is generally determined case by case, by the treating physician, based on clinical judgement. Treatment breaks are part of the RDI because they determine total exposure to the drug, as well as the total cost of treatment. As such, treatment breaks are compatible with 100% compliance when this is strictly defined as '21 daily doses in 28 days. This aspect is unique to the use of lenalidomide in maintenance and is supported by evidence in Myeloma XI and corroborated by clinical advice. In essence, in Myeloma XI, the RDI was lower than 100% despite the general assumptions that the large majority of cycles was costed @ 100% compliance.
- The method also accounts for patients who skipped entire cycles or fractions of them, because the RDI uses the theoretical number of cycles at full compliance at the denominator. Therefore, a skipped cycle would contribute as 0mg consumed in the cumulative doses received and as 210mg in the theoretical cumulative doses. For example, a patient who is treated for two cycles (56 days) and skips one cycle will accrue 210mg, instead than the 420mg as per protocol.

Protocol 5

For Protocol 5 only, the total cumulative dose per cycle was analysed to calculate the number of packs received.

For each patient, the number of cycles received at each dose (10mg or 5mg) were accounted for separately. The total number of packs was calculated from the sum of total doses per cycle, for all 10mg or 5mg cycles. The number of packs consumed was calculated using the total cumulative dose per cycle divided by 210 (10mg) or 105 (5mg).

This produced an integer number of packs for the majority of patients, representing full or reduced doses, per protocol cycles, with 100% compliance and no wastage, intended as doses that are not accounted for in the costing of the economic model.

The approach we took assumes the following:

- For patients treated with the 210 mg dose (21 days in 28), a full pack of 21 doses per cycle is dispensed. Similarly, for people prescribed the 105mg total dose, a full pack of 21 5mg daily doses (pills) per cycle would be dispensed.
- Cycles in patients who were treated for one cycle only were assigned one full 21 days 10mg pack, regardless of whether the dose reported was equal or less than 210mg. This was because the initial per protocol cycle would not be likely to be adjusted.

In our methods, we used carry-over in the following cases:

- For cycles who had less than 105mg total doses, we assumed that the daily dose would be 5mg if the total dose was a multiple of 5 and 10mg if the total dose was a multiple of 10. All cycles with doses higher or equal to 110 and multiple of 10 (with the exception of 140, as per Table 1) were assumed cycles based on the 10mg. In both cases above, for cycles that were less than 210mg or 105mg, pack carry-over to the next cycle was assumed as it is reasonable to think that patients would only be given a subsequent prescription limited to the amount required to complete one further cycle. For these patients, the total number of packs was calculated rounded to the next multiple of 105 or 210. This would take account of the pills discarded at treatment discontinuation, which constitutes wastage. This approach also takes into account the doses wasted as patients are switched from the 10mg dose to the 5mg dose or vice-versa, as the two are calculated separately for each patient, based on the number of cycles that the patient had with 10mg and the number of cycles at 5mg.

Carry-over was also used in a specific number of cases, based on the sequences of cycles observed in some patients, as represented in Table 43 below.

Table 43: Sequences of dosages in, Myeloma XI

Treatment dosing	Total lenalidomide dose
Cycles of [REDACTED]	Interpreted as one cycle [REDACTED]

	Assumed equal to full pack if an isolated cycle; carry-over assumed if several one-dose cycles were repeated for the patient (minimum one 21 DDDs box)
Patients who received [REDACTED] at the same dose	[REDACTED]
[REDACTED] consecutively (three patients), after regular cycles at 105mg.	[REDACTED] One patient had 0 wastage (14 cycles) [REDACTED]

Finally, we assumed that all packs dispensed were of 21 daily doses for both 10mg and 5mg. Although lenalidomide 10mg and 5mg is also available in packs of 7 daily doses, we did not consider these packs in the calculation of RDI because they were not documented in the data: doing so would require assumptions on physicians' dispensing behaviours. In some cases, when carry-over applies, the approaches would be equivalent for the purpose of costing. In real practice, it is possible that 7 daily doses packs would be dispensed, limiting wastage. The approach in this analysis therefore should be considered conservative.

Protocol 6

For protocol V6, cycles classified as 'per protocol' were assigned a regular treatment cycle at full dose (210mg, over 21/28 days). Cycles marked 'not per protocol' and classified by the investigator, were handled as follows:

- Delayed: cycle length was obtained from cycle dates. No doses changes were assumed for these cases
- Dose reduction: if a dose reduction was reported and cycle dates were regular, then the full 105mg dose was applied

- Dose omitted: these patients were therefore assigned a 0mg dose, as it is unrealistic to assume that a prescription would be issued if a clinician decided to omit a cycle.
- For cycles that were 'omitted' and were the first and last cycle for a patient (i.e. patients who only had one cycle) a 210mg dose was imputed for conservative reasons, as we are assuming that patients were at least started on treatment. Importantly, these patients were not included in the safety dataset, as they were unlikely to have been initiated on treatment, however a pack of lenalidomide was still costed (100% wastage), to exclude the possibility of underestimation.
- For patients with consecutive omitted cycles, the first cycle was imputed as full dose; remaining cycles were imputed as a 0 dose, as it is unrealistic to assume that clinicians would continue to prescribe for patients temporarily taken off treatment.

Six cycles were classified as 'not per protocol' with no further information therefore they were assigned a 'per protocol' dose (210mg 21/28).

As there was no dose information from Protocol 6 data, no assumptions were made on dose adjustments and alternative treatment patterns; therefore packs calculated for all Protocol 6 cycles were assumed dispensed in full dose (210 or 105mg) and filled for all prescriptions. Essentially, for all Protocol 6 cycles, the RDI was entirely driven by treatment intervals between a cycle and another and treatment breaks.

1.1. Missing data

Missed doses were reported for a very small number of patients; therefore no formal imputation method was used but the most likely or the most conservative dose per cycle was imputed instead. These cases are summarized here (Table 44) for completeness and transparency.

Table 44. Methods used to handle missing data

Missing data	Assumption	Impact
One patient with 3 consecutive doses missing	210mg as per regular treatment was assumed	Highest possible dose; conservative cost

Patients with missing doses during intermediate cycles	the same dose as for the adjacent cycles applied	Plausible cost assumed instead than no cost
One patient had a different dose prior to and after the missing value	highest dose was imputed	
Cycle start date missing	a regular cycle was assumed (21/28 days),	Conservatively (n=41)
Single cycle and no start date	start date was the first cycle in the data set.	

1.5 Calculation of RDI

The RDI for lenalidomide maintenance was calculated separately for the 10mg and the 5mg cycles.

Step 1. Effective volume of doses consumed

Based on doses per cycle for all cycles in the dataset, the (effective) total number of packs was calculated for each patient, separately for the 210mg and 105mg cycles.

For each patient, the (effective) total number of cycles was also calculated, as the difference between first and last date for treatment with 10mg and first and last date for treatment with 5mg. For example, if a patient had 4 cycles of treatment at 10mg and 4 cycles of treatment at 5mg, then the '10mg treatment period' would be the total of time spent on cycle 1 to 4 (between start of cycle 1 and start of cycle 5), and the 5mg period would be the total of time spent in cycle 5 to 8.

As the purpose of this analysis was to incorporate the average cost of therapy in the model as a proportion of the maximum cost (100% regular cycles) for the two doses separately (100% dosing at @ 210mg or @ 105mg), we used a pragmatic approach to account for the duration of mixed cycles (i.e. alternate days at 10mg and at 5mg) counting them as 10mg cycles. This was because there were very few mixed cycles and furthermore, accounting them as two separate cycles would constitute double counting and therefore would dilute the RDI.

Step 2. Maximum volume of doses that could be prescribed (full compliant cumulative dose)

A cycle at 100% compliance and 100% regular, with no treatment break, corresponds to 210mg every 28 days, or 105 every 28 days.

Using the time spent on 10mg and on 5mg (as illustrated in the Paragraph XX above), we calculated the number of cycles that the patient would receive if treated with no interval breaks, based on regular 210mg cycles in 28 days, or regular 105mg in 28 days cycles, separately for each dose.

For example, if the treatment duration for a patient was 112 days (with one of the two doses), the 'per protocol' number of cycles that this patient would receive would be 4 ($=112/28$). Therefore, when treated according to 'protocol' and with no treatment breaks, this patient would receive 4 cycles.

The RDI was calculated dividing the number of packs per patient used in real practice by the number of packs that the same patient would require if the same patient were fully compliant with the 210mg (or 105mg) dosage and regular 21/28 days use.

For example, if a patient received 3 packs (@ 10mg) during the 112 days, the RDI would be equal to 75% ($=3/4$).

As the model is set up using a cycle of 28 days, the RDI was used as a correction factor to estimate the total treatment cost for that patient. We applied the RDI in the model separately for 10mg and 5mg, as a weighted average of RDI and proportion of cycles on 10mg and on 5mg from the Myeloma XI data.

So, applying the cost of one pack per cycle, weighted by the RDI (75%) would generate 3 packs over 4 cycles and therefore the cost of lenalidomide in the model would be correctly estimated.

2 Results

The Myeloma XI dataset included data for ██████ in total, ██████ of which had dosing information (Table 45). Of the remaining ██████ cycles, ██████ had a total dose multiple of 5 or 10 but not otherwise adjudicated; ██████ cycles had missing (dose or date) or non-classifiable information. (Table 45 and 46 below)

All patients in the dataset had at least one 10mg cycle; about [REDACTED] patients) had a dose reduction to 5mg at least once. Most cycles were either 'regular' ([REDACTED]) or followed a recognizable 'pattern' (Table below). Slightly less than [REDACTED] of treatment cycles were at 5mg.

Table 45. Total numbers of patients and cycles, by dose of lenalidomide

Result	10mg dose	5mg dose
Number of patients with 10mg/5mg cycles	[REDACTED]	[REDACTED]
Total number of cycles (%) including dosage of 10mg/5mg	[REDACTED]	[REDACTED]

Table 46. Adjudicated cycle doses, recognizable patterns

Treatment dose (by cycle)	Total dose	Number of cycles	
		Protocol 5	Protocol 6
[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]

On average, patients had [REDACTED] at 10mg. The patients that had a reduced dose of 5mg had on average [REDACTED] at the reduced dose. Most patients had treatment delays of an average [REDACTED], both with the 10mg and 5mg dose, with a mean cycle duration [REDACTED].

The mean RDI was [REDACTED] (Table 47)

Table 47: Results of the RDI analysis

Result	10mg dose	5mg dose
Number of patients with 10mg/5mg cycles	████████	████████
Total number of cycles (%) including dosage of 10mg/5mg	████████	████████
RDI (SD)	████████	████████
Mean number of cycles (SD) per patient	████████	████████
Mean cycle length (including 7 days treatment-free period) (SD, min-max)	████████	████████

3 Conclusion

Data from myeloma XI show that in real clinical practice, the dosing of maintenance with lenalidomide is adapted to patients by means of the extensive use of treatment delays and dose reduction. In practice, this approach supports a cost-effective use of lenalidomide.

3.1 Wastage

Due to the costing methodology used in the model, all drugs packs prescribed have been costed as part of NHS perspective. From this viewpoint, the costing methods accounted for all drug that is wasted as part of delivery of treatment.

Because there is no information regarding whether patients are compliant in everyday use is impossible to ascertain; nevertheless, this would not be a factor in the costing of treatment from the perspective of the UK NHS.

The reduced RDI is the result of spaced intervals between a cycle and another and all daily doses dispensed to patients are fix dose tablets. Therefore, we conclude that all possible sources of wastage have been accounted for in the model.

Following the requests of the committee further analyses have been undertaken. Table 1 details the change in company base case ICER when applying these preferences.

Using the assumed committee preferred scenarios lenalidomide maintenance treatment is a cost-effective use of NHS resources.

Table 1: Committee preferred assumptions and impact on the cost-effectiveness estimates

Scenario	Committee preferred assumption	Assumptions used in scenario	ICER (change vs base case)
Base case	Not applicable	Company base case analysis (revised at technical engagement)	████████
Base case (updated)	Not applicable	Company base case analysis (revised at technical engagement) ██ ██	████████
1	Clinical-effectiveness evidence from the CALGB-100104 and GIMEMA trials of lenalidomide maintenance treatment	No change required to company base case (Section 1)	████████
2a	Survival estimates in the economic model based on Myeloma XI data, with CALGB 100104 data (adjusted to reflect the Myeloma XI population as closely as possible, and conditional on the underlying survival of patients in Myeloma XI) used to help longer-term extrapolation	Myeloma XI data used until 60 months with adjusted CALGB data (using propensity score weighting) from 60 months to inform parametric survival analysis (Section 3)	████████

		Long term extrapolation distribution for OS: Weibull	
2b		Myeloma XI data used until 60 months with adjusted CALGB data (using propensity score weighting) from 60 months to inform parametric survival analysis (Section 3) Long term extrapolation distribution for OS: Gamma	████████
3	The company's justification for using the rank-preserving structural-failure time model over other methods to adjust for treatment switching in CALGB 100104	No change required to company base case (Section 7)	████████
4	Waning of the treatment effect of lenalidomide	Treatment waning applied after 10 years (Section 4)	████████
5a	5% to 10% of people having a second autologous stem cell transplant, and a range of assumptions for the proportions of people having other subsequent therapies for the purposes of estimating subsequent treatment costs in the model	5% of patients receiving a second autologous stem cell transplant (Section 5)	████████
5b		10% of patients receiving a second autologous stem cell transplant (Section 5)	████████
6	Detailed methods for how the relative dose intensity was calculated using Myeloma XI data	No change required to company base case (Section 8)	Not applicable

7	A model scenario reflecting a 1-day to 28-day lenalidomide treatment regimen	Treatment given on every day of 28 day cycle (Section 6)	████████
2b, 4, 5a combined	Assumed committee preferred scenario (with treatment waning and 5% of people having second ASCT)		████████
2b, 4, 5b combined	Assumed committee preferred scenario (with treatment waning and with 10% of people having second ASCT)		████████
2b, 4, 5a, 7 combined	Assumed committee preferred scenario (with treatment waning, 5% of people having second ASCT and treatment given on every day of 28 day cycle)		████████
2b, 4, 5b, 7 combined	Assumed committee preferred scenario (with treatment waning, 10% of people having second ASCT and treatment given on every day of 28 day cycle)		████████
2b, 5a and alternative subsequent therapy assumptions	New company base case (5% of people having second ASCT with 10% of people in the observation arm receiving lenalidomide + dexamethasone in second line)		████████

Lenalidomide for the maintenance treatment of multiple myeloma after autologous stem cell transplantation [ID475]

Consultation on the appraisal consultation document – deadline for comments 5pm on Tuesday 29 September email: NICE DOCS

	<p>Please read the checklist for submitting comments at the end of this form. We cannot accept forms that are not filled in correctly.</p> <p>The Appraisal Committee is interested in receiving comments on the following:</p> <ul style="list-style-type: none"> • 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? <p>NICE is committed to promoting equality of opportunity, eliminating unlawful discrimination and fostering good relations between people with particular protected characteristics and others. Please let us know if you think that the preliminary recommendations may need changing in order to meet these aims. In particular, please tell us if the preliminary recommendations:</p> <ul style="list-style-type: none"> • could have a different impact on people protected by the equality legislation than on the wider population, for example by making it more difficult in practice for a specific group to access the technology; • could have any adverse impact on people with a particular disability or disabilities. <p>Please provide any relevant information or data you have regarding such impacts and how they could be avoided or reduced.</p>
<p>Organisation name – Stakeholder or respondent (if you are responding as an individual rather than a registered stakeholder please leave blank):</p>	<p>Myeloma UK</p>
<p>Disclosure Please disclose any past or current, direct or indirect links to, or funding from, the tobacco industry.</p>	<p>N/A</p>
<p>Name of commentator person completing form:</p>	<p>████████████████████</p>
<p>Comment number</p>	<p>Comments</p>

Please return to: **NICE DOCS**

Lenalidomide for the maintenance treatment of multiple myeloma after autologous stem cell transplantation [ID475]

Consultation on the appraisal consultation document – deadline for comments 5pm on Tuesday 29 September email: NICE DOCS

	<p>Insert each comment in a new row. Do not paste other tables into this table, because your comments could get lost – type directly into this table.</p>
	<p>Myeloma UK Response to Lenalidomide Maintenance ACD Consultation</p> <p>Myeloma UK is very disappointed that lenalidomide maintenance for newly diagnosed multiple myeloma patients who are eligible for high-dose therapy and stem cell transplantation (HDT-SCT) has not been approved for routine commissioning.</p> <p>Given the scale of new modelling and evidence requested from the company, and the need for this to be informed by expert clinical opinion and relevant patient insight, we ask that clinicians and Myeloma UK are invited to take part in the second Committee meeting.</p>
1	<p><u>Has all of the relevant evidence been taken into account?</u></p> <p>No. We note the requests in the Appraisal Consultation Document (ACD) for new modelling and evidence to be presented by the company to address areas of uncertainty identified by the Committee.</p> <p>We welcome the following findings in the ACD based on the evidence presented:</p> <ul style="list-style-type: none"> - lenalidomide is the only potential option for maintenance treatment for multiple myeloma after an autologous stem cell transplant. This underlines the high level of unmet need at this point in the pathway - Lenalidomide maintenance would be the standard treatment for newly diagnosed myeloma patients who are eligible for an HD-SCT - The demonstrable need for and benefits of this treatment are strengthened by the fact that first remission is often the best remission, as patients are at their fittest and can maintain their highest quality of life and, unfortunately, the numbers of patients able to access further lines of treatment diminishes - The dosing schedule used in clinical practice would be 21 out of 28 days. We understand why, for completeness, the Committee has asked for further data on the 28 days schedule since it reflects the marketing authorisation. However, we note that the unanimous support of clinicians for the 21-day schedule was strengthened by NHS England’s confirmation that it would commission a 21-day schedule. There can be no practical doubt that the 21-day schedule is the one that would be used in clinical practice - Lenalidomide is an effective maintenance treatment for people who have had an autologous stem cell transplant. We emphasise that lenalidomide is not just an effective treatment, it is highly effective; the magnitude of additional clinical benefit it delivers is rarely seen in the treatment of myeloma, which remains an incurable cancer - Myeloma XI should be used in relation to relative dose intensity
2	<p><u>Are the summaries of clinical and cost effectiveness reasonable interpretations of the evidence?</u></p> <p>Given the scale of further information requested from the company at this point in the appraisal process, it is not possible to come to a definitive view on whether the summaries are reasonable.</p> <p>We understand why the Committee has requested further evidence from the company to enable further scrutiny and understanding of uncertainty, for example to enable assumptions about the effects of subsequent therapies to be fully explored.</p> <p>However, this work must be seen in the context of the continuously and rapidly evolving myeloma treatment pathway. (Impacted by routine commissioning, CDF approvals and also by recent approvals of alternative oral treatments during COVID- 19.)</p>

Please return to: **NICE DOCS**

Lenalidomide for the maintenance treatment of multiple myeloma after autologous stem cell transplantation [ID475]

Consultation on the appraisal consultation document – deadline for comments 5pm on Tuesday 29 September email: NICE DOCS

	<p>In this context, it is inevitable that trial design and subsequent data will be unable to fully reflect UK clinical practice at any given time. The Committee has also acknowledged the challenges in modelling the costs of subsequent therapies given that Cancer Drugs Fund (CDF) approved therapies should not be used in economic modelling.</p> <p>A proportionate approach to the inevitable uncertainty which arises from the welcome development of the myeloma treatment pathway is therefore vital. Not to do this would be unreasonable.</p> <p>We note that the issue of treatment sequencing and the inability to reflect real world use of CDF funded drugs is impacting all myeloma appraisals and that current procedures may not be sustainable. There is a danger that, in the absence of new approaches to dealing with this complexity, NICE decisions will become increasingly disconnected from established real world practice and therefore difficult to present as meaningfully “reasonable”.</p>
3	<p><u>Are the provisional recommendations sound and a suitable basis for guidance to the NHS?</u></p> <p>No. Notwithstanding the further information that has been requested from the company, we do not accept that a decision not to recommend lenalidomide maintenance post HDT-SCT is sound.</p> <p>There is a clear and significant unmet need for lenalidomide maintenance post HDT-SCT for the treatment of myeloma which is standard best practice internationally.</p> <p>In Myeloma XI we have a mature and comprehensive UK data set which, alongside the CALGB trial, provides compelling evidence of the significant benefit delivered by lenalidomide maintenance post HDT-SCT.</p> <p>This is an incredibly effective, life-extending, safe treatment, administered orally which, in the current COVID-19 environment, delivers further benefits to patients, families and to the NHS.</p> <p>It would be deeply concerning if, with this rich data, the company, Evidence Review Group (ERG) and NICE were not collectively able to resolve sufficiently these areas of uncertainty to enable a positive recommendation.</p> <p>In the next stage of the appraisal it is therefore essential that the company provide the additional evidence the Committee has requested; and that, in interpreting the evidence, the Committee takes into account the inevitable limitations of both the data, and of appraisal methods and processes in providing a complete and certain picture of clinical benefit and practice.</p> <p>Not to do so runs the risk of arriving at a decision which can be justified methodologically, but which is not reasonable when it is applied in a real-world context.</p>
4	<p>Myeloma UK have sought comments directly from patients about the impact of the draft no recommendation and these have been included Appendix A of our response to the ACD.</p>
5	
6	

Insert extra rows as needed

Checklist for submitting comments

- Use this comment form and submit it as a Word document (not a PDF).
- Complete the disclosure about links with, or funding from, the tobacco industry.
- Combine all comments from your organisation into 1 response. We cannot accept more than 1 set of comments from each organisation.

Please return to: **NICE DOCS**

Lenalidomide for the maintenance treatment of multiple myeloma after autologous stem cell transplantation [ID475]

Consultation on the appraisal consultation document – deadline for comments 5pm on Tuesday 29 September email: NICE DOCS

- Do not paste other tables into this table – type directly into the table.
- Please underline all confidential information, and separately highlight information that is submitted under **'commercial in confidence' in turquoise** and all information submitted under **'academic in confidence' in yellow**. If confidential information is submitted, please also send a 2nd version of your comment with that information replaced with the following text: 'academic / commercial in confidence information removed'. See the Guide to the processes of technology appraisal (section 3.1.23 to 3.1.29) for more information.
- Do not include medical information about yourself or another person from which you or the person could be identified.
- Do not use abbreviations
- Do not include attachments such as research articles, letters or leaflets. For copyright reasons, we will have to return comments forms that have attachments without reading them. You can resubmit your comments form without attachments, it must send it by the deadline.
- If you have received agreement from NICE to submit additional evidence with your comments on the appraisal consultation document, please submit these separately.

Note: We reserve the right to summarise and edit comments received during consultations, or not to publish them at all, if we consider the comments are too long, or publication would be unlawful or otherwise inappropriate.

Comments received during our consultations are published in the interests of openness and transparency, and to promote understanding of how recommendations are developed. The comments are published as a record of the comments we received, and are not endorsed by NICE, its officers or advisory committees.

Patient Feedback to Lenalidomide Maintenance ACD

Myeloma UK have sought feedback from patients regarding the draft no recommendation from NICE in the lenalidomide maintenance ACD. Below is a summary of responses gathered from email/social media channels. (All responses have been anonymised)

Patient A

I am extremely disappointed in the NICE/NHS decision to issue a draft "no" to lenalidomide maintenance.

This is part of a real whammy for that cohort of myeloma patients who were diagnosed too late to be part of the UK myeloma XI trial, through which so many of us accessed lenalidomide as maintenance. Maintenance that some of us are still on.

We who have had SCTs since 2017 will have shorter remission times and shorter overall survival times than patients diagnosed earlier.

Our lives are currently and in the near future restricted due to COVID and our immune deficiencies. We all want to survive long enough to have quality of life again, but the recent NICE decisions appear to be sacrificing us.

I want to live long enough to become a grandparent, to see my children married and in homes of their own, to have a period of active retirement with my husband, to support him caring for his elderly mother. Currently I am well, I appear fraudulently ill. I have no pain, no physical manifestation of myeloma and I want to remain in that state for as long as possible.

I want to see other myeloma patients having long periods of remission. I don't want to keep reading of, and meeting, patients who have short remission times because they are denied drugs that would be available to them in other first world countries. It is painful bearing witness to other patients suffering when I know that statistically they would have had longer periods of good health if maintenance had been available to them.

I also worry about how the NHS is going to recruit and retain good myeloma specialist Doctors and nurses if those professionals are continually faced with being unable to offer patients the treatments that they know they need.

It makes me feel angry, sad but mostly disenfranchised.

Patient B

I'm a married ■-year-old myeloma patient and mum to two gorgeous ■ aged ■ and ■. I was diagnosed ■, aged ■. My SCT was delayed due to COVID but took place in July 2020. I have been told I have high risk myeloma and have to come to terms with the fact that there is no cure, though treatable. However, while this may be the case I am much younger than the 'normal' sufferer added to the fact I am high risk.

I NEED to do everything I can to extend my life to bring my boys up as close to adulthood as possible, My boys are happy boys, who work hard in school, love sport and as already said happy. This disease isn't just mine but that of my family (extended as well). My only priority in life is my boys, and I cannot allow this disease to deprive them of the childhood I have always dreamed for them. My aim has always been to bring them up knowing right from wrong, to work hard and to be happy. While this disease is non-curable, I know that at some point I will have to destroy their happy world by leaving them. To know there is a drug out there that could extend my life and my time with my children to get them closer to adulthood is so encouraging. However, to know that I currently can't access this drug is devastating.

I can only plead with NICE to approve this drug that has been proven to help people like me.

Thank you for listening.

Patient C

I am a patient who was diagnosed with multiple myeloma in [REDACTED] I underwent VDT and was ready for an SCT in April 2020 but this was delayed due to COVID 19. I was prescribed thalidomide as maintenance treatment but my Paraprotein levels started to rise so I received a further 1.5 cycles of VDT which brought the levels down again and I underwent an SCT in August this year.

I note the Committee accepts that Lenalidomide as a maintenance treatment after SCT improves PFS and OS. The uncertainty and reason for the draft no seems to be about the cost benefits of treatment and other issues surrounding the models used, and the use of existing data by the pharmaceutical company in support of their application for approval.

As a patient all I can do is outline my views. I am of working age. When I have recovered from my SCT I intend to return to work. As myeloma is incurable, I recognise, as all patients do, that myeloma will return some day. What we don't know is when it will return so it is like having the sword of Damocles hanging above you 24/7. I cannot underestimate the effect of this on patients.

I have only received VDT to date, which was tough going and prevented me from working. I worry that when I relapse, (assuming my SCT has worked, I won't know until after a bone marrow biopsy in November) I will be unable to work again. Many patients in a similar position will have to apply for state benefits once they relapse, and there is also the cost to the NHS of treatment for active disease which the data you have considered indicates will be incurred sooner without maintenance treatment. I do hope the relevant data will be provided to you about the possible costs.

The evidence you have already reviewed strongly supports the proposition that the first remission after SCT is the longest and provides the best quality of life for most patients. Lenalidomide will provide most patients, if it's available as maintenance treatment after SCT on the NHS, with:

a) Longer quality of life, which is so important to us.

B) More time in work for those of working age. For many of us continuing to work and be able to financially support ourselves and our families for as long as possible is very important. It also benefits the economy, and reduces reliance on state benefits.

C) Longer time off intensive treatment before the disease becomes active once more. Intensive treatment believe me is usually tough for most of us.

The data you have reviewed indicates the vast majority of patients would take Lenalidomide as a maintenance treatment if it was available on the NHS. This treatment is available privately in the UK. It is available in Europe. It should be available on the NHS to all myeloma patients who have had an SCT, there should be no arbitrary cut off date, and whose PFS and OS it is considered clinically, taking into account the risks to the patient from receiving Lenalidomide, will be extended by receiving maintenance treatment.

Patient D

“I'm thoroughly disappointed in the NICE decision not to approve Lenalidomide for maintenance. I was really holding out for a 'yes'. Its been proven that Lenalidomide as maintenance following a SCT offers an average of two additional years of remission before further treatment is required. That's two more years of a full life without undertaking what I understand to be more expensive treatments or hospitalisation. I'm a very fit ■ year old and I want to use any extra time I can get to spend with my family and serve my community. I can do that when I'm in remission, but not when I'm having treatment.”

Patient E

“When I had my SCT in ■■■■■ my consultant said he would ideally have liked to have put me on Lenalidomide as maintenance but he couldn't. I have a rare mutation of Myeloma and NHL and he wanted to give me as long a remission as possible as my treatment options are more limited than some. If my consultant thinks it's worth having and is annoyed that he can't prescribe it then it seems crazy that NICE won't agree; he can't be the only consultant who is of that opinion.”

Patient F

“It's really disappointing, especially when it's been proven to extend remission, I'm a ■ year old mother of ■ so the longest remission the better so I can be treatment free and live a somewhat normal life.”

Patient H

“Disappointed to hear this- my husband has been on lenalidomide maintenance, post SCT, via the myeloma XI trial, for the last 4.5 years. Don't want to tempt fate but he's doing so well.... can't understand why access is restricted following the great results from this trial - such a shame others can't benefit from this drug.”

Patient I

An extremely disappointing decision. My mum has relapsed 11 months post SCT and this based on trial data would have been progression free for much longer had she had the option of Lenalidomide as maintenance.

Progression for her is not just a rise in paraprotein, she now has progression of multiple bone lesions with vertebral fractures and a prophylactic IM nail in her right femur due to high risk of fracture. Her chronic pain, anxiety and depression have all worsened as well as her quality of life. To know that there is a drug that could have prevented this for years but is not licensed in the U.K. but available elsewhere is heart breaking.

I understand funding is an issue. However she now has to have second line treatment which itself is not cheap; weekly hospital appointments for months, increasing burden on primary and secondary care and my dad as her primary carer who has had to leave work.

I'd ask NICE to please reconsider and let someone else have the chance to lead a normal life just a little longer."

Comments on the ACD received from the public through the NICE Website

Lenalidomide for the maintenance treatment of multiple myeloma after autologous stem cell transplantation [ID475]

Name	██████████
Role	
Other role	
Organisation	
Location	
Conflict	
Notes	
Comments on the ACD:	
<p>I am a patient who was diagnosed with multiple myeloma in ██████████. I underwent VDT and was ready for an SCT in April 2020 but this was delayed due to Covid 19. I was prescribed thalidomide as maintenance treatment but my Paraprotein levels started to rise so I received a further 1.5 cycles of VDT which brought the levels down again and I underwent an SCT in August this year.</p> <p>I note the Committee accepts that Lenalidomide as a maintenance treatment after SCT improves PFS and OS. The uncertainty and reason for the draft no seems to be about the cost benefits of treatment and other issues surrounding the models used, and the use of existing data by the pharmaceutical company in support of their application for approval.</p> <p>As a patient all I can do is outline my views.</p> <p>I am of working age. When I have recovered from my SCT I intend to return to work. As myeloma is incurable, I recognise, as all patients do, that myeloma will return some day. What we don't know is when it will return so it is like having the sword of Damocles hanging above you 24/7. I cannot underestimate the effect of this on patients.</p> <p>I have only received VDT to date, which was tough going and prevented me from working. I worry that when I relapse, (assuming my SCT has worked, I won't know until after a bone marrow biopsy in November) I will be unable to work again. Many patients in a similar position will have to apply for state benefits once they relapse, and there is also the cost to the NHS of treatment for active disease which the data you have considered indicates will be incurred sooner without maintenance treatment. I do hope the relevant data will be provided to you about the possible costs.</p> <p>The evidence you have already reviewed strongly supports the proposition that the first remission after SCT is the longest and provides the best quality of life for most patients.</p> <p>Lenalidomide will provide most patients, if it's available as maintenance treatment after SCT on the NHS, with:</p> <p>A) Longer quality of life, which is so important to us.</p> <p>B) More time in work for those of working age. For many of us continuing to work and be able to financially support ourselves and our families for as long as possible is very important. It also benefits the economy, and reduces reliance on state benefits.</p> <p>C) Longer time off intensive treatment before the disease becomes active once more. Intensive treatment believe me is usually tough for most of us.</p>	

The data you have reviewed indicates the vast majority of patients would take Lenalidomide as a maintenance treatment if it was available on the NHS. This treatment is available privately in the UK. It is available in Europe. It should be available on the NHS to all myeloma patients who have had an SCT, there should be no arbitrary cut off date, and whose PFS and OS it is considered clinically, taking into account the risks to the patient from receiving Lenalidomide, will be extended by receiving maintenance treatment.

Name	██████████
Role	
Other role	
Organisation	
Location	
Conflict	
Notes	
Comments on the ACD:	
<p>Has all of the relevant evidence been taken into account? Yes</p> <p>Are the summaries of clinical and and cost effectiveness reasonable interpretations of the evidence?</p> <p>No - see detailed comments regarding the intrinsic flaw in the argument that the data does not include current England-specific next treatment comparisons. There is a "catch 22" of wanting long term survival data yet wanting that data to reflect current treatment options. The committee needs to decide which argument it is going to use to decline a drug but cannot use opposing reasons in different appraisals.</p> <p>Are the recommendations sound and a suitable basis for guidance to the NHS?</p> <p>No. The recommendations fail to recognise the most significant step forward in the management of myeloma in this country in the last decade. The UK is one of the only developed countries in the world to not enable access to a clearly highly effective maintenance option. There are no robust arguments against approval. I implore the committee to rethink its decision in the interests of patient care and clinical outcomes rather than use fatuous arguments to block access to a highly effective drug.</p> <p>committee-discussion</p> <p>The dosing schedule that would be used in clinical practice is different to that in the marketing authorisation</p> <p>Lenalidomide 10 mg days 1-21 is the dose that is used across the UK with many patients still currently receiving it in the Myeloma XI clinical trial and used in the private sector in the UK for those who have access to this. There is substantial familiarity with its use across the country and of the management of possible toxicity.</p> <p>The company should have presented evidence from other trials of lenalidomide maintenance treatment</p>	

Meta-analysis of all the trials mentioned using different dosing schedules of lenalidomide show a clear advantage of lenalidomide over no maintenance.

The safety profile of lenalidomide as a maintenance treatment compared with monitoring alone is likely to be acceptable

Data from the Myeloma XI study (Jackson et al, Lancet Oncology 2019) indicates a manageable toxicity profile with lenalidomide maintenance. There is significant familiarity in the UK managing these now and therefore it is not the case that there is clinical uncertainty about the safety profile of the drug amongst those who treat people with myeloma.

The company's model structure does not allow assumptions about subsequent treatments to be explored

The ERG argument is intrinsically flawed. The Myeloma XI trial is the largest of its kind and is directly relevant to UK practice. In a rapidly evolving field such as myeloma it cannot be expected that next line treatment options will remain static over time as newer treatments are developed. It is integral to the presentation of long term survival data that patients have to have been treated several years ago in order to generate robust survival modelling. If this argument were followed to its logical conclusion, no drug could ever be evaluated unless the survival associated with that drug were measured in a matter of months because treatment options would have changed over that time. Patients should not be penalised because the modelling preferred by the ERG cannot take account of this.

Survival extrapolations should use Myeloma XI data as the main source of evidence but could be supplemented with CALGB 100104 data

There is no clinical rationale to suggest that 28 day continuous dosing is any more or less effective than 21/28 dosing. That is mere speculation and should not be used in an appraisal where robust evidence should be examined

Costs of subsequent treatments are highly uncertain so scenarios should be presented

Given the range of therapies available on the CDF in myeloma, it is clinically inappropriate not to include these options as they do reflect current standard of care in myeloma with proven clinical efficacy and good UK clinical experience. This exclusion needs urgent revision as makes assumptions about next treatment irrelevant to the population in question treated in England and any modelling entirely fictitious.

Carfilzomib is available as a second line treatment option in myeloma - see TA457.

Name	
Role	
Other role	
Organisation	
Location	
Conflict	
Notes	
Comments on the ACD:	

As someone whose father has recently been diagnosed with Active Multiple Myeloma and will need Lenalidomide as maintenance treatment in order to live for longer, I don't believe that these summaries are "reasonable". If "reasonable" is defined as: "having sound judgement; fair and sensible" I would absolutely argue that these summaries are unreasonable. They couldn't be understood as fair or sensible by anyone who is really thinking about the consequences. The recommendations state that the science shows that Lenalidomide maintenance treatment prolongs life and at a high quality. There isn't another drug or any other treatment available that does this, so the only alternative is to die faster. Thousands of people would live quality lives for years longer than they are currently able to with this drug available on the NHS. If there wasn't a pandemic, people would be able to access this drug as part of a "trial" - my father would be one of them. Therefore, to not approve this drug at a time when there is a pandemic on seems inhumane to me. Particularly given that it is taken as a capsule at home, which means it doesn't require a hospital visit. My understanding is that this drug is prescribed as maintenance treatment as a matter of practise if a patient is being treated privately, rather than through the NHS. I don't think it is "reasonable" that someone who has a lot of money could afford to keep themselves alive for longer than someone who doesn't have a lot of money. If the cost-effectiveness needs to be understood more clearly from additional modelling, I think this should be done asap to turn this decision into a "yes". If Celgene / BMS are charging more than the NHS modelling can approve, they should offer this treatment to the NHS at a lower fee, or NICE and the NHS should find a way to consider this through a model that allows for a higher threshold. There is something wrong here when a treatment that works so well is being withheld from people who undeniably need it and can't get it any other way. I understand that Celgene / BMS have stakeholders to answer to and that this drug is a "blockbuster" for them, but they will still make profits, while allowing thousands of families to know their loved ones for longer. I honestly believe that to say "no" to approval of this drug would go against the Hippocratic Oath which surely used to be at the heart of the pharmaceutical industry as well as healthcare more generally. If it is coming out of patent soon anyway, why allow thousands of people to suffer and die prematurely, just to make more money in the meantime? How can anyone in a decision-making position allow that to happen? If NICE's mission statement is to: "Protect people's money, ensure their safety and improve their experiences", they must be able to find a way to approve Lenalidomide for maintenance on the NHS. Otherwise will be forced into paying life-changing prices, selling their homes, to try and live for longer. This isn't "safe" for mental or physical health and, rather than "improving experiences", would most definitely negatively affect their experiences and those of their wider circle of friends and family too. The science says that this drug really helps people who are in desperate need, the only potential problem is the price. Please, please reconsider this draft no and find a way through the corporate red tape and stakeholder interests to remember that real people's lives are at stake here. My dad is one of them and I beg you to reconsider.

Lenalidomide for the maintenance treatment of newly diagnosed multiple myeloma after autologous stem cell transplantation [ID475]

ERG Review of Company's Response to ACD

30 October 2020

Produced by	Peninsula Technology Assessment Group (PenTAG) University of Exeter Medical School South Cloisters St Luke's Campus Heavitree Road Exeter EX1 2LU
Authors	Caroline Farmer¹ Emma Knowles² Helen Coelho¹ Justin Matthews¹ Sophie Robinson¹ Naomi Shaw¹ Claudius Rudin³ Jenny Bird⁴ Simone Critchlow² Louise Crathorne¹ G.J. Melendez-Torres¹
Correspondence to	Caroline Farmer Email: C.Farmer@exeter.ac.uk

¹ Peninsula Technology Assessment Group (PenTAG), University of Exeter Medical School, Exeter

² Delta Hat Ltd, Nottingham UK

³ Dorset County Hospital, UK

⁴ University Hospitals Bristol NHS Foundation Trust, UK

1. SUMMARY

In its appraisal consultation document (ACD), the committee raised a number of concerns regarding the evidence base underlying the appraisal of lenalidomide for the treatment of multiple myeloma after autologous stem cell transplantation (ASCT). In this document, the evidence review group (ERG) reviewed additional evidence provided by the company to address these concerns, in advance of a second committee meeting.

2. KEY ISSUES

Issue 1: The company should have presented evidence from other trials of lenalidomide maintenance treatment

The ERG and company disagreed about the status of information from the GIMEMA trial (Palumbo *et al.* 2014).¹

Both the ERG and company agreed that the ‘primary analysis’ of Palumbo *et al.* 2014 was not specific to the subpopulation of patients who received an ASCT. However, the ERG had argued in its report that data could be extracted for the post-ASCT subgroup from figures in Palumbo *et al.*, and obtained estimates for progression-free survival (PFS) and overall survival (OS). At technical engagement (TE) the ERG further reported that an analysis given in McCarthy *et al.* 2017 using individual patient data (IPD) provided a direct estimate in the ASCT subgroup of HR=0.50 (95%CI 0.31 to 0.80) for PFS and 0.72 (95%CI 0.37 to 1.38) for OS.

The company argued that the design of the GIMEMA study invalidates its use in the decision problem. GIMEMA was designed with 2x2 factorial randomisation to ASCT or MPR, and to maintenance or not, at enrolment. The company pointed out that some patients randomised to ASCT did not subsequently receive it or failed it. The ERG notes that Figure 1 of Palumbo *et al.* (2014)¹ indicates these numbers are small (141 patients were randomised to ASCT of whom six discontinued), and understands those discontinuing would have been excluded from the estimate in McCarthy *et al.* (2017),² which analysed known ASCT patients. The company also argued that “there is a possibility that randomisation was broken following exclusion of patients who failed induction and ASCT” (ACD addendum, p.20), but the ERG believed this would not happen because “randomisation to maintenance ... was concealed until the end of the MPR or ASCT phase when patients were assessed for maintenance therapy” (ACD addendum, p.9).

Regardless of the above arguments, the ERG accepted that the available data for the ASCT subgroup in GIMEMA was limited (e.g. no Kaplan-Meier curve). The result from GIMEMA for the post-ASCT subgroup reported by McCarthy *et al.* (2017)² suggested a benefit to lenalidomide maintenance, with the estimates (given above) broadly in line with the Myeloma XI results [REDACTED].

In its report, the ERG queried the rationale for excluding the CALGB 100104 trial from the company’s clinical effectiveness review in the CS. At technical engagement the company presented pooled evidence from the CALGB 100104 trial and the Myeloma XI trial, without

presenting full clinical efficacy evidence from CALGB 100104, and without rationale for the change in position. In its response to the ACD, the company has presented the clinical efficacy evidence from CALGB 100104 in full. The relevance of this evidence for understanding the clinical evidence of lenalidomide maintenance therapy has previously been discussed in the ERG report. In this response, the ERG discussed the appropriateness of pooling evidence from Myeloma XI and CALGB 100104 in Issue 3, and the methods of treatment switching used in CALGB 100104 in Issue 3 and Issue 4.

Issue 2: The company's model structure does not allow assumptions about subsequent treatments to be explored. The committee concluded that the company's model structure had limitations. It also concluded that there was likely to be uncertainty around the cost-effectiveness estimate because assumptions about the effects of subsequent therapies on survival could not be fully explored

The ERG disagreed with the company that the cost-effectiveness model was able to account for the clinical efficacy of subsequent therapies and believed this remained a limitation of the analysis. Furthermore, the ERG did not consider the subsequent treatments administered within the Myeloma XI and CALGB 100104 trials to be similar (discussed in further detail in Issue 7).

The ERG considered there to be uncertainty in the cost-effectiveness estimate as a result of the chosen model structure. However, it was the ERG's opinion that the model could still be useful for decision making, provided alternative subsequent therapy assumptions are explored adequately. Subsequent therapy estimates and scenarios are further discussed in Issue 7 and Section 4.

Issue 3: The company's methods and rationale for pooling Myeloma XI and CALGB 100104 data, and adjusting for treatment switching, are unclear

The ERG discussed the company's methods and rationale for pooling the Myeloma XI and CALGB 100104 data within Issue 3. The appropriateness of adjusting for treatment switching in the placebo arm of CALGB 100104 is mentioned briefly within Issue 3; however, further critique is provided in Issue 4.

In order to address concerns regarding differences in the patients enrolled in the CALGB 100104 and Myeloma XI studies, the company performed four analyses that account for differences in patient populations. These analyses are based on two Propensity Score Weighting (PSW) methods and two 'Matching Adjusted Indirect Comparison' (MAIC) methods.

PSW methods:

The propensity score analysis is broadly well conducted. In the first analysis the CALGB study was reweighted to match the Myeloma XI study, based on:

- Age (<60 years)
- Sex
- ISS
- Prior lenalidomide usage

The second PSW analysis also included 'response to ASCT' in the weighting, and therefore CALGB was reweighted to match the Myeloma XI study, based on:

- Age (<60 years)
- Sex
- ISS
- Prior lenalidomide usage
- Response to ASCT

Only the results of the analysis not including the 'response to ASCT' term were presented, though the company stated results were similar between analyses. Although the reweighted data exhibited a good match, the ERG was concerned with the large number of patients that were given a weight of close to zero, implying they were different. Under such circumstances Propensity Score Matching may have been a more suitable approach.

MAIC methods:

Limited information was provided on how the MAIC analyses were implemented, therefore the ERG was unable to comment on how well the approach had been conducted. Although not explicitly stated, the ERG believed that the MAIC analysis implemented by the company used an anchored MAIC as they described the analysis to be conducted using *the 'common comparator (placebo/observation)'*. As with the PSW approach, two analyses were presented with the same components matched for – one analysis including the term for 'response to ASCT' and one without. Again, the results presented were only for the analysis without the 'response to ASCT' term. In this analysis a good match is achieved, and similar results are observed as with the PSW analysis.

Summary of adjustment methods:

In choosing between the two approaches, given concerns raised about the performance of anchored MAIC recently by Phillipppo *et al.* (2020),³ and the established nature of propensity scoring, the ERG would elect to use a PSW based approach. This is despite the limited nature of what was provided with regards to the large quantity of zero weights which the ERG believed could have been handled differently.

What should be noted, however, is that these analyses are only able to adjust for differences in the observed characteristics of patients. Any other differences in population characteristics or in study design, such as the difference in dosing regimen (21 of 28 day [Myeloma XI] vs. 28 of 20 day [CALGB 100104]) and subsequent treatments (as presented in Table 28 of the company's post-ACD addendum) are not accounted for (aside from the adjustment for treatment switching, which remained a key concern for the ERG [detailed in Issue 4]). Furthermore, any unobservable differences are unable to be accounted for by the methods, for instance any selection bias in the patients enrolled into either study.

Following the adjusted analyses, a difference can be seen between the lenalidomide Myeloma XI and the lenalidomide CALGB 100104 adjusted KM curves in Figures 6 and 9 of the company's post-ACD addendum, where previously the arms (unadjusted) showed very similar survival. While the arms were still similar, the ERG had concerns as to why, in this population that had been adjusted for patient characteristics, the intervention arms showed a greater difference in survival. As a result, the ERG was unable to rule out the possibility that differences in the study designs such as the dosing (21 of 28 day [Myeloma XI] vs. 28 of 28 day [CALGB 100104]), or other differences (such as the benefits accrued from receiving different subsequent therapies) could be the driving force for this. It was the ERG's understanding that no re-weighting analysis could reconcile these differences.

Furthermore, in the company ACD comments (Issue 3) the company stated: "*The robustness of the KM curve with respect to matching adjustment also constitutes indirect support to the difference between the lenalidomide arm in CALGB and that in Myeloma XI could be largely attributed in lenalidomide dosing.*". For this reason, the company performed a weighted pooled analysis using MXI and (to the best of the ERGs understanding) the PSW CALGB 100104 data, including terms for trial, treatment and an interaction term between the two. It was unclear to the ERG whether the source of clinical data - 'Pooled MXI and CALGB (Myeloma XI prediction)' included in the model used PSW as suggested in the company's addendum to ACD comments. The ERG was able to obtain the company's AC1 base case ICER [REDACTED] using this data source in the model; however, no adjustments for patient characteristics or interaction term were considered prior to ACD suggesting that this data source was

unchanged from AC1. The ERG found no additional data source in the model that corresponded to a weighted analysis of Myeloma XI data pooled with PSW CALGB data, including an interaction term for trial and treatment. The company reported that a non-statistically significant difference was found for the interaction term for the new analysis, yet the company did not provide the corresponding p-values (Table 23 of the company ACD addendum), for the ERG to interpret the extent of non-significance. However, the ERG noted that the 95% confidence interval reported for the interaction was large, therefore it is likely that the level of non-significance is reasonable for the interaction term. Kaplan-Meier (KM) data, Akaike information criterion (AIC) and Bayesian information criterion (BIC) were not presented in the company's addendum to ACD for this analysis (weighted pooled analysis of MXI and PSW CALGB with interaction term), nor were they included in the model received by the ERG, thus the ERG was unable to sufficiently utilise this data source as there was no evidence to support model selection.

The company's updated preferred base case data source was Myeloma XI pooled with PSW adjusted CALGB 100104, excluding the 'response to ASCT' term. The ERG had concerns regarding the comparability of the Myeloma XI and CALGB 100104 trials that were not fully addressed by the new analyses presented by the company. In addition, the ERG was not convinced on the appropriateness of the RPSFTM used to adjust for treatment switching (discussed in detail in Issue 4). Therefore, the ERG has presented two base case scenarios for the committee:

1. The first option implements the ERG's original preferred data source; Myeloma XI data only,
2. The second option presented by the ERG incorporates Myeloma XI data for the first 60 months (length of MXI follow up) with the PSW adjusted CALGB (excluding the 'response to ASCT' term) used to inform the remainder of the extrapolation. The ERG have chosen to exclude the 'response to ASCT' term to align with the company's base case however, notes that the inclusion of the term has little influence on the ICER.

Results and scenario analyses exploring different data sources performed by the ERG are presented in Section 4.

Issue 4: The justification for using the rank preserving structural failure model should be provided

The company provided an extended explanation of methodology and results with respect to treatment switching. The company also explained (ACD Addendum p33) that RPSFTM adjusted survival estimates for CALGB were used as the basis for the MAIC/PW analyses

(Issue 3), which have informed the latest economic model. The ERG's main concern with treatment-switching related to making adjustments on later treatment lines (see subsection below).

The company reported their consideration of available treatment-switching methods, which are those outlined in TSD16. The simple methods (including censoring and excluding data) advised against by NICE in its methods guidance (NICE Methods Guide, 2013)⁴ were not used. Detailed rationale and results from RPFSTM analysis have been provided. The choice of RPSFTM appears to largely follow the guidance of TSD16.⁵ A further complex switching method, IPE, that extends RPFSTM, was carried out, though not explicitly requested by the committee.

The company considered the assumptions of RPSFTM including that of the important 'common treatment effect' described in TSD16.⁵ The company explained that the offer of treatment to non-progressed placebo patients in CALGB was made at the point of a planned two-year interim analysis, and not as a response to patient prognosis or disease progression. The ERG agrees that this provides reassurance regarding the assumption.

Within RPFSTM, the company considered a 'treatment-group' approach (selected for base case) in which the effect of lenalidomide is retained until death, and an 'on-treatment' approach in which the effect of lenalidomide only applies when it is being received.

The treatment-group approach selected for base case by the company where the treatment effect lasts until death may be in contradiction with long-term waning or abrupt discontinuation of the treatment effect (e.g. 10 years) (Issue 6). However, the effect estimates (HRs) under RPSFTM appear insensitive to this choice (ACD Addendum Table 41).

The interim analysis at unblinding in CALGB at or just before switching commenced gave HR estimates of 0.37 (95% CI, 0.26 to 0.53) for PFS and 0.52 (95% CI: 0.26 to 1.02) for OS (McCarthy et al. 2012), compared to the RPSFTM-adjusted values (on-treatment, no covariate adjustment; other cases give similar point values) of [REDACTED] for PFS and [REDACTED] for OS (ACD Addendum table 41). The similarity of the HR estimates for OS before switching was offered (0.52) and after adjustment for switching is made ([REDACTED]) seems supportive of the approach. Conversely there is some difference for PFS (0.37 at unblinding, [REDACTED] after treatment-switching adjustment).

Subsequent treatments

The RPSFTM adjusts the survival time of those who switch to lenalidomide from the placebo arm which the ERG understands applies the adjustment to all lines of lenalidomide treatment

among switchers (maintenance or subsequently). For second line, this is a logical step when mapping from a US (CALGB) to UK (Myeloma XI) context, since the company provide evidence that [REDACTED] in the UK (ACD Addendum table 28 and CS Appendices Table 64). However, UK [REDACTED] (CS Appendices Table 64) and the company's revised subsequent therapy estimates in response to the ACD indicate that 65% of observation patients would be anticipated to receive lenalidomide at third line (following a second relapse). The company, ERG and the committee all acknowledged uncertainty in subsequent therapy estimates across the course of the TA process to date, yet the range of subsequent lenalidomide at second relapse for the observation arm has stayed consistently high (varying between 50% using the MXI data to 70% applied by the ERG).

Based on this, the ERG questioned whether adjusting the effect of third line lenalidomide among switchers may be problematic (because it thereby does not correspond to the real-world UK context when lenalidomide is not available for maintenance, in which third line lenalidomide usage may be high). No information for third line usage in CALGB was presented. With the treatment-group approach, no clinical benefit associated with lenalidomide is applied at any time point; however, 65% of observation patients are assigned lenalidomide + dexamethasone at the third line in the model (in the company's base case) and thus are costed for this treatment. Based on the company's base case analysis (using adjusted CALGB curves to inform survival), and the revised subsequent therapy estimates, the ERG questioned the appropriateness of applying extrapolations for the observation arm, where the benefit of subsequent lenalidomide is removed, yet costs are accrued associated with subsequent lenalidomide.

Issue 5: Survival extrapolations should use Myeloma XI data as the main source of evidence but could be supplemented with CALGB 100104 data

The company presented a range of analyses to explore survival extrapolations within the cost-effectiveness model. The following data sources are available for selection within the model:

- Myeloma XI only (ERG's AC1 base case)
- CALGB 100104 only
- Pooled Myeloma XI with unadjusted CALGB 100104 (company's AC1 base case)
- Pooled Myeloma XI with adjusted CALGB 100104
 - PSW, without 'response to ASCT' term

- PSW, with ‘response to ASCT’ term
 - MAIC, without ‘response to ASCT’ term
 - MAIC, with ‘response to ASCT’ term
- Pooled Myeloma XI with weighted CALGB 100104 with terms for trial, treatment and interaction (trial with treatment)

The approaches taken to adjust the CALGB 100104 data to match to the Myeloma XI data as closely as possible are discussed in detail in Issue 3. The company’s preferred data source is pooled Myeloma XI with adjusted CALGB 100104 using PSW, without the ‘response to ASCT’ term. The ERG has provided a critique of the survival curve extrapolation choices discussed in the company’s addendum to ACD below.

Treatment effect:

Alternative assumptions regarding the effect of lenalidomide maintenance treatment were implemented and explored by the company for the pooled Myeloma XI and adjusted CALGB 100104 analyses, listed below:

	Treatment effect	
Options	0-60 months	60+ months
Option 1	Myeloma XI	Adjusted CALGB 100104 data (with either PSW or MAIC)
Option 2	Myeloma XI	
Option 3	Weighted pooled analysis of Myeloma XI and CALGB with covariates for trial, treatment and trial-interaction	

Abbreviations: MAIC, matched adjusted indirect comparisons; PSW, propensity score weighting

For the pooled adjusted data sources that the treatment effects are applied to, the ERG considers the first approach to be the most appropriate. Myeloma XI is the key trial for this appraisal, therefore the ERG believes the treatment effect from this study should be utilised for the observed period (60 months). From this point, the treatment effect could be taken from the chosen adjusted CALGB 100104 data. However, the ERG is still of the view that the Myeloma XI trial alone should be considered for decision making.

The ERG questions the robustness of utilising a treatment effect taken from one data set and applying it to the parameters obtained from fitting a model to a second data set, as with the second and third treatment effect options. The ERG believes it could be useful for exploration, however, has concerns around the potential added uncertainty and thus, believes the analysis

produced using option one (listed in the table above) is much more robust than the latter choices.

Overall survival:

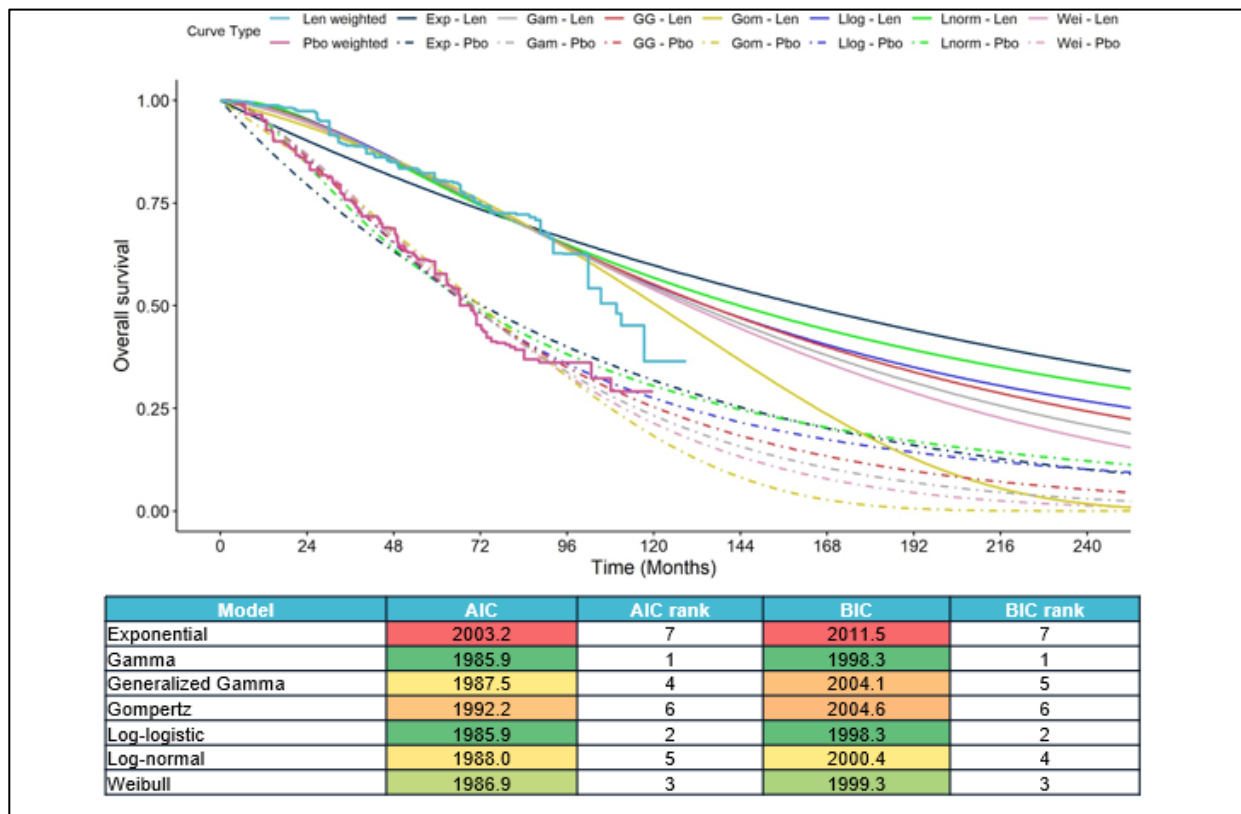
A range of scenarios, dependent on data source and survival model choices, were presented in the company's ACD addendum. For the PSW and MAIC analyses (without 'response to ASCT' term), evidence was provided in Figures 10 and 11 of the company's ACD addendum to support the choice of a joint (dependent) model.

The ERG noted that the company could have explored a piecewise approach to fitting models to the pooled analyses to allow different assumptions prior to and following 60 months, which may have provided a better fit to the data. For example, different parametric survival curves could have been used for the Myeloma XI 0-60 months and a different curve choice for the CALGB 100104 60+ months. Instead, the model is constrained so that the follow on (60+ month period) curve selection must be the same as the curve choice selected for the Myeloma XI.

Pooled Myeloma XI and PSW CALGB 100104 model fit

Figure 1 presents the fit of the joint models to the PSW adjusted CALGB 100104 data, along with the corresponding AIC and BIC goodness-of-fit statistics, presented in the company's ACD addendum. The KM data for this analysis was not available in the model sent to the ERG, therefore the ERG is unable to present alternative plots for the committee or validate the visual fit.

Figure 1: OS joint model fits to pooled MXI and PSW CALGB



Note: Image taken from the company's ACD addendum, Figure 12, page 61. Data source uses pooled MXI and adjusted PSW CALGB data without the 'response to ASCT' term.

The company fit 7 parametric models to the data (as shown in Figure 1) and selected the joint gamma model as their preferred base case as it provided the best AIC/BIC. However, the ERG note that the AIC/BIC statistics are identical to two decimal places (dp) for the gamma and log-logistic models and see no reason to prefer the joint gamma over the joint log-logistic model. Selecting the joint log-logistic model for OS extrapolation results in an increase of approximately [REDACTED] to the ICER (using the company base case) due to an increase in survival predicted for the observation arm.

Although the ERG acknowledges the similarity in AICs for the gamma distribution and the log-logistic distribution, the cost-effectiveness model sent to the ERG only has 6 parametric distributions to choose from, with the gamma distribution (the company's preferred OS selection) missing. The front settings sheet to the model has the gamma selection (with 3 parameters) as an option to select but not the gamma distribution (2 parameters). This is further confused by the model calculation sheets referring to the generalised gamma (3 parameters) as the gamma distribution. Therefore, whilst the company have stated a preference for the gamma distribution based on AIC statistics, this is not an option within the

model and instead the ERG believe the generalised gamma has been used, which has a different AIC.

The ERG is unclear why the company's decision to select the 'gamma' curve (noting the discrepancy in labelling resulting in the generalised gamma curve actually being selected in the company model) for the extrapolation of OS did not take account of the extrapolations when the piecewise approach is selected (e.g. MXI data to 60 months with CALGB 100104 PSW adjusted data thereafter). The company has based their selection on Figure 1 (Figure 12 of the company's ACD Addendum) which appears to be the CALGB 100104 adjusted data only. However, the company base case settings include a part 1 (MXI to 60 months) and part 2 (CALGB 100104 adjusted). Despite this, the company have not provided any supporting information or rationale for their decision to select the generalised gamma to extrapolate OS for the pooled-adjusted data.

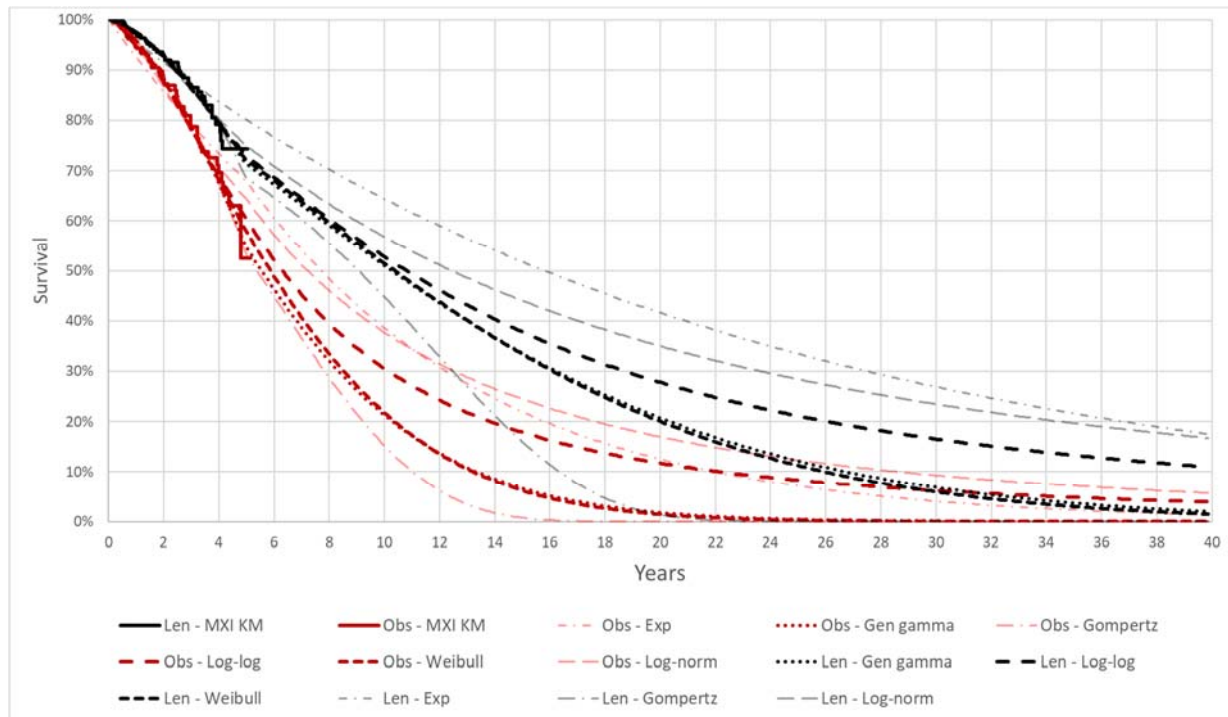
The ERG believes the appropriate curve fit should have been based on a combination of factors including:

1. AIC/BIC statistics of Myeloma XI parametric extrapolations
2. AIC/BIC statistics of the CALGB 100104 adjusted parametric extrapolations
3. Visual fit to the KM data of Myeloma XI
4. Plausible extrapolation of the CALGB 100104 adjusted parametric extrapolations
5. Plausible extrapolation of the two curves fitted together

From their description, the company's approach only appears to use factors 2 and 4. Figure 2 presents the piecewise curve fits when the Myeloma XI data is followed by PSW adjusted CALGB 100104 data for all parametric models. The KM data presented is that of Myeloma XI. Ideally the ERG would have plotted the KM of Myeloma XI for the first 60 months, followed by the PSW adjusted CALGB 100104 KM from 60+ months however, pooled data with any adjustment is not available within the company's submitted model. In observing the piecewise approach in Figure 2, the long-term extrapolations for both arms vary greatly dependent on the model selected to fit to the data. When examining Figure 2, there is substantial uncertainty across the curves fits and the two arms which can be seen not only in the disparity in the curve fits to each separate arm (lenalidomide maintenance and observation), but also the multiple times the different parametric curves cross each other (e.g. where a lenalidomide maintenance extrapolation overlaps or crosses an extrapolation for the observation arm). The ERG notes that while the inclusion of the piecewise approach combining MXI with CALGB

100104 offers the use of more long-term data, this does not directly translate into reduced long-term uncertainty.

Figure 2: Myeloma XI followed by PSW adjusted CALGB OS curves



As the pooled adjusted data was unavailable within the model, the ERG has been limited in making a fully informed decision to populate the preferred assumptions for their base case using this data. Previously, when reviewing the extrapolation of the Myeloma XI data alone, the ERG concluded that the joint log-logistic and joint Weibull models provided the best fit to the Myeloma XI data. Based on the AIC/BIC statistics and visual fit to the KM, the log-logistic and Weibull curves provide a good fit to the adjusted CALGB 100104 data also (Figure 1). Therefore, in Figure 3 the ERG have presented the joint generalised gamma (company's base case), joint log-logistic and joint Weibull curves for comparison, with Figure 4 providing a closer look at the observed period for Myeloma XI. This approach uses the selected curve for both time periods (pre 60 months and post 60 months).

Figure 3: Myeloma XI followed by PSW adjusted CALGB OS extrapolations – Company and ERG preferred extrapolations

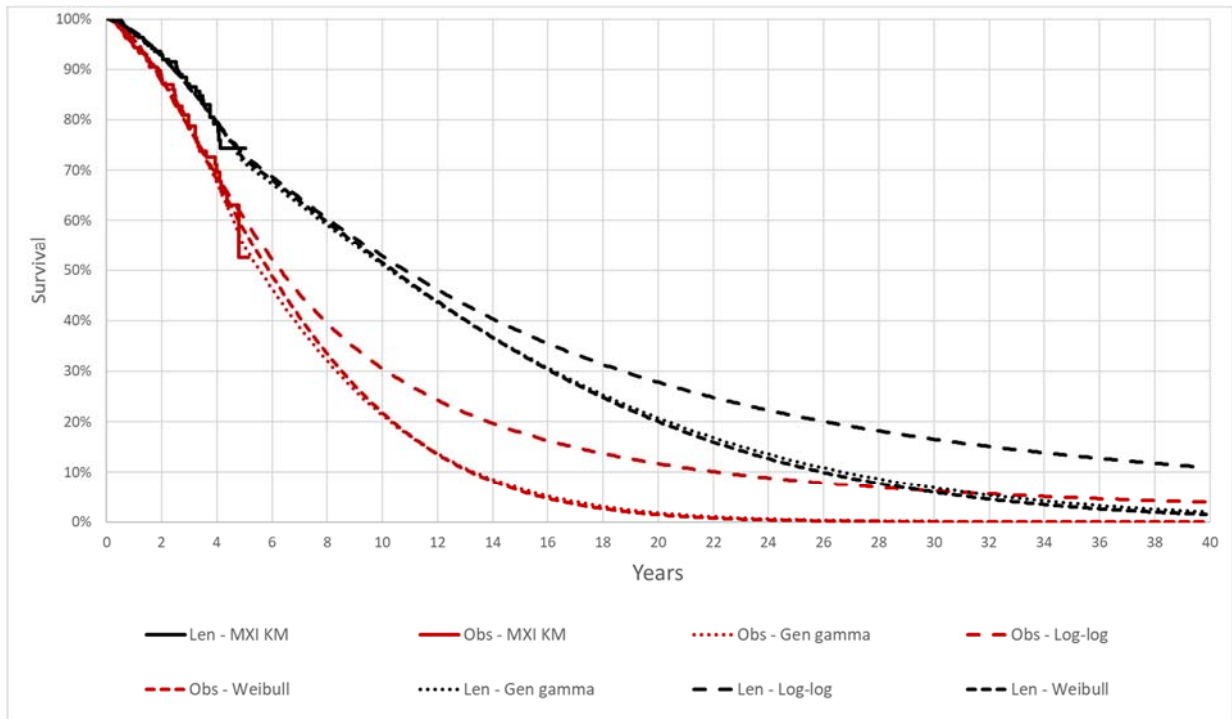
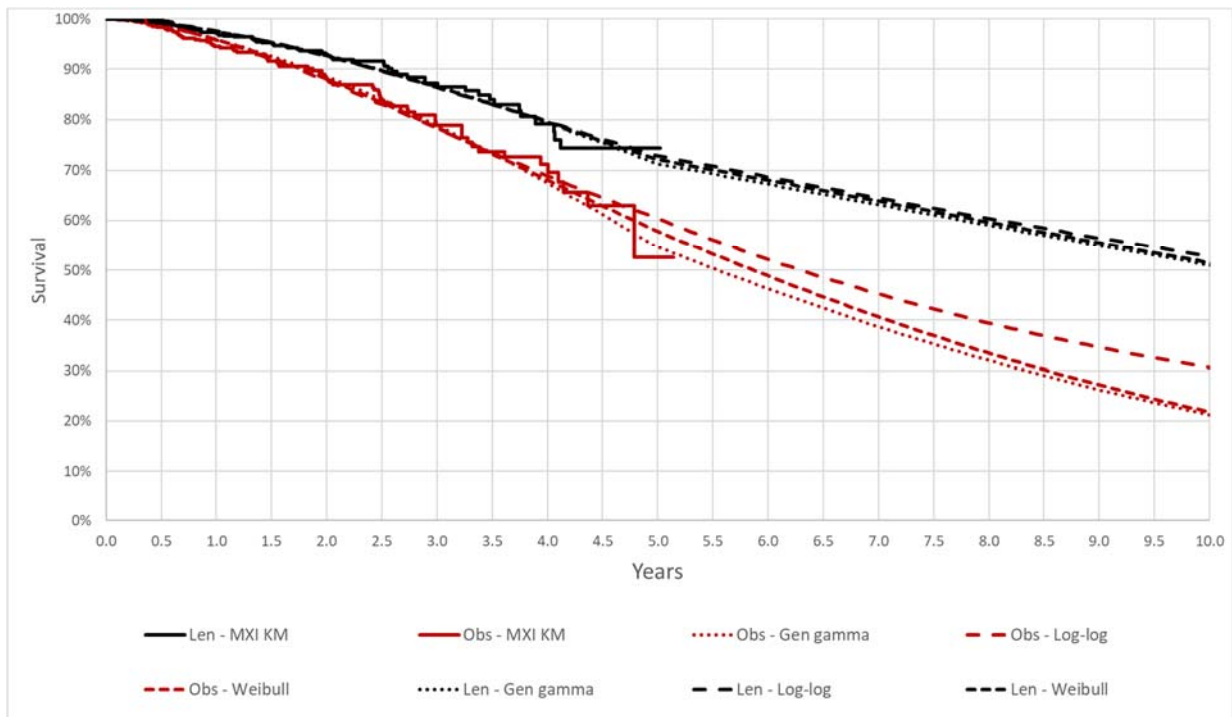


Figure 4: Myeloma XI followed by PSW adjusted CALGB OS extrapolations - Observed period close up



In observing the Myeloma XI followed by PSW adjusted CALGB 100104 curve fits to the Myeloma XI data, it can be seen that the models follow a similar trajectory for the period up to approximately 4.5 years. From this point, the generalised gamma and Weibull follow a similar

path, with the log-logistic providing a more optimistic prediction of long-term survival in both arms. As a result, the ERG considers the joint log-logistic model the most appropriate to capture the expected long-term overall survival for this data source (PSW), based on the individual fits to the Myeloma XI data and the PSW adjusted CALGB 100104 data.

The log-logistic predicts ■ and ■ OS at 10 years for lenalidomide maintenance and observation, respectively. The ERG notes that the joint Weibull model also provides a plausible (albeit more pessimistic) fit to each set of data, predicting ■ and ■ OS at 10-years for lenalidomide maintenance and observation. Acknowledging the original MXI (only) log-logistic curve, selected as the ERG base case produced estimates at ■ and ■ for lenalidomide maintenance and observation respectively, showing that the inclusion of the CALGB 100104 adjusted data for the post 60-month period has little impact on the observation arm but increases anticipated survival for the lenalidomide maintenance arm at the 10-year time point.

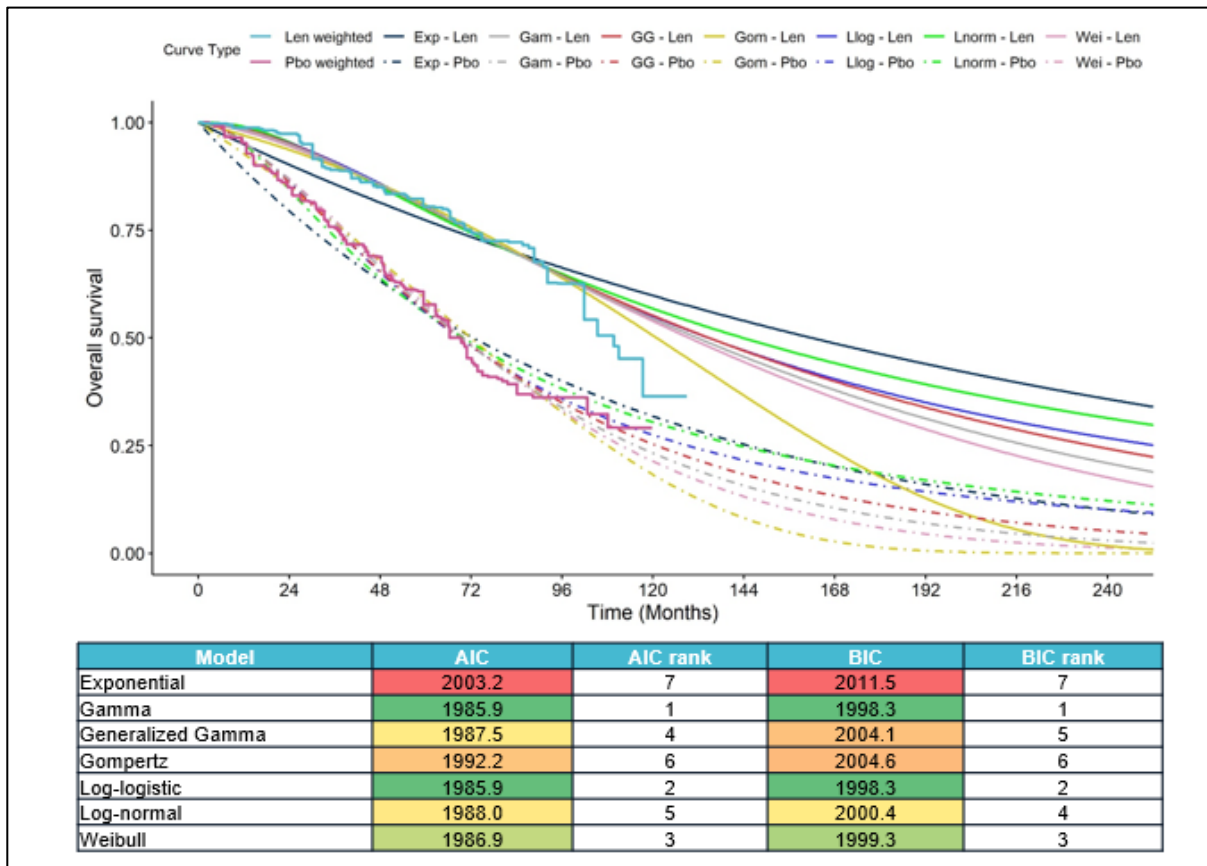
Based on visual fit, AIC/BIC statistics and plausible extrapolation the ERG believe the Weibull and log-logistic to provide reasonable estimates of OS for lenalidomide and the observation arm. Therefore, the ERG presents a scenario analysis using the Weibull model fit to this data source (PSW) in Section 4. Although the generalised gamma curve does not provide a good statistical fit to the data, the curves (for both lenalidomide maintenance and observation) do lie between the log-logistic and Weibull curve projections.

The modelling approach undertaken by the company is restricted so that the same parametric function has to be selected for both the Part 1 (MXI data to 60 months) and Part 2 (the CALGB 100104 adjusted data, 60+ months). The ERG would have preferred additional functionality within the model to explore different curve functions for the different time periods based on the different data cuts.

Pooled Myeloma XI and MAIC CALGB 100104 model fit

Error! Reference source not found. presents the fit of the joint parametric models to the MAIC adjusted CALGB data, along with the corresponding AIC and BIC goodness-of-fit statistics, presented in the company's ACD addendum. As with the PSW analysis, the ERG assumes that the treatment effect from Myeloma XI is used for the first 60 months with the MAIC adjusted CALGB (without 'response to ASCT') treatment effect applied thereafter for this figure although this is not explicitly stated. The KM data for this analysis was also unavailable in the model sent to the ERG, therefore the ERG is unable to present alternative plots for the committee or validate the visual fit.

Figure 5: OS joint model fits to pooled MXI and MAIC CALGB



Note: Image taken from the company's ACD addendum, Figure 13, page 62. Data source uses pooled MXI and adjusted MAIC CALGB data without the 'response to ASCT' term.

As detailed with the PSW analysis above, the company presented the fit of 7 parametric models to the data however, only 6 model choices were available in the cost-effectiveness model. In addition, the log-logistic and gamma AIC/BIC were again seen to be identical to 2dp. Selecting the joint log-logistic OS extrapolation results in an increase of approximately [REDACTED] to the ICER due to an increase in survival predicted for the observation arm.

Furthermore, as with the PSW analysis, the company appear not to have explored the curve selection fully and have presented the extrapolations of the MAIC adjusted CALGB 100104 data only for the full time period rather than being based on Part 1 (Myeloma XI up to 60 months) and Part 2 (MAIC adjusted CALGB 100104 data thereafter).

Figure 6, Figure 7 and Figure 8 provide the pooled MAIC adjusted curve extrapolations compared with the KM curves from Myeloma XI to explore the visual fit for the observed period. Similar to the PSW adjusted analysis, the long-term extrapolations vary greatly with the log-logistic providing a more optimistic survival prediction than the generalised gamma and Weibull curves in the longer term.

Figure 6: Myeloma XI followed by PSW adjusted CALGB OS curves

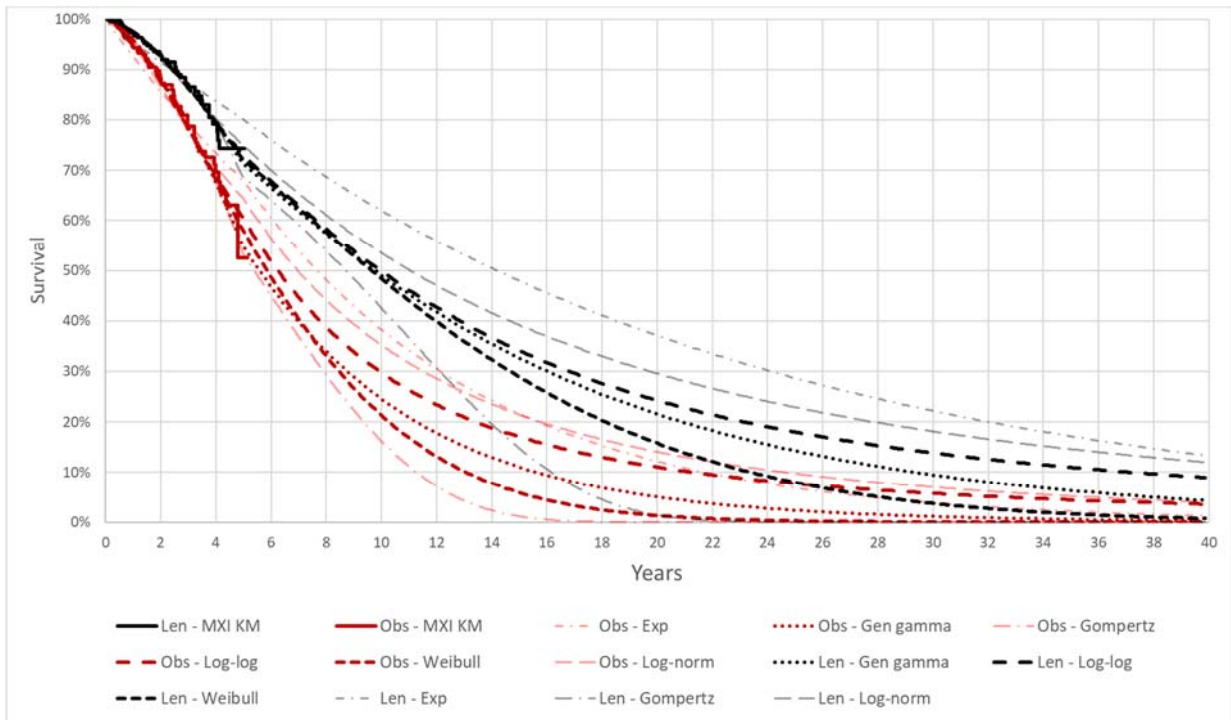


Figure 7: Myeloma XI followed by PSW adjusted CALGB OS extrapolations – Company and ERG preferred extrapolations

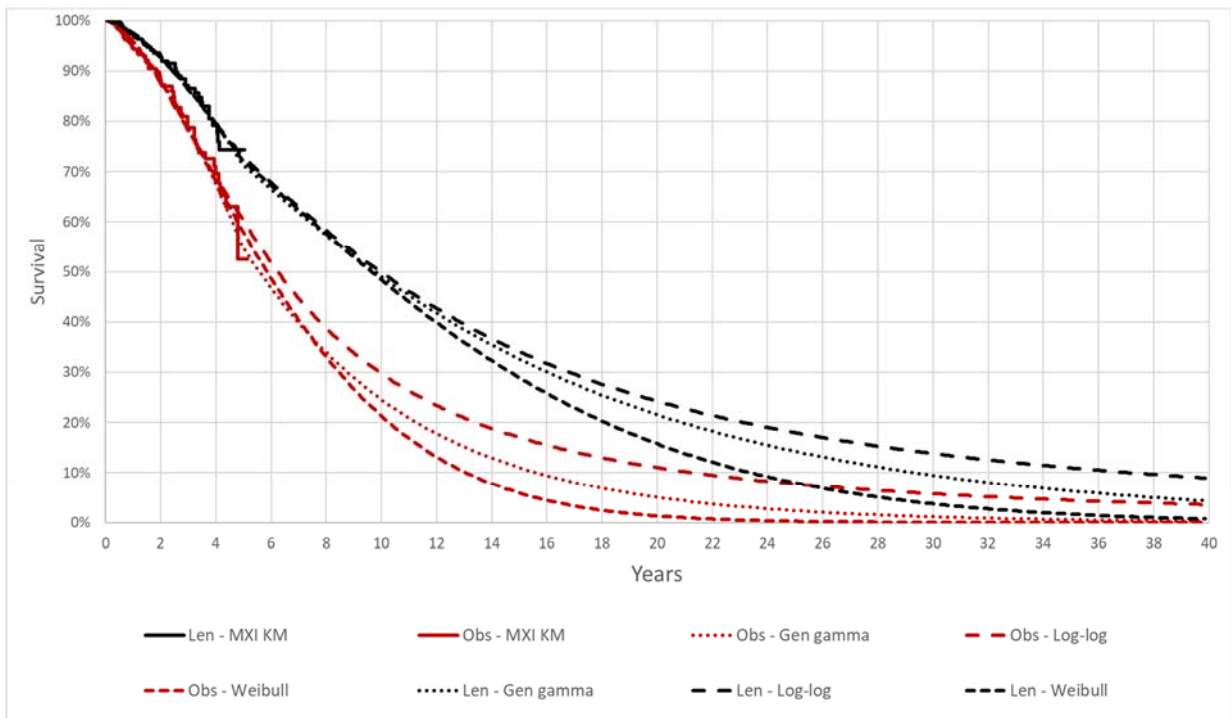
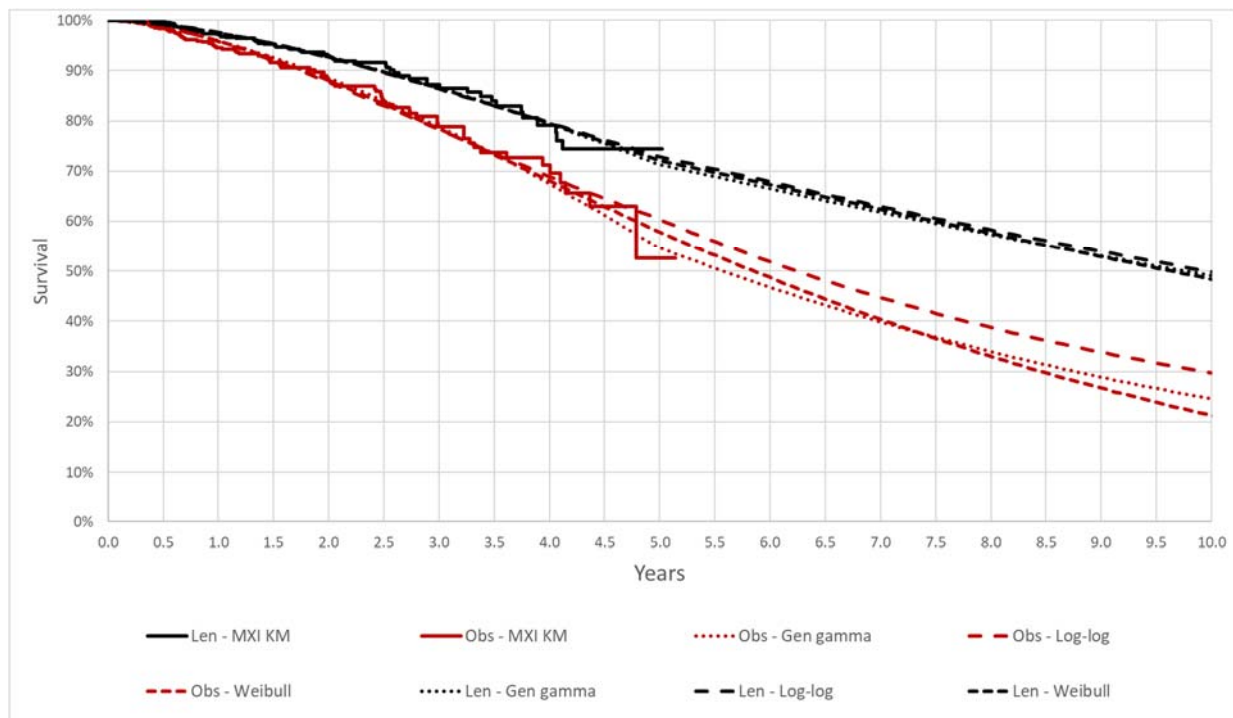


Figure 8: Myeloma XI followed by MAIC adjusted CALGB OS extrapolations - Observed period close up



As the ‘gamma’ distribution preferred by the company is not available to the ERG in the cost-effectiveness model (discussed above with PSW model fit), the ERG considers the joint log-logistic model the most appropriate to capture the expected long-term overall survival for this data source (MAIC), based on the individual fits to the Myeloma XI data and the PSW adjusted CALGB 100104 data. The log-logistic model predicts ■ and ■ OS at 10 years for lenalidomide maintenance and observation, respectively. As with the PSW analysis, the ERG considers the joint Weibull model to also provide a plausible (albeit more pessimistic) fit to each set of data, predicting ■ and ■ OS at 10-years for lenalidomide maintenance and observation. Therefore, the ERG presents a scenario analysis using the Weibull model fit to this data source (MAIC) in Section 4.

Myeloma XI only

As the ERG considers the pooling of Myeloma XI and CALGB 100104 to remain limited by the different dosing regimens and subsequent therapies received in the trials, the ERG still believes the analysis performed using the Myeloma XI only analysis (ERG original base case) to be relevant to this appraisal. The ERG’s preferred OS extrapolation for the Myeloma XI data remains the joint log-logistic model, with the joint Weibull model also providing reasonable (yet more pessimistic) estimates. As previously mentioned, the log-logistic model predicts ■ and ■ OS at 10 years for lenalidomide maintenance and observation, with the Weibull model

predicting ■ and ■, respectively. Detailed rationale for this choice of curve can be found within the ERG report and the ERG response to technical engagement.

To provide the committee with results from a Myeloma XI data only and a pooled data analysis, the ERG has presented two base case scenarios for the committee to consider in Section 4 based on different data sources. Both base case analyses utilise the joint log-logistic model for OS extrapolation, with the joint Weibull extrapolation explored in the scenario analyses.

Progression-free survival

The company opted to use the same PFS curve chosen at the technical engagement stage of the appraisal, the joint gamma model (suspected by the ERG to be the generalised gamma). The company provided a range of ICERs dependent on data source and OS model selection in tables 24, 25 and 26 of the company's ACD addendum using either the company's preferred PFS extrapolation (gamma) or the ERGs (Weibull).

The addendum states *'Although matched-adjusted PFS was included in the new version of the model, PFS remains a relatively uninfluential parameter in the cost-effectiveness... The results of the PFS analyses are presented in the Appendices for completeness.'* The adjusted PFS curves were not provided within the model or presented in the addendum appendices. Therefore, the ERG has been unable to validate whether the previously chosen PFS model selections are the most appropriate extrapolations for the adjusted data sources. At the previous stages of this appraisal, the choice of PFS curve was found to have little impact on the cost-effectiveness results therefore, the ERG does not have any great concerns having not seen the adjusted PFS data. However, the ERG is unable to validate the PFS curve selections made by itself and the company for the adjusted analyses.

Issue 6: The treatment effect of lenalidomide maintenance may wane over time and this should be included in the model

The ERG agreed with the company that for the observed period in both trials (~10years in the longest study [CALGB 100104]), there is evidence of a constant treatment effect for lenalidomide maintenance. However, despite the rationale provided in the company ACD comments, the ERG is still of the opinion that there is no evidence to support the proportional hazards (PH) assumption holding indefinitely. The ERG has provided scenario analyses relating to the loss of a treatment effect; however, did not incorporate this in its base case.

At the request of the committee, the company explored a waning treatment effect of lenalidomide maintenance over time, presenting results of waning at 10 years. The company adapted the ERG's scenario to explore treatment waning to obtain the results. At the specified

time point, the hazard ratio between arms becomes equal to one such that no treatment effect is applied to lenalidomide after that time.

The company presented the results of a 10-year waning effect for a range of data source and OS extrapolation options in Table 27 of the company's ACD Addendum.

However, the ERG found discrepancies between the ICERs reported in the company's ACD Addendum and those found when applying the waning effect themselves. Therefore, the ERG presented the results of some key scenarios that explored a treatment waning effect on both the company and ERG base cases in Section 4.

Issue 7: Costs of subsequent treatments are highly uncertain so scenarios should be presented

The company's revised base case ICER is based on 10% of observation patients receiving lenalidomide + dexamethasone and 15% receiving 'other treatment' in the second line.

The ERG was unable to replicate the ICERs presented in Tables 30 and 31 of the Addendum using the estimates mentioned above. As a result, the ERG reproduced the key scenarios in Section 4 below for consideration by the committee.

Comparability between trials:

Table 28 in the company ACD Addendum presented the second line therapies received by patients in the Myeloma XI and CALGB 100104 trials. The ERG found this table highly confusing. The proportion of patients reported to have received 'any second line anti-myeloma treatment (AMT)' and 'no second line AMT' for the CALGB 100104 trials equaled 100% for all arms as expected; however, for the Myeloma XI trial these were 91.7% and 85.5% for lenalidomide maintenance and observation, respectively. Beneath this table, the company stated: '*Some patients received monoclonal antibodies in very small proportions as subsequent therapies (CALGB: Myeloma XI:* '; however, this is not reported explicitly within Table 28. Furthermore, the ERG was unable to find the reference linked to this statement due to formatting in the Addendum.

The ERG did not consider the second line therapies to be similar between Myeloma XI and CALGB 100104. Second line lenalidomide use in CALGB 100104 was and for lenalidomide maintenance and placebo (in placebo switchers); however, in Myeloma XI these values were and for lenalidomide maintenance and observation, showing much lower use in the UK based trial. Bortezomib + dexamethasone in the second line was and

█ for lenalidomide maintenance and observation, with lower proportions administered this treatment in CALGB 100104 (█ - lenalidomide maintenance, placebo, placebo switchers). Though the ERG understands there is uncertainty in subsequent therapy estimates in a UK treatment setting, the differences across the two trials was very apparent. The ERG further noted that these differences in subsequent treatment would be likely to impact extrapolations of the two trials, of which adjustment methods cannot account for, and as such the appropriateness of pooling the trials may be questionable. As acknowledged by the company, the differences in subsequent treatment (alongside the dosing regimens) may explain the separation of the MXI and CALGB adjusted curves. Conversely the company also stated that clinician feedback indicated that subsequent therapy distributions between the two trials would not translate into 'material differences' in overall survival. Given these points seem contradictory, the ERG was unclear as to whether subsequent therapy would influence overall survival differences between the two trials.

Company scenarios:

The company has presented three scenarios to explore the effect of subsequent therapies in the model, which are discussed in turn below:

- Scenario 1: 'Celgene's revised estimates' (this was presented as the company revised base case in Table 29 of version 2.0 of the company response addendum, however subsequent clarification by the company noted that this was not actually used for the base case)
- Scenario 2: Increased ASCT rates – from 5% to 10% in both arms
- Scenario 3: Increased lenalidomide rates – from 0% to 10%, observation arm only (actual company revised base case).

Company's revised subsequent therapy estimates (Scenario 3):

The company noted they found the ICER to remain cost effective when using their revised base case estimates for subsequent therapies.

Table 1 presents a comparison of the company's revised estimates (Scenario 3) and ERG's revised base case subsequent therapy estimates for patients following lenalidomide maintenance.

Table 1: Comparison of company’s revised estimates (Scenario 3) and ERG revised base case assumptions for lenalidomide maintenance subsequent therapies

Treatment arm	Lenalidomide maintenance			
Option	Company revised estimates		ERG revised base case	
Line	Post 1st relapse (2 nd line)	Post 2nd relapse (3 rd line)	Post 1st relapse (2 nd line)	Post 2nd relapse (3 rd line)
Len + dex				
Bor + dex	■	■	■	■
Car + dex				
Pan + bor + dex		■		■
ASCT	■		■	
Other	■	■	■	■
No treatment	■	■	■	■

Abbreviations: ASCT, autologous stem-cell transplantation; bor, bortezomib; car, carfilzomib; dex, dexamethasone; ERG, evidence review group, len, lenalidomide; pan, panobinostat; pom, pomalidomide

Previously the ERG and company disagreed on the proportion of patients expected to receive a second ASCT at the second line following lenalidomide maintenance treatment, with the company favouring a smaller proportion (■) and the ERG a larger proportion (■). Due to the high level of uncertainty in the estimates for subsequent therapies, the ERG considered the company’s revised estimate of ■ to be reasonable. However, based on clinical advice received, the ERG noted that a greater proportion of patients treated with lenalidomide maintenance are more likely to be in a health state that is eligible for a second ASCT than those on observation. Therefore, the ERG produced scenarios to explore the impact in Section 4.

The distribution of subsequent therapies at the third line, following lenalidomide maintenance, remained aligned between the ERG’s original and company’s technical engagement base case assumptions (Table 2).

Table 2 presents a comparison of the company’s revised base case (Scenario 3) and ERG’s revised base case subsequent therapy estimates for patients following observation.

Table 2: Comparison of company (Scenario 3) and ERG revised base case assumptions for observation subsequent therapies

Treatment arm	Observation			
Option	Company revised base case		ERG revised base case	
Line	Post 1st relapse (2 nd line)	Post 2nd relapse (3 rd line)	Post 1st relapse (2 nd line)	Post 2nd relapse (3 rd line)
Len + dex	■	■		■

Bor + dex	■	■	■	■
Car + dex	■			
Pan + bor + dex		■		■
ASCT	■		■	
Other	■	■	■	■
No treatment	■	■	■	■

Abbreviations: ASCT, autologous stem-cell transplantation; bor, bortezomib; car, carfilzomib; dex, dexamethasone; ERG, evidence review group, len, lenalidomide; pan, panobinostat; pom, pomalidomide

The estimated second line therapies following observation differs only in the proportion of patients assumed to receive lenalidomide, carfilzomib and ‘other’ treatment between the company’s and ERG’s base case assumptions. The ERG has previously raised concerns over the use of carfilzomib + dexamethasone following ASCT as the majority of patients would be administered an induction regimen including the use of bortezomib. NICE TA457⁶ guidance states that carfilzomib is recommended only for patients that have not previously received bortezomib, thus the ERG do not consider this option to be relevant assuming patients are managed per current practice. Lenalidomide is not recommended by NICE in the second line however, based on clinical opinion, may be used here if Cancer Drug’s Fund (CDF) treatments were unavailable for use.

In the third line, the distributions differ more so. Clinical advice to the ERG indicated the majority of observation patients would likely receive lenalidomide + dexamethasone in the third line, hence the high proportions observed in both the company’s (■) and ERG’s (■) base cases. The ERG’s original assumptions were based primarily on the company’s base case following clarification and amended based on clinical advice. The company’s original estimate of patients to receive lenalidomide + dexamethasone and panobinostat + bortezomib + dexamethasone were ■ and ■, respectively. Clinical experts presented with the ERG’s previous assumptions did not identify any reason to change the proportions assigned to these regimens at the third line, therefore the ERG’s estimates (■ and ■) are unchanged.

Furthermore, the ERG also previously stated that it is unlikely there would be a difference in the proportion of patients to receive *any* subsequent treatment between arms. The ERG was presented with any evidence to support a higher proportion of lenalidomide maintenance patients receiving no treatment in the third line compared to observation patients (■ vs. ■). Therefore, the ERG’s preferred assumption is unchanged, with the proportion of patients receiving no treatment following a second relapse aligned between the arms (■).

Company’s subsequent therapy scenarios:

Table 30 of the company’s ACD Addendum presents a range of scenarios (dependent on data source and OS extrapolation model) using the company’s ‘higher second ASCT rates’

subsequent therapy distribution. In this scenario (Scenario 2) the company reduced the proportion of patients assigned to 'other' treatment by ■ and redistributed this to second ASCT (from ■ to ■). All ICERs presented were under the £30,000 willingness-to-pay threshold (using the company's revised estimates for all other therapy distributions).

Table 31 of the company's ACD Addendum presented the same range of scenarios however, applies both the 'higher second ASCT rates' (Scenario 2) and 'higher lenalidomide second line rates' scenarios (Scenario 3). The ICERs in Table 31 remain below £30,000 for all but two scenarios.

The ERG noted that it was unable to replicate the ICERs presented in either Table 30 or 31 of the company's ACD Addendum.

The ERG notes that in the company's ACD addendum it states: *'An increase in the rates of lenalidomide second line in observation improves the ICER substantially. This improvement may be substantial if, in the future, the proportion of lenalidomide in second line, should maintenance not be approved, may become much higher than rates from Myeloma XI and CALGB'*. The ERG noted that lenalidomide use in the second line is not currently reimbursed by NICE so it is unclear to the ERG why the company might expect an increase in lenalidomide use at that line in the pathway in the future should lenalidomide maintenance not be approved.

There are high levels of uncertainty surrounding the distribution of subsequent therapies with the assumptions having a high influence on the cost-effectiveness results. Therefore, the ERG produced an updated heat map exploring the cost-savings relating to subsequent therapies (Section 3.1), in addition to scenarios exploring alternative assumptions, to aid the committee with decision making.

Issue 8: Myeloma XI trial data should be used to estimate relative dose intensity

The company's ACD Addendum provides an at length description of how RDI was estimated from the Myeloma XI data. This covers sections including:

- RDI from Myeloma XI data
- RDI calculation
- Lenalidomide prescription data collected during Myeloma XI
- Lenalidomide drug consumption data analysis

- From dose prescribed to packs dispensed
- Missing data
- Calculations of RDI
- Results
- Conclusions
- Wastage

The company stated that lenalidomide is a fixed dose therapy (not personalized by weight) and that patients begin treatment on a 10 mg pill taken daily as standard for 21 days with a 7-day break. This is aligned with the Myeloma XI study.

The company expressed the importance of ensuring tolerability to treatment and, as such, explain that in the case where the initial dose of 10 mg is not well tolerated, and according to the label, the daily dose can be reduced to 5 mg for 21 days in 28 days, with a 7-day break. The company further explained that according to clinical opinion, physicians may take other measures to dose adjustments, by spacing doses or cycles, or mixing doses however no explanation is provided to the clinical opinion which was sought to validate this.

The company stated that the appropriate calculation of the relative dose intensity for lenalidomide used in maintenance should include the following factors:

1. Drug regimen prescribed, either 10 mg or 5 mg dose;
2. Frequency of dose, adjusted by the physician on a patient basis;
3. Spacing out of intervals between a cycle and another, in addition to the 7-days break, for treatment-related or treatment unrelated reasons.

The ERG agreed that these considerations would factor into the calculation of RDI for lenalidomide maintenance treatment.

To estimate RDI the company used the Myeloma XI treatment prescription data estimated as the proportion of the number of packs used over the duration of therapy for a patient, and the number of packs that would be required to cover 100% compliance. The company then used this information to estimate the RDI based on the proportion of cycles that were prescribed as 10 mg and the proportion that were prescribed as 5 mg over the number of treatment cycles. Given that the dosage in the submission is 10 mg once daily (and aligns with the dosing regimen stated as part of the Myeloma XI trial), the ERG was unclear why RDI was separated

out by dosing regimens of 5 mg and 10 mg, as the ERG understand that typically any reduction of the 10 mg dose would represent a corresponding reduction in RDI.

To estimate the RDI from the Myeloma XI trial the company considered differences in dosing and data collection between Protocol 5 and Protocol 6. These are quoted below:

- Protocol 5

Consumption data collected in Myeloma XI, Protocol V5, were:

1. Start and end date for each treatment cycle, collected alongside Myeloma XI.
2. Total accumulated dose prescribed for the cycle (i.e. 210 mg for the full per protocol dose, 105 mg for a reduced dose, regular treatment, and a variety of other total cumulative doses between 5 mg per cycle and 525 mg per cycle).

- Protocol 6

Drug dosing data were subjected to a protocol amendment with Protocol 6. The variables collected were:

3. Start and end date for each treatment cycle, collected alongside Myeloma XI
4. Whether the cycle was a 'per protocol' cycle or not
5. Whether the treatment was reduced or delayed or omitted.

For both Protocol 5 and Protocol 6, the start and end of each treatment cycle was the time of the first and last doses taken for each cycle, and not the date on which the treatment was actually dispensed. As part of data collection, and outside of the 10 mg daily dose (and subsequently acknowledged 5 mg daily dose), eight alternative dosing schedules for lenalidomide were noted in Table 42 of the response which ranged from [REDACTED] per treatment cycle to [REDACTED]. Given the purpose of maintenance treatment is to offer a tolerable treatment to sustain response, the ERG was unclear as to why so many dose adjustments and different treatment regimens were required and subsequently why they may be necessary to factor in to the RDI calculation.

The company presented several assumptions which have to be made to estimate RDI, and broadly, the ERG considers that the approach taken by the company may be conservative in some instances, e.g. where no dosing data can be interpreted from the trial analysis, the full 210 mg (10 mg for 21 days) is assumed. Unfortunately, given the length and lack of clarity of the description presented by the company, the ERG was still unclear about the approach

taken overall to estimate RDI. Further, as part of the results subsection of Issue 8 the company highlighted the numbers presented as part of estimating RDI. Of these there are several inconsistencies where numbers do not total. Examples of this include the description of the prescribing information and the total columns in Table 46. These factors limited the ERG's ability to reliably validate the approach taken, and the ERG are unable to interpret (or re-calculate) exactly how the RDIs ([REDACTED]) were estimated (particularly given the different varying dosing options e.g. alternated 10 mg and 5 mg /21 days, 7 days interval as presented in Table 42 of the company response).

While the two RDIs appear similar ([REDACTED]), due to the reduced dose of 5 mg being accounted for separately, the true assumed average dose in the economic model when accounting for this weighted average is lower still. Based on the company assumptions the average dose applied per 10 mg dose is [REDACTED] (as explained in the calculation Table 3 and formula below).

Table 3: Dosing assumptions applied within the company model*

Dosing	RDI for respective dose	% assumed to receive each dose	Total RDI	Total assumed dose as an average per patient
10 mg	[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]
5 mg*	[REDACTED]	[REDACTED]		

Notes:

+ Please note that the values presented within Table 45 of the company response to the ACD differ slightly from those in the model ([REDACTED]).

* This is also labelled as 10 mg within the company's economic model however is aligned to the assumptions (RDI and cost) applied for 5 mg so the ERG have assumed this is a typo.

[REDACTED]

The original RDI proposed by the company at the company submission stage was [REDACTED] corresponding to an average dose of [REDACTED]. The ERG believed that given the simplicity of the 10mg and 5mg dosing (i.e. not weight based individual dosing), and the suggested tolerability of the 10mg dose (as supported by clinicians) this revised RDI value may be too low to reflect real UK practice. This estimate is further questioned by the TMM1 study which although in a later line of treatment (arguably where patients are sicker) and with a higher dose (25 mg - which presumably may increase the risk of patients experiencing toxicity), had a corresponding RDI of 94.9%.

On account of this, the ERG proposes a simplified (albeit limited) view of RDI which may be relevant for consideration based on prescribing of packs as opposed to doses received. Based on Table 45 of the response to the ACD where it's is acknowledged the

[REDACTED], the corresponding RDI assuming that all patients received 21 days at 10 mg or 21 day at 5 mg would be [REDACTED] calculated as:

[REDACTED]

Whilst a simplifying assumption has been made, the [REDACTED] accounts for patients receiving a reduced dose of lenalidomide and also allows accountability for the non-linear pricing of lenalidomide. Using this revised simplified approach also produces an RDI which is far closer to the 94.9% RDI observed within the TMM1 study (which had a corresponding higher dose of 25 mg) and was presented as the original ERG base case analysis.

Given the model is sensitive to assumptions around RDI and the uncertainty around the estimate, the ERG believe it is important to consider the impact on the ICER from both a trial perspective (applying values observed in MXI derived by the company) and also values which may be more reflective / plausible for a real-world UK setting. Therefore, the ERG has explored a range of scenarios presented in Table 4. The impact on the ICERs are presented within Section 4.

Table 4: RDI assumptions explored by the ERG

Scenarios applied	RDI			
Description	Original company base case using MXI	Company base case See Table 3 using MXI	Revised ERG base case using simplified approach to MXI	Original ERG base case from TMM1
Company revised base case (all other settings equal)	[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]
ERG revised base case (all other settings equal)	[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]

Issue 9: A model scenario reflecting a 1-day to 28-day lenalidomide treatment regimen.

The lenalidomide marketing authorisation recommends a dosage of 10 mg once daily on Days 1 to 28 of repeated 28-day cycles, while the company's submission and corresponding

trial data from Myeloma XI is based on 10 mg once daily on Days 1 to 21 of a repeated 28-day cycle. The CALGB 100104 study used to validate long-term survival in this indication matches the licensed dose. As part of the first appraisal committee meeting, there was ‘unanimous support’ for the use of lenalidomide with the 21-day regimen (dosed at 10 mg daily) from clinicians, patient experts and other stakeholders (ACD, Section 3.4 , p7). Despite this support, given the misalignment between trials and licensed dosing versus proposed UK practice, the committee requested that the company provide scenario analysis accommodating a 28-day dosing schedule. While the company have provided this as scenario analysis, it does not inform their revised base case assumptions.

The company ACD Addendum stated: ‘*A scenario analysis was conducted using the matched-adjusted CALGB data and the corresponding matched-adjusted time on treatment curve.*’. The company did not present the matched-adjusted time on treatment curves in the addendum or cost-effectiveness model, nor have they stated the preferred extrapolations for each data source. Furthermore, it was unclear to the ERG whether the time on treatment curves used were taken from Myeloma XI for the first 60 months and from the matched-adjusted CALGB curve thereafter. The ERG believed this approach to using time on treatment curves would be the most appropriate; however, were unable to confirm whether this was what the company had implemented. It was the ERG’s understanding that all other analyses presented in the addendum use the Myeloma XI time on treatment curve.

The company has not explicitly stated how this scenario was implemented within the model. It was the ERG’s understanding that only the dosing was changed with RDI, medical resource use (MRU) and adverse event (AE) rates remaining equal to the 21-day dosing regimen. The ERG considers it likely that similar MRU would be seen across a 21-day or 28-day dosing regimen however, RDI and AEs may differ between the regimens. AEs have a relatively low impact on the ICER however, RDI assumptions are highly influential on the results. Furthermore, it is unclear whether the company have utilised the RDI from the CALGB 100104 trial for any timepoints in this scenario or used the Myeloma XI RDI throughout.

The company’s results (Table 32) showed ICERs greater than £30,000 when using the Gompertz model and the unmatched pooled analysis of Myeloma XI and CALGB 100104 for the log-normal and log-logistic distributions. All other ICERs presented were lower than £30,000.

The ERG had several concerns with the analysis presented. Firstly, none of the ICERs presented in the company’s ACD Addendum were able to be replicated by the ERG due to a lack of scenario description and only the Myeloma XI time on treatment curves being available in the cost-effectiveness model. Secondly, it was unclear how RDI, MRU and AE are

considered for this scenario. The ERG performed a naïve scenario analysis using the Myeloma XI; data, time on treatment, RDI and AE rates to explore the effect of a 28-day regimen. While this scenario is limited to using the same efficacy data as the 21-day regimen (and thus cannot account for potential differences in efficacy and RDI between dosing regimens), the ERG believed this scenario was a more realistic depiction of the impact of a 28-day regimen compared to what was presented by the company, despite its limitations. In addition, the ERG was unable to explore other data sources as the corresponding time on treatment curves are not options in the model.

While the ERG had concerns over this scenario analyses, clinical evidence provided to the ERG implied that should lenalidomide maintenance be approved following ASCT, it would be dispensed via a 21-day dosing regimen.

3. COMPANY MODEL FOLLOWING ACD

In response to the ACD, the company presented an updated base case in the document 'ID475 Celgene ACD Addendum v3'.

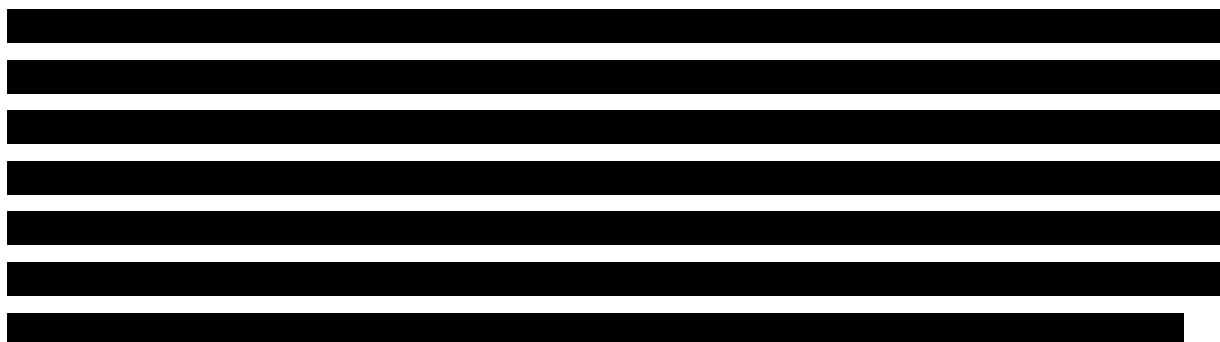
The company's revised base case incorporated the ERG corrections made post-TE (detailed in the ERG's response to TE) and the changes detailed in Table 5.

Table 5: Company base case changes

Change	Company's previous base case	Company's revised base case
Change of launch date	1 st November 2020	1 st January 2021
Data source	Myeloma XI and CALGB 100104 data pooled without adjustment	Myeloma XI and CALGB 100104 data pooled with PSW adjustment
Treatment effect	Predicting for Myeloma XI	Myeloma XI for 0-60 months, adjusted CALGB for 60+ months
OS extrapolation	Joint Weibull	Joint Gamma (ERG suspect generalised gamma)
Subsequent therapy distribution*	2% of patients receiving a second ASCT at the 2nd line	5% of patients receiving a second ASCT at the 2nd line

Key: ASCT, autologous stem cell transplant; ERG, evidence review group; OS, overall survival.

Note: *Discrepancies were found between the subsequent therapy estimates reported by the company in the ACD Addendum and the estimates contributing to the base case in the cost-effectiveness model



The ERG considered this a reasonable change.

The new company base case lowered the ICER from [REDACTED] to [REDACTED], however, the ERG noted the analysis was subject to the following errors:

- The reported OS extrapolation model applied to achieve the revised ICER is the joint generalised gamma, not the company preferred gamma distribution (further detailed in Issue 5).
- The proportion of patients assigned to receive lenalidomide maintenance 10 mg dose was rounded to 1 decimal place.

As the company’s preferred OS model, ‘gamma’, was unavailable in the cost-effectiveness model, the ERG was unable to provide a fully corrected ICER. However, a partly corrected analysis using the unrounded proportion to receive the 10 mg dose is presented by the ERG, though has little impact on the cost-effectiveness results.

The company revised and ERG partly corrected (for proportion assigned 10 mg only) base cases are presented in Table 6.

Results using the actual cPAS’s are provided in the addendum to this report.

Table 6: Company revised and ERG corrected base case (company PAS settings)

Arm	Total			Incremental			ICER (£/QALY)
	Costs (£)	LYs	QALYs	Costs (£)	LYs	QALYs	
Company revised base case (deterministic)							
Observation	████	████	████				
Lenalidomide	████	████	████	████	████	████	████
Company revised base case – ERG partly corrected (deterministic)							
Observation	████	████	████				
Lenalidomide	████	████	████	████	████	████	████

Key: ERG, evidence review group; ICER, incremental cost-effectiveness ratio; LY, life year; QALY, quality adjusted life year.

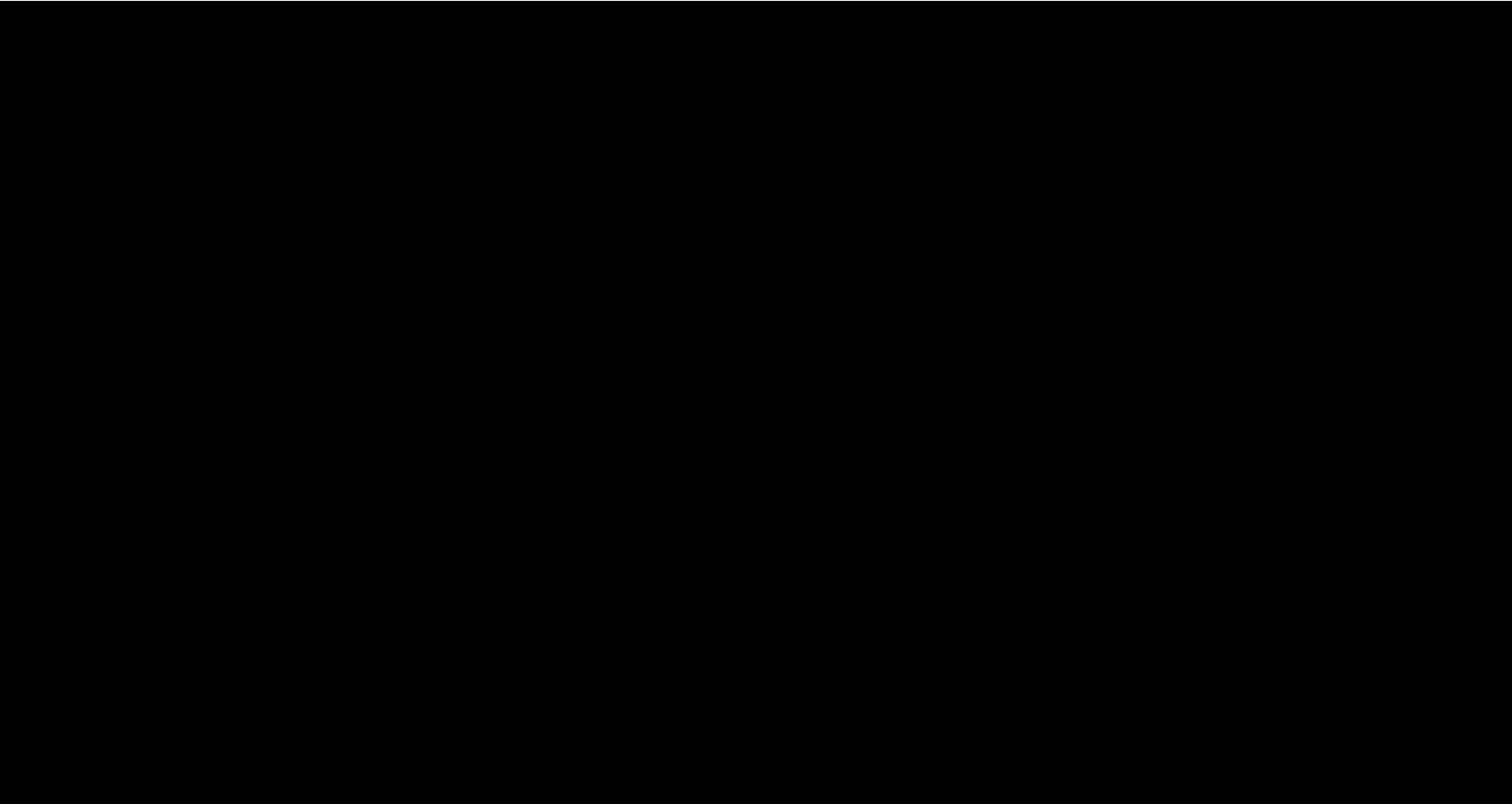
Note: ERG partly corrected analysis only used unrounded proportion of patients assigned to lenalidomide maintenance– OS extrapolation remains with a joint generalised gamma model

The main drivers of the reduced ICER compared to the post-TE results were the changes made to the data source and subsequent therapy distributions.

3.1. Updated heat map

The ERG reproduced the heat map (Figure 9), provided in the original submission, with the company’s updated base case preferences. The heat map illustrates the combined effect of subsequent therapy and treatment duration assumptions on the company’s base case ICER. Figure 9 highlighted the ICERs when the treatment duration and cost savings were set to the company’s base case, when the treatment effect duration was set to 10 years and highlighted the company’s previously estimated cost savings.

Figure 9: Updated heat map of the company's base case (company PAS settings)



4. ERG MODEL FOLLOWING ACD

In response to the ACD, the ERG decided to present two base case analyses. The first incorporated data from Myeloma XI only as the ERG believed these data to be the most relevant for the appraisal. The second used the company's Myeloma XI followed by PSW adjusted CALGB 100104 data. While the ERG recognised that the follow-up in CALGB 100104 was significantly longer than Myeloma XI (ten years vs. five years), the ERG still had reservations on the suitability of using the CALGB 100104 data to predict clinical outcomes for this appraisal. These are discussed in detail in Issue 3 and Issue 4 but included; the inability to account for differences in dosing regimen and subsequent therapies, the lack of data available for the ERG to validate assumptions and the appropriateness of adjusting for treatment switching in the placebo arm of the CALGB 100104 trial.

All analyses below use the company PAS settings. Results when using the actual cPAS settings are presented in the Addendum to this report.

4.1. Changes to the ERG base case

Aside from changing the clinical data source for one of the base case scenarios presented for the committee, the ERG made the following amendments to its base case:

- Subsequent therapy distributions (Issue 7)
 - Aligned with the company's estimates for a second ASCT at the second line (reduced from 15% to 5% for lenalidomide maintenance)
 - Removed the use of lenalidomide + dexamethasone in the second line for observation (reduced from 30% to 0%)
 - Increased the estimate for bortezomib + dexamethasone in the second line for observation (increased from 40% to 60%)
 - Increased the estimates for 'other treatment' in the second line (from 20% to 30% for both arms)
- RDI (Issue 8)
 - Decreased from 94.9% to 91.62%. The ERG noted that this analysis is limited however, feels the estimate is a more realistic reflection of real-world practice than the estimate used by the company.

4.2. ERG base case 1 – Myeloma XI data

At all previous stages of this appraisal the ERG has preferred the use of the Myeloma XI data only to inform the cost-effectiveness model. Assumptions for survival extrapolations remained the same as at TE stage, with OS extrapolation performed by a joint log-logistic model and PFS by a joint Weibull model.

Table 7 presents the ERG’s preferred assumptions and the cumulative impact on the ICER from the company’s (ERG part-corrected) revised base case.

Table 7: ERG’s preferred model assumptions - Base case 1 (company PAS settings)

Preferred assumption	Section in ERG report or response to ACD	Cumulative ICER £/QALY*
Company revised base case	<i>Response to ACD: Section 3</i>	██████
Set clinical data source to Myeloma XI	<i>Response to ACD: Section 3</i>	██████
Set OS curve to joint log-logistic	<i>Report: Section 4.2.6.1</i>	██████
Set PFS curve to joint Weibull	<i>Report Section 4.2.6.2</i>	██████
ERG’s preferred subsequent treatment settings	<i>Response to ACD: Section 2.7</i>	██████
Set RDI for lenalidomide maintenance to 91.62%	<i>Response to ACD: Section 2.8</i>	██████

Abbreviations: ACD, Appraisal Committee Document; ERG, Evidence Review Group; ICER, incremental cost-effectiveness ratio; OS, overall survival; PFS, progression-free survival; QALY, quality adjusted life year; RDI, relative dose intensity; TE, Technical Engagement.

Note: * Errors were found with the company’s base case result estimate (detailed in Section 3). The ERG has corrected the proportion assigned 10mg of lenalidomide maintenance.

4.3. ERG base case 2 – Pooled Myeloma XI and PSW adjusted CALGB 100104 data

In this base case the ERG presented the same base case assumptions using the Myeloma XI data followed by PSW adjusted CALGB 100104 data (without the ‘response to ASCT’ term) as the clinical data source instead of the Myeloma XI only data. Based on the information provided in Figure 12 of the company’s ACD Addendum, the ERG chose the log-logistic curve to extrapolate OS. As no information for PFS was provided to the ERG, the Weibull curve was implemented (in line with the ERG’s previous base case preferred assumptions). The treatment effect applied to the extrapolations follows the Myeloma XI data for the first 60 months, with the PSW adjusted CALGB 100104 (without the ‘response to ASCT’ term) followed thereafter.

Table 8 presented the ERG’s preferred assumptions and the cumulative impact on the ICER from the company’s (ERG part-corrected) revised base case.

Table 8: ERG's preferred model assumptions - Base case 2 (company PAS settings)

Preferred assumption	Section in ERG report or response to ACD	Cumulative ICER £/QALY*
Company revised base case	<i>Response to ACD: Section 3</i>	██████
Set clinical data source to Myeloma XI followed by PSW adjusted CALGB without response to ASCT term	<i>Response to ACD: Section 3</i>	██████
Set OS curve to joint log-logistic	<i>Report: Section 4.2.6.1</i>	██████
Set PFS curve to joint Weibull	<i>Report Section 4.2.6.2</i>	██████
ERG's preferred subsequent treatment settings	<i>Response to ACD: Section 2.7</i>	██████
Set RDI for lenalidomide maintenance to 91.62%	<i>Response to ACD: Section 2.8</i>	██████

Abbreviations: ACD, Appraisal Committee Document; ERG, Evidence Review Group; ICER, incremental cost-effectiveness ratio; OS, overall survival; PFS, progression-free survival; QALY, quality adjusted life year; RDI, relative dose intensity; TE, Technical Engagement.

Note: * Errors were found with the company's base case result estimate (detailed in Section 3). The ERG has corrected the proportion assigned 10mg of lenalidomide maintenance; ** Company's base case data source applied so no change vs previous, hence no change in ICER

4.3.1. Comparison of the company and ERG base case analyses

A comparison of the company base case with the two ERG base case analyses is provided in Table 9. In addition, the ERG partly-corrected version of the company's base case (unrounded proportion assigned to receive 10 mg of lenalidomide) is also provided.

Table 9: Comparison of company and ERG base cases (company PAS settings)

Arm	Total			Incremental			ICER (£/QALY)
	Costs (£)	LYs	QALYs	Costs (£)	LYs	QALYs	
Company revised base case (deterministic)							
Observation	████	██	██				
Lenalidomide	████	██	██	████	██	██	████
Company revised base case – ERG partly corrected (deterministic)							
Observation	████	██	██				
Lenalidomide	████	██	██	████	██	██	████
ERG revised base case – Myeloma XI only (deterministic)							
Observation	████	██	██				
Lenalidomide	████	██	██	████	██	██	████
ERG revised base case – Pooled adjusted Myeloma XI and CALGB 100104* (deterministic)							
Observation	████	██	██				
Lenalidomide	████	██	██	████	██	██	████

Key: ERG, evidence review group; ICER, incremental cost-effectiveness ratio; LY, life year; QALY, quality adjusted life year.

Note: ERG partly corrected analysis only corrects the proportion of patients assigned to 10mg of lenalidomide maintenance– OS extrapolation remains with a joint generalised gamma model.

* Chosen data source is 'Myeloma XI followed by PSW adjusted CALGB 100104, without 'response to ASCT' term'

4.4. ERG scenario analyses

A range of scenario analyses applied to the company's (partly corrected by ERG) and ERG base case preferred assumptions to explore the assumptions of the key issues in the ACD are provided in Table 10. The three columns represent the impact the scenario has when applied to the company revised base case, the ERG revised base case 1 (see Table 7) and the ERG revised base case 2 (see Table 8). Some scenarios, such as changing the data source, may not be informative as the scenarios apply the preferred OS extrapolations for each base case. For example, the company did not select the joint-generalised gamma for OS in the original submission when it considered Myeloma XI as the data source; however, the ICER in the scenario (Table 10) reflected those settings.

Table 10: Comparison of scenarios on company and ERG preferred assumptions (company PAS settings)

Scenario	ICER (£/QALY)		
	Company *	ERG 1 **	ERG 2 ***
Base-case	████	████	████
<i>Issue 3 - Data source</i>			
Myeloma XI	████	████	████
CALGB	████	████	████
Pooled MXI and CALGB (Myeloma XI prediction)	████	████	████
MXI followed by unadjusted CALGB	████	████	████
MXI followed by CALGB PSW without response to ASCT term	████	████	████
MXI followed by CALGB MAIC without response to ASCT term	████	████	████
MXI followed by CALGB PSW with response to ASCT term	████	████	████
MXI followed by CALGB MAIC with response to ASCT term	████	████	████
<i>Issue 5 – Overall survival extrapolation</i>			
Joint generalised gamma	████	████	████
Joint log-logistic	████	████	████
Joint Weibull	████	████	████
Joint Weibull – MAIC adjusted data (without ‘response to ASCT’ term)	████	████	████
<i>Issue 6 – Treatment waning</i>			
Equal hazard at 10 years	████	████	████
Equal hazard at 20 years	████	████	████
<i>Issue 7 – Subsequent therapy distribution±</i>			
0% patients to receive subsequent therapies	████	████	████
Increase second ASCT probability from 5% to 10% - both arms	████	████	████

5. REFERENCES

1. Palumbo A, Cavallo F, Gay F *et al.* Autologous Transplantation and Maintenance Therapy in Multiple Myeloma. *New England Journal of Medicine* 2014; **371**: 895-905.
2. McCarthy PL, Holstein SA, Petrucci MT *et al.* Lenalidomide Maintenance After Autologous Stem-Cell Transplantation in Newly Diagnosed Multiple Myeloma: A Meta-Analysis. *J Clin Oncol* 2017; **35**: 3279-3289.
3. Phillippo DM, Dias S, Ades AE *et al.* Assessing the performance of population adjustment methods for anchored indirect comparisons: A simulation study. *Statistics in Medicine* **n/a**.
4. National Institute for Health and Care Excellence. *Guide to the methods of technology appraisal*. 2013. <https://www.nice.org.uk/process/pmg9/chapter/foreword>
5. Latimer NR, Abrams KR. *Adjusting survival time estimates in the presence of treatment switching (TSD16)*. Sheffield: 2014. <http://nicedsu.org.uk/technical-support-documents/treatment-switching-tds/>
6. National Institute for Health and Care Excellence. *Carfilzomib for previously treated multiple myeloma: Technology appraisal guidance [TA457]*. London: 2017. <https://www.nice.org.uk/guidance/ta457>

National Institute for Health and Clinical Excellence

Centre for Health Technology Evaluation

Pro-forma Response

ERG report

**Lenalidomide for the maintenance treatment of newly
diagnosed multiple myeloma after autologous stem cell
transplantation [ID475]**

Page	ERG comment	Factual error	ERG response
3	<p>The company pointed out that some patients randomised to ASCT did not subsequently receive it or failed it. The ERG notes that Figure 1 of Palumbo <i>et al.</i> (2014)¹ indicates these numbers are small (141 patients were randomised to ASCT of whom six discontinued), and understands those discontinuing would have been excluded from the estimate in McCarthy <i>et al.</i> (2017),² which analysed known ASCT patients.</p>	<p>The ERG argues that we consider GIMEMA invalid because some patients did not receive ASCT. This is incorrect.</p> <p>We consider GIMEMA invalid because the maintenance group includes people randomised to maintenance and did not receive maintenance. The McCarthy meta-analysis (2017) remains flawed, as it included patients in the ASCT + maintenance arm that did not receive ASCT, maintenance or neither.</p> <p>Contrarily to what stated by the ERG, there were 25 patients who either failed induction, and / or failed ASCT and / or failed to remain eligible for treatment in the maintenance vs no maintenance comparison, 18% of the initially randomised cohort (See calculations and Figure below). The ERG calculation fails to include the 19 participants who discontinued between consolidation and maintenance (i.e. who did not receive allocated treatment) or relapsed after ASCT or were withdrawn from treatment before maintenance / no maintenance.</p> <p>Contrarily to what stated by the ERG, these subjects were not excluded from McCarthy et al (2017). McCarthy excluded those who failed consolidation prior to melphalan high dose therapy and ASCT, but</p>	<p>This is not a factual inaccuracy. The ERG has presented its view on the relevance of the GIMEMA trial in previous documents; the company response does not change that view.</p>

		<p>not those that did not receive maintenance (despite being randomised to it). Specifically, there were 10/68 patients in the maintenance group received no maintenance. (McCarthy et al, 2017, Table 1) who remained in the analysis.</p> <p>We also stated specifically the bias deriving from this design should be interpreted as 'dilution' i.e. a treatment effect would be lower than that calculated using data for people who received treatment.</p> <p>The dilution bias was driven 10/68 people who did not receive maintenance in the maintenance group. this proportion is a number large enough to bias the estimate.</p> <p>We acknowledge that not all the reasons stated for not receiving maintenance may be described in Table 9 of our response; a more comprehensive description would be "3. Patients who responded to induction, failed ASCT, and received no maintenance, or received no maintenance for any other reason, including post-ASCT progression".</p> <p>Nevertheless, the substantial fact remains that these 19 patients were not on the allocated treatment when the maintenance/no maintenance comparison started.</p> <p><u>Calculation</u></p>	
--	--	--	--

		<p>As explained in detail in our Response (Table 9) the GIMEMA ASCT + maintenance and ASCT + no maintenance groups include the following:</p> <ol style="list-style-type: none">1. Patients who succeeded induction, succeeded ASCT and received maintenance2. Patients who failed induction, received no ASCT, received no maintenance3. Patients who responded to induction, failed ASCT, and received no maintenance <p>The graph below provides a summary of attrition from the initial randomisation to the actual receipt of maintenance / no maintenance.</p> <p>25 people were discontinued before maintenance/ no maintenance. This is 18% of the total number of participants allocated to the ASCT +/- maintenance comparison.</p> <p>As the ERG correctly identified, all but one of the 135 people from the Palumbo et al (2014) study were included in the McCarthy analysis (n=134, McCarthy et al 2017, Table 1). This means that McCarthy data clearly do include all people in GIMEMA that were withdrawn any time after melphalan and before the time when the maintenance phase in GIMEMA started.</p>	
--	--	--	--

		<p>Reasons for attrition are stated in Palumbo, Figure 1: received other treatment (n=4); protocol violation (n=1); progressed before maintenance / no maintenance (n=3); withdrew consent (therefore not receiving any interventions after melphalan which could be ASCT or maintenance) (n=8); did not adhere to maintenance protocol (n=1); had toxic effects (n=1 from melphalan as in the no maintenance group there was no other pharmacological treatment); and lost to follow-up (n=1).</p> <p>These reasons clearly support the interpretation that 10 people in the maintenance group received no maintenance.</p>	
	<p>Based on the company's base case analysis (using adjusted CALGB curves to inform survival), and the revised subsequent therapy estimates, the ERG questioned the appropriateness of applying extrapolations for the observation arm, where the benefit of subsequent lenalidomide is removed, yet costs are accrued associated with subsequent lenalidomide.</p>	<p>The ERG incorrectly states that the benefits of subsequent lenalidomide use are removed.</p> <p>The RPSFTM adjusts for lenalidomide used in maintenance (i.e. pre-progression) but not for lenalidomide used as subsequent therapy, therefore post-progression. The RPSFTM correction applies to people who received lenalidomide as maintenance; because of the non-retreat rule, these people could not receive lenalidomide as subsequent therapy; conversely, the cases that remained eligible for lenalidomide in later lines were not adjusted.</p> <p>This is clearly explained in the ACD response, Page 58: "The RPSFTM/IPE methods assume that the experimental treatment effect is the same</p>	<p>Based on the information provided by the company, the ERG do not consider this to be a factual inaccuracy. The information provided by the company in their FAC response still does not full clarify the methods used by the company to account for treatment switching. The ERG have therefore not changed their view.</p> <p>The ERG understands the company</p>

		<p>regardless of when it is administered; that is, the treatment effect of lenalidomide is the same for patients randomised directly to lenalidomide as for those who switching from placebo to lenalidomide. As such, only patients who received lenalidomide prior to investigator assessed progression (N=76) have had their survival times adjusted using the methodology described in this Section. Patients who receive lenalidomide post PD and/or as combination therapy do not have their survival times adjusted. “</p>	<p>chose a ‘treatment-group’ approach to RPSFTM as their base case, and the company describe (addendum p.61-2) the formation of time assigned to lenalidomide (T_L):</p> <p><i>“Patients randomised to placebo who subsequently switching to lenalidomide prior to PD had all time until death (or censoring) following first dose of lenalidomide assigned to T_L.”</i></p> <p>The company explained (addendum p60) that the RPFSTM adjusts T_L so that all counterfactual survival times are equalised across randomised groups.</p> <p>In their FAC response, the company draw attention to the following (addendum p58):</p> <p><i>“Patients who receive lenalidomide post PD and/or as combination</i></p>
--	--	---	--

			<p>therapy do not have their survival times adjusted”</p> <p>The ERG finds the assignment and adjustment of T_L in the base case analysis (the treatment group approach) confusing.</p> <p>Based on the ERG’s understanding of the company’s response to the ACD Addendum, the ‘treatment-group’ adjusts all time from switching, and the ‘on-treatment’ approach adjusts for the time spent receiving lenalidomide maintenance only. Based on the company’s selection of the ‘treatment-group’ approach and the description provided by the company in the FAC response, the ERG is unclear on the difference between the two approaches.</p>
--	--	--	--

			<p>The ERG notes that 33% of placebo patients who switched to lenalidomide maintenance received lenalidomide+dexamethasone at the second line in CALGB (Table 28 of the company's response to ACD Addendum), with third line subsequent therapies not reported. Therefore, in the context of subsequent treatments, the ERG interpret that if the 'treatment-group' RPFSTM approach had been undertaken (all survival time adjusted from point of switching) then the effect of lenalidomide as a subsequent therapy at the second line (and third line) would have been adjusted for. Despite this, the costs of lenalidomide as a subsequent therapy (in the observation arm) will still have been incurred due to the assigned use of lenalidomide+dexametha</p>
--	--	--	---

			<p>some as a subsequent therapy (at the second and third lines) in the company's base case assumptions.</p>
<p>Page 22</p>	<p>The company noted they found the ICER to remain cost effective when using their revised base case estimates for subsequent therapies. The company's TE ICER reduced by [REDACTED] when applying the revised base case assumptions (Celgene ACD Addendum version 3 – base case assumptions). However, as mentioned previously, the ERG found discrepancies between the estimates reported in the ACD Addendum and the cost-effectiveness model. Correcting the model to apply the distributions presented in the Addendum results increases the ICER by [REDACTED].</p>	<p>The ERG states that the company base case for subsequent therapies and specifically use of len+dex in second line set to 0%.</p> <p>This is inaccurate as the first version of the ACD response included 'scenarios' and did not specify a base case, as described in the paragraph below. "We also tested a scenario where lenalidomide in second line is given to 10% of people who would not receive maintenance (Table 31). This proportion was chosen to be closely representing the proportions seen in Myeloma XI and CALGB and in alignment with the observed clinical efficacy in the studies; this scenario faithfully reflects the current situation with respect to the use of lenalidomide in second line." (ACD response, page 47)</p> <p>However subsequently a base case was set (Executive Summary) including 10% len+dex in second line for people who receive observation. Therefore, the base case quoted by the ERG is not the base case used in the response. As such, the model should not be corrected. Table 2, and Table 2, Note in the ERG response should be modified accordingly</p>	<p>Based on the information provided by the company in their FAC response, the ERG note that in their ACD response the company adjusted the subsequent treatment distribution in their base case to allow for 10% of patients in the comparator arm to receive len+dex at 2nd line. The company noted this in version 3.0 of their ACD response addendum, whereas previously the ERG had understood the subsequent treatment distribution to be as presented in 'Celgene's revised estimates' in Table 29 of the company's response to ACD Addendum version 2.0.</p>

			The ERG has updated their response to incorporate this change.
27	The company further explained that according to clinical opinion, physicians may take other measures to dose adjustments, by spacing doses or cycles, or mixing doses however no explanation is provided to the clinical opinion which was sought to validate this.	<p>This statement is factually incorrect.</p> <p>The ACD response presents <i>data</i> from the Myeloma XI trial that substantiate patterns of cycles spacing.</p> <p><u>It is unclear why the ERG expresses a preference for clinical opinion when data exist.</u> Clinical validation is useful however it is not a replacement for data, when data are presented.</p> <p>Nevertheless, clinical opinion was sought. The ‘atypical’ cycles (n=174 out of n=10,911) were interpreted by the Myeloma XI PI and are reported in Table 46.</p> <p>We do not see a valid reason for arguing that clinical trial data and specifically investigator reported start dates of each cycle should not be used. These were valid data and were used to calculate the duration of spacing between each treatment cycle.</p> <p>The ERG proceeds to assume that treatment cycles are not spaced, overriding the evidence presented from Myeloma XI.</p>	<p>This is not a factual inaccuracy.</p> <p>The ERG does not prefer clinical opinion over data. The company’s response to ACD Addendum noted that clinical opinion was sought to inform dosing adjustments. The statement made by the ERG refers only to the unknown nature of the clinical validation.</p>

<p>29</p>	<p>On account of this, the ERG proposes a simplified (albeit limited) view of RDI which may be relevant for consideration based on prescribing of packs as opposed to doses received. Based on Table 45 of the response to the ACD where it's is acknowledged the [REDACTED] the corresponding RDI assuming that all patients received 21 days at 10 mg or 21 day at 5 mg would be [REDACTED] calculated as: [REDACTED]</p> <p>Whilst a simplifying assumption has been made, the [REDACTED] accounts for patients receiving a reduced dose of lenalidomide and also allows accountability for the non-linear pricing of lenalidomide.</p>	<p>This statement is erroneous as it assumes that the dose of lenalidomide and price of lenalidomide are in a linear relationship.</p> <p>The ERG computation assumes that the 5mg dose counts half of the 10mg dose when computing compliance (reverting to dose exposure rather than cost, see table 4 in the RDI table in the ERG response) or that one 5mg pack is equivalent to one 10mg pack, when reverting to costs.</p> <p>So, the formula [REDACTED]</p> <p>Is a representation of the dose-exposure (and yet ignoring data for periods when exposure to drug = 0, during treatment breaks), and is not relevant for the computation of costs in the model.</p> <p>It is unclear why a reversal of position is introduced at this stage, moving the discussion back to the pre-technical engagement position, with the cost of lenalidomide now calculated based on dose exposure and not on 'packs'.</p> <p>When correctly computing the 10mg and 5mg as "pack", instead than milligrams, the ratio presented by the ERG is 100%. It is unclear what this calculation clarifies.</p>	<p>This is not a factual inaccuracy.</p> <p>The company have misunderstood the scenario undertaken by the ERG.</p> <p>The ERG's scenario does account for the non-linear pricing of lenalidomide. The ERG's calculations account for the number of packs prescribed at 10mg and 5mg for Myeloma XI and have used this to directly inform the number of packs costed for in the modelling. Using data from MXI, the ERG's calculation takes account of a proportion of patients being prescribed 5mg (which is below the 10mg dose under consideration as part of this appraisal).</p>
-----------	--	---	---

When the price is incorporated in this formula, the price applied in the model is [REDACTED] of the cost of treatment with the 10mg lenalidomide only

[REDACTED]

(£3,780 and £3,570 are the list prices of 10mg and 5mg lenalidomide respectively).

Therefore, the adjustment made on the dose scale de facto hides the assumption that all drug is costed at 10mg and removes the 5mg price and use in the model computation.

The statement that the formula correctly accounts for the non linearity in price is therefore inaccurate.

This approach is both a logical error and a biased evidence selection, as it overrides the overwhelming evidence (data, clinical trial protocols, label, clinical opinion) that the 5mg dose is used in practice.

The impact of this erroneous calculation is that the model is constrained to 100% compliance with label, so the costing implemented by the ERG assumes the 10mg dose only and a strict 21 doses/28 days posology. This, again, is contrary to all evidence and clinical opinion.

The application of these values in the model also predicts incorrect drug costs.

		<p>This is because the model uses a cost per cycle, whilst data from Myeloma XI are observed over the total time to progression.</p> <p>Because the ERG discarded evidence of cycles spacing from Myeloma XI, the amount of drug consumed in Myeloma XI is apportioned to a much shorter time on treatment, equivalent to the observed number of cycles x 28 days. Yet the time on treatment in the model remains as observed from Myeloma XI.</p> <p>Consequently,</p> <ol style="list-style-type: none">1. The cost per cycle applied in the model is inflated2. The cost of active treatment is applied to the model for the periods between cycles during which (as per Myeloma XI data) patients receive no drugs as they are on treatment breaks. <p>The application of these erroneous estimates in the model substantially inflate the cost of drug, as a result of this formulae not taking the time off drug into account, that the ERG finds 'unexplicable'.</p> <p><u>Calculations</u></p> <p>To clarify the error, we report here a very simple calculation, accessible to anyone without specialist modelling knowledge.</p>	
--	--	--	--

		<p>For a hypothetical patient that receives 9x 10mg cycles and 2x 5mg cycles of maintenance per year, the ERG formula returns:</p> <p>RDI: $(10 \times 9 + 5 \times 2) / (11 \times 10) = 90.9\%$ compliance Number of packs: $(9+2)/11 = 100\%$ of dose prescribed</p> <p>Applying this number to a model where the patient cycles for 1 year in 28 days cycles,</p> <p>Number of cycles received: $365/28 = 13.04$ (rounded) Number of packs received: $13.04 \times 100\% = 13.04$</p> <p>Therefore, a patient observed to receive a total of 11 cycles in the data (9 @10mg and 2 @5mg) is modelled to receive 13 packs.</p> <p>When applying the non linear pricing, 1. Observed cost: $9 \times £3,780 + 2 \times £3,570 = £41,160$ 2. Modelled cost: $13.04 \times £3,780 \times 90.9\% = £44,795$, approx £3,600 larger than the true cost.</p> <p>A cost per cycle of £3,436 is then extrapolated over the total time on treatment in the model, as a result of ignoring treatment free periods.</p> <p>These two factors cause a miscalculated and incorrect explosion of drug costs in the model.</p> <p>The correct calculation instead is as follows: Average cycle duration: $365 / 11 = 33.2$ days</p>	
--	--	---	--

		<p>Total time on 10mg cycles: $33.2 \times 9 = 298.6$ days Total time on 5mg cycles: $33.2 \times 2 = 66.3$ days</p> <p>Total number of cycles in model: as above, 13.04 Expected (per label) cycles @ 10mg: $13.04 \times 9/11 = 10.67$ Expected (per label) cycles @ 5mg: $13.04 \times 2/11 = 2.37$</p> <p>Average time on treatment / off treatment: 28 days / 33.4 days = 84.38%</p> <p>Modelled cost per year:</p> $\frac{10.67 \text{ cycles} \times \text{£}3,780 + 2.37 \text{ cycles} \times \text{£}3,570}{13.04} \times 84.38\% = \text{£}3,157.48$ <p>In this formula, the total drug costs from Myeloma XI is adjusted by the total time spent before progression. This is exactly the correction applied in the company's calculation of RDI.</p>	
30	<p>28-days scenario</p> <p>While the company have provided this as scenario analysis, it does not inform their revised base case assumptions</p>	<p>The 28 days scenario was presented as part of the tabulation included in the 'Executive summary' document, where the changes in incremental ratio are provided 'incrementally' - proceeding step by step, with all assumptions clarified. The table was submitted at a later point in time with respect to the ACD response.</p> <p>Each scenario proceeds from the previous (or as indicated) hence taking all preceding changes into account. It is unclear whether the comment of the</p>	<p>This is not a factual inaccuracy.</p> <p>The statement made by the ERG that the 28-day dosing scenario does not inform the company's revised base case assumptions is correct.</p>

		<p>ERG is pertinent to the information submitted by the company or just to the initial ACD response.</p> <p>The following scenarios were submitted:</p> <ol style="list-style-type: none"> 1. A model scenario reflecting a 1-day to 28-day lenalidomide treatment regimen. In this scenario, we accounted for 28 pills instead than 21, all other assumptions in the base case being equal (Scenario 7) 2. Assumed committee preferred scenario (with treatment waning, 5% of people having second ASCT and treatment given on every day of 28 day cycle) - including all other assumptions as per table in executive summary (Scenario 2b, 4, 5a, 7 combined) 3. Assumed committee preferred scenario (with treatment waning, 10% of people having second ASCT and treatment given on every day of 28 day cycle) including all other assumptions as per table in executive summary (Scenario 2b, 4, 5b, 7 combined) 	<p>The ERG received two versions of the model and three versions of an Addendum post ACD. The third version of the Addendum was a 3-page document containing the revised company base case and the effect on the ICER of various scenarios requested by the Committee. This document contains a scenario that explores the impact of a 28-day dosing regimen, but the dosing regimen does not form a component in the company's revised base case.</p> <p>To confirm, the comments made by the ERG in response to the company's ACD are based on Addendum version 2, Addendum version 3 and model version 2.</p>
30	<p>The company has not explicitly stated how this scenario was implemented within the model. It was the ERG's understanding that only the dosing was changed with RDI.</p>	<p>An explicit statement indicates that the 28 days dose scenario was calculated using the cost of 28 doses for each cycle, all other things being equal. (Executive Summary, Scenario 7)</p>	<p>Based on the information that the ERG has received from the company, this is not a factual inaccuracy. The ERG are unclear what</p>

		<p>This scenario was not updated in the ACD response.</p> <p>No changes to the RDI were made (as the RDI is driven by treatment spacing).</p>	<p>'Executive Summary' document the company refer to, however the statement with the phrase "<i>all other things being equal</i>" does not appear in any of the documents submitted by the company to the ERG in its response to the ACD.</p> <p>This response from the company suggests that the ERG's assumption was correct. The company appear to suggest that the ERG can remove any uncertainty about this; however, the ERG have not updated this sentence as a full explanation of the methods used by the company has not been provided to the ERG.</p>
Cover letter	Unrounded proportion of people receiving the 10 mg dose	The value this sentence refers to is not stated, we are not able to comment.	<p>This is not a factual inaccuracy.</p> <p>Page 33 of the ERGs response to ACD states 'The proportion of patients assigned to</p>

			receive lenalidomide maintenance 10 mg dose was rounded to 1 decimal place'. The proportion of patients assigned 10mg of lenalidomide maintenance can be seen in Cost data, cell J16 and is 83.2%.
Cover letter	Aligning the model subsequent therapy distributions with those reported in your Addendum, see section 3 of ERG critique	A response regarding the base case subsequent distributions is provided in this document	This is addressed above.