

Pomalidomide with Dexamethasone for Treating Relapsed and Refractory Multiple Myeloma Previously Treated with Lenalidomide and Bortezomib: An Evidence Review Group Perspective of an NICE Single Technology Appraisal

Nasuh C. Büyükkaramikli¹  · Saskia de Groot¹ · Debra Fayter² · Robert Wolff² · Nigel Armstrong² · Lisa Stirk² · Gill Worthy² · Fernando Albuquerque de Almeida¹ · Jos Kleijnen^{2,3} · Maiwenn J. Al¹

Published online: 31 October 2017

© The Author(s) 2017. This article is an open access publication

Abstract The National Institute for Health and Care Excellence (NICE), as part of the institute's single technology appraisal (STA) process, invited the manufacturer of pomalidomide (POM; Imnovid[®], Celgene) to submit evidence regarding the clinical and cost effectiveness of the drug in combination with dexamethasone (POM + LoDEX) for the treatment of relapsed and refractory multiple myeloma (RRMM) after at least two regimens including lenalidomide (LEN) and bortezomib (BOR). Kleijnen Systematic Reviews Ltd (KSR) and Erasmus University Rotterdam were commissioned as the Evidence Review Group (ERG) for this submission. The ERG reviewed the evidence submitted by the manufacturer, validated the manufacturer's decision analytic model, and conducted exploratory analyses in order to assess the robustness and validity of the presented clinical and cost-effectiveness results. This paper describes the company submission, the ERG assessment, and NICE's subsequent decisions. The company conducted a systematic review to identify studies comparing POM with comparators outlined in the NICE scope: panobinostat with bortezomib and dexamethasone (PANO + BOR + DEX), bendamustine with thalidomide and dexamethasone (BTD) and conventional chemotherapy (CC). The main clinical effectiveness

evidence was obtained from MM-003, a randomized controlled trial (RCT) comparing POM + LoDEX with high-dose dexamethasone (HiDEX; used as a proxy for CC). Additional data from other studies were also used as non-randomized observational data sources for the indirect treatment comparison of POM + LoDEX with BTD and PANO + BOR + DEX. Covariate or treatment switching adjustment methods were used for each comparison. The model developed in Microsoft[®] Excel 2010 using a semi-Markov partitioned survival structure, submitted in the original submission to NICE for TA338, was adapted for the present assessment of the cost effectiveness of POM + LoDEX. Updated evidence from the clinical-effectiveness part was used for the survival modelling of progression-free survival and overall survival. For POM + LoDEX, the patient access scheme (PAS) discount was applied to the POM price. Three separate comparisons were conducted for each comparator, each comparison using a different dataset and adjustment methods. The ERG identified and corrected some errors, and the corrected incremental cost-effectiveness ratios (ICERs) for POM + LoDEX versus each comparator were presented: approximately £45,000 per quality-adjusted life-year (QALY) gained versus BTD, savings of approximately £143,000 per QALY lost versus PANO + BOR + DEX, and approximately £49,000 per QALY gained versus CC. The ERG also conducted full incremental analyses, which revealed that CC, POM + LoDEX and PANO + BOR + DEX were on the cost-effectiveness frontier. The committee's decision on the technology under analysis deemed that POM + LoDEX should be recommended as an option for treating multiple myeloma in adults at third or subsequent relapse of treatments including both LEN and BOR, contingent on the company providing POM with the discount agreed in the PAS.

✉ Nasuh C. Büyükkaramikli
buyukkaramikli@imta.eur.nl

¹ Institute for Medical Technology Assessment (iMTA), Institute of Health Policy and Management (iBMG), Erasmus University Rotterdam, Rotterdam, The Netherlands

² Kleijnen Systematic Reviews Ltd, York, UK

³ Department of Family Medicine, School for Public Health and Primary Care (CAPHRI), Maastricht University, Maastricht, The Netherlands

Key Points for Decision Makers

Non-randomized evidence can be used to assess comparative effectiveness and inform cost-effectiveness analyses in the absence of randomized evidence. However, reviewers and decision makers should pay special attention to the adjustment methods used for accounting for differences between patient populations at baseline as they may have considerable impact on effectiveness and cost-effectiveness results.

Treatment recommendations can be changed in time, based on additional clinical and cost-effectiveness evidence and changes in the clinical landscape, as well as in prices.

Sometimes, the end-of-life criteria might not hold for all of the comparators. In such a situation, the decision maker might base the recommendation according to the end-of-life threshold for the comparisons that satisfy the end-of-life criteria. For the other comparisons, the decision maker might instead apply the standard threshold to the ICERs calculated (e.g. 'costs per QALY gained' or 'costs saved per QALY lost').

1 Introduction

The National Institute for Health and Care Excellence (NICE) is an independent organization providing national guidance on promoting good health and preventing and treating ill health [1]. The single technology appraisal (STA) process is designed to provide recommendations on a single product, device or other technology with a single indication. The process covers new technologies and enables NICE to produce guidance shortly after the technology is introduced into the UK. The NICE Appraisal Committee (AC) obtains relevant evidence from several sources: the company submission (CS), a report from the appointed independent Evidence Review Group (ERG) and advice from consultees (i.e. patients, experts and other stakeholders). The CS includes a written report and a decision analytic model that describes the clinical and cost effectiveness of the technology under investigation. The ERG, an external organization independent of the NICE, reviews the CS and produces a summary report and provides a critique of the submitted evidence. After consideration of all the relevant evidence, the AC formulates

preliminary guidance, in the form of the Appraisal Consultation Document (ACD), as to whether or not to recommend the intervention. The stakeholders are invited to comment on this ACD and the submitted evidence. A subsequent ACD may be produced or a Final Appraisal Determination (FAD) issued. Once published, NICE technology guidance provides a legal obligation for National Health Service (NHS) providers to reimburse technologies that have been approved. This paper presents a summary of the ERG report and the development of NICE guidance based on the findings of the AC for the STA of pomalidomide (POM) in combination with low-dose dexamethasone (DEX), for treating relapsed and refractory multiple myeloma (RRMM) after at least two regimens including lenalidomide (LEN) and bortezomib (BOR). Full details of all relevant appraisal documents can be found on the NICE website [2.]

2 The Decision Problem

The underlying indication of this appraisal is RRMM. Multiple myeloma is a rare, incurable malignant haematological disease arising from the monoclonal expansion of plasma cells in the bone marrow [3]. It represents approximately 1% of all incident cancers globally and results in more than 43,000 deaths annually worldwide. Multiple myeloma is primarily a disease of the elderly, and the median age at diagnosis ranges from 69 to 74 years [4–6]. Patients suffer from a range of debilitating symptoms, including bone pain and damage including fractures, mobility problems, anaemia and general ill health [7–9].

The course of the disease is characterised by cycles of remission and relapse [10]. With increasing lines of therapy, there is a decreasing duration of response and, ultimately, development of refractory disease in combination with greater symptomatic burden [10–12]. The main aims of therapy are to control disease, maximise health-related quality of life (HRQoL) and prolong survival [10–13]. Despite the availability of MM agents such as thalidomide (THAL), BOR and LEN, prognosis remains bleak, especially for patients who relapse or are refractory to these agents, with a median survival of 3–9 months [14–18.]

The clinical pathway of care for multiple myeloma patients differs between transplant-eligible and transplant-ineligible patients, especially in the choice of the first-line therapy. For transplant-eligible patients, NICE guidance TA311 recommends BOR + DEX or BOR + THAL + DEX induction therapy followed by autologous stem cell transplant (ASCT) [19], whereas for transplant-ineligible patients, THAL in combination with an alkylating agent (e.g. melphalan) and a corticosteroid (e.g. prednisone) is recommended. However, for THAL-contraindicated

patients, THAL may be replaced with BOR (NICE guidance TA228 [20]).

In the second line, BOR monotherapy is recommended (TA129 [21]) as a treatment option for people who are at their first relapse, with a patient access scheme (PAS). Note that even though LEN is not recommended by NICE, patients might have received LEN in the first two lines of therapy as part of clinical trials, as well as through previous funding from the Cancer Drugs Fund (CDF). From the third line onwards, LEN + DEX has been recommended with the condition of a maximum of 26 cycles (TA171 [22]), while PANO + BOR + DEX has been recommended as an option for treating “adult patients with relapsed and/or refractory multiple myeloma who have received ≥ 2 prior regimens including BOR and an immunomodulatory agent, when the company provides PANO with the discount agreed in the patient access scheme” (TA380) [23]. Besides the recommended treatments above, conventional chemotherapy (CC) and bendamustine (BEN) combinations (e.g. BEN in combination with THAL and DEX-BTD) are commonly used in UK clinical practice, with the latter being funded via the CDF [24].

POM (Imnovid[®]) is an oral immunomodulatory agent for RRMM. POM in combination with DEX has a UK marketing authorisation for the following indication: “pomalidomide in combination with low dose dexamethasone for the treatment of adult patients with relapsed and refractory multiple myeloma who have received at least two prior treatment regimens, including both lenalidomide and bortezomib, and have demonstrated disease progression on the last therapy” [25].

Following a previous appraisal in 2015, NICE did not recommend POM + LoDEX within its market authorisation (TA338). The agent was re-evaluated in the current submission following a new PAS price for the technology and additional clinical and cost-effectiveness evidence. The remit of this appraisal was specified by NICE’s final scope [26], which was to assess the clinical and cost effectiveness of POM + LoDEX within its licensed indication. The proposed positioning of POM + LoDEX is from third-line treatment onwards. For patients who have had two prior therapies, the comparators listed in the scope were PANO + BOR + DEX, whereas for patients who have had three or more prior therapies the comparators were CC, BTD and PANO + BOR + DEX.

3 The Independent Evidence Review Group (ERG) Report

Kleijnen Systematic Reviews Ltd (KSR), in collaboration with Erasmus University Rotterdam, acted as the ERG, and reviewed the evidence on the clinical and cost effectiveness

of POM + LoDEX for the treatment of adult patients with RRMM who have received at least two prior treatment regimens, including both LEN and BOR, and have demonstrated disease progression on the last therapy as submitted by the company (Celgene).

The review embodied three aims:

- to assess whether the CS conformed to the methodological guidelines issued by NICE [1];
- to assess whether the company’s interpretation and analysis of the evidence were appropriate;
- to indicate the presence of other sources of evidence or alternative interpretations of the evidence that could inform NICE guidance.

The ERG critically reviewed the evidence in the CS, in the response to clarification questions, and evidence provided after publication of the ACD. Furthermore, it conducted additional searches, explored the impact of assumptions on the incremental cost-effectiveness ratio (ICER), revised the economic model and explored additional scenario analyses.

3.1 Summary of the Clinical Evidence

The company conducted a systematic review to inform the submission. The main direct evidence was taken from the MM-003 randomised controlled trial, which was the only trial that compared POM + LoDEX with any of the comparators listed in the final scope (i.e. HiDEX; used as a proxy for CC). Individual patient data (IPD) from patients receiving POM + LoDEX in the MM-002, MM-003 and MM-010 studies, and from the MUK-One [27], Gooding et al. [15] and Tarant et al. [16] studies of BEN formed the source of the comparison between POM + LoDEX and BTD. For the comparison of POM + LoDEX versus PANO + BOR + DEX, IPD from the POM + LoDEX arms of the MM-003 [28], MM-002 [29] and MM-010 [30] studies, as well as the aggregate data from the single-arm PANORAMA-2 study, were used. An overview of the trials used can be seen in Table 1.

3.1.1 Direct Evidence

The MM-003 randomised controlled trial was the only study that directly compared POM + LoDEX with any of the comparators listed in the final scope [26]. This trial compared POM + LoDEX (POM 4 mg/day on days 1–21 and DEX 40 mg each week in a 4-week cycle) with high-dose dexamethasone (HiDEX; DEX 40 mg on days 1–4, 9–12 and 17–20 of a 4-week cycle). The company considered HiDEX as a proxy for CC. MM-003 included 455 participants and was an open-label, multinational trial including participants recruited in 93 study sites, 68 of

Table 1 Eligible studies included in the quantitative analysis

Intervention	Study	IPD	Description
POM + LoDEX	MM-003 [28]	Available	Phase III, RCT ($n = 455$, 2:1) comparing POM + LoDEX vs. HiDEX in RRMM patients
	MM-002 [29]	Available	Phase II, RCT ($n = 221$, 1:1) comparing POM + LoDEX vs. POM in RRMM patients
	MM-010 [30]	Available	Phase IIIb, single-arm ($n = 682$), open-label study of POM + LoDEX in RRMM patients
BEN	MUK-One [27]	Available	RCT ($n = 95$) comparing different doses of BEN (100 mg/m^2 vs. 60 mg/m^2) of two BTD regimens applied to RRMM patients
	Tarant et al. [16]	Available	Retrospective study ($n = 55$) on the effectiveness and safety of various therapies (including BTD) in RRMM patients
	Gooding et al. [15]	Available	Retrospective study ($n = 39$) on the effectiveness and safety of various therapies (including BTD) in RRMM patients
PANO+BOR+DEX	PANORAMA-2 [31]	Not available	Two-stage, single-arm, open-label study ($n = 55$) on the effectiveness and safety of PANO + BOR + DEX in RRMM patients

POM pomalidomide, LoDEX low-dose dexamethasone, BEN bendamustine, PANO panobinostat, BOR bortezomib, DEX dexamethasone, RCT randomised controlled trial, RRMM relapsed and refractory multiple myeloma, BTD bendamustine + thalidomide + dexamethasone, IPD individual patient-level data, HiDEX high-dose dexamethasone

which are located in Europe. The number of centres located in the UK and number of patients recruited in the UK was unclear.

Using the latest data cut-off of MM-003 (1 September 2013, investigator assessment) at 15.4 months follow-up, there was an increase in median survival with pomalidomide. Overall survival (OS) was significantly better for patients treated with POM + LoDEX compared with those receiving HiDEX [13.1 months vs. 8.1 months; hazard ratio (HR) 0.72, 95% confidence interval (CI) 0.56–0.92]. POM + LoDEX significantly extended progression-free survival (PFS) compared with HiDEX (4 vs. 1.9 months; HR 0.50, 95% CI 0.41–0.62).

Since 56% of patients in the HiDEX arm received subsequent therapy with POM in the MM-003 trial, the company provided crossover-adjusted OS results, which are given in Table 2.

The company found that 247 of 300 patients (82.3%) in the POM + LoDEX group had at least one adverse event (AE) considered by the investigator to be related to POM. Furthermore, 190 patients (63.3%) had grade 3–4 treatment-emergent AEs (TEAEs) considered to be related to POM. However, the company stated that “with dose

modifications and supportive care the safety profile was predictable, manageable and generally well tolerated”. Events occurring more frequently in the POM + LoDEX group included neutropenia (51.3 vs. 20.0% for neutropenia and 9.3 vs. 0% for febrile neutropenia). The main cause of treatment discontinuation was progressive disease, and discontinuations related to AEs were uncommon. There were more dose interruptions in the POM + LoDEX group than in the HiDEX group (67 vs. 30%).

3.1.2 Indirect Evidence

No studies directly compared POM + LoDEX with either BEN or PANO + BOR + DEX. Furthermore, no studies could provide a common comparator to support an indirect comparison or mixed treatment comparison (MTC). As a consequence, the company selected individual treatment arms from the available studies and performed separate analyses comparing POM + LoDEX with each of the comparators independently.

3.1.2.1 Indirect Comparison with Bendamustine Comparison of IPD from patients receiving POM + LoDEX in

Table 2 Intention-to-treat and crossover-adjusted OS results from MM-003 trial

Median OS in months	POM + LoDEX (95% CI)	HiDEX (95% CI)	Difference (HR, 95% CI)
Intent-to-treat, (independent assessment, earlier data cut)	12.7 (10.4–15.5)	8.1 (6.9–10.8)	4.6 (0.74, 0.56–0.97)
Crossover adjustment, two-stage method	12.7 (10.4–15.5)	5.7 (4.2–7.5)	7.0 (0.52, 0.39–0.68)
Crossover adjustment, RPSFTM method	12.7 (10.4–15.5)	6.7 (4.6–10.5)	6.0 (HR 0.49, 0.33–1.00)

Source: Company submission, Table 24 (page 75) [32]

OS overall survival, POM pomalidomide, LoDEX low-dose dexamethasone, HiDEX high-dose dexamethasone, HR hazard ratio, RPSFTM rank-preserving structural failure time model, CI confidence interval

the MM-002 [29], MM-003 [28] and MM-010 [30] studies with IPD from the MUK-One [27], Gooding et al. [15] and Tarant et al. [16] studies of BTD was achieved using Cox proportional hazards regression models to adjust for factors thought to be prognostic of OS and PFS (based on mean covariate values according to the baseline of the pooled trials). The prognostic factors to be included in the regression model in the CS were determined according to a selection procedure based on data availability, information derived from the systematic literature review of prognostic factors in RRMM and clinical expert consultation. Following the selection procedure, the CS reported the following covariates in the final analysis given in Table 3.

The MM-002 trial [29] alone was selected for use for POM + LoDEX within the base-case analysis due to the lower levels of refractoriness exhibited within this trial (78%) compared with the remainder of the POM + LoDEX data (95%). This lower level of refractoriness was considered more comparable with the BTD data (18–25% across sources). As this covariate was identified by clinicians as the most important prognostic factor and is difficult to adjust for with the current datasets (given that the overlap between datasets is low), it was considered more important to select the more comparable dataset for analysis than to retain the maximum number of patients for analysis in the POM + LoDEX arm.

The company conducted three sensitivity analysis. In the first sensitivity analysis, the company incorporated data from all available POM + LoDEX trials (MM-002 [29], MM-003 [28] and MM-010 [30]), while in the second sensitivity analysis, the company incorporated International Staging System (ISS) stage as an additional covariate, on the dataset that includes patients whose ISS stage was available (i.e. the MM-002 [29] and Gooding et al [15.] studies were excluded). The last sensitivity analysis was conducted on the same dataset used in sensitivity analysis 2, but this time ISS stage was not included as a covariate. The unadjusted and covariate-adjusted results from the base-case analysis and sensitivity analyses are given in Table 4.

From the results of the base-case and sensitivity analyses, the company concluded that POM + LoDEX improves OS and PFS compared with BTD, and the inclusion of POM + LoDEX data from the MM-003 [28] and MM-010 [30] trials did not substantially alter the results of the PFS, whereas it reduced the OS benefit of POM + LoDEX compared with the base-case. Furthermore, the company suggested that, based on the results of sensitivity analyses 2 and 3, even though ISS stage is a very important predictive factor of an increased hazard of death and progression, it had little impact on the HR of PFS and OS.

3.1.2.2 Indirect Comparison with PANO + BOR + DEX

For the comparison of POM + LoDEX versus PANO + BOR + DEX, IPD from the POM + LoDEX arms of the MM-003 [28], MM-002 [29] and MM-010 [30] studies, as well as the aggregate data from the single-arm PANORAMA-2 study [31], were used. For this comparison, not all patients in the POM + LoDEX trials, but a subgroup (n = 886), who were refractory to BOR but not primary refractory, was included to align with the PANORAMA-2 population. Since no IPD were available from the PANORAMA-2 study publication [31], covariate adjustment methods used in the comparison with BTD were not possible, and the company used the matching-adjusted indirect comparison (MAIC) method to adjust for the differences in patient characteristics between studies. The MAIC method reweights the patient-level data for POM + LoDEX to reflect a population of similar baseline characteristics as the PANO + BOR + DEX population. The results from the unadjusted and MAIC-adjusted analysis are given in Table 5.

As can be seen, the application of the MAIC method resulted in a 1-month increase in median OS for patients receiving POM + LoDEX compared with the unweighted analysis, and, in both cases, the median OS was shorter than the median OS of patients receiving PANO + BOR + DEX. The application of MAIC seems to have little effect on the median PFS of patients treated with POM + LoDEX compared with the unweighted analysis. These results suggest that POM + LoDEX reduced the risk of

Table 3 Covariates used in the BTD vs. POM + LoDEX indirect comparison

Covariate	Type (e.g. continuous/categorical)
Treatment arm	Categorical (POM + LoDEX or BTD)
Age at the start of the treatment	Continuous
Number of prior treatment lines of therapy	Continuous
Receipt of prior THAL	Categorical (yes/no)
Refractory to LEN	Categorical (yes/no)
ISS stage (in sensitivity analyses)	Categorical (1/2/3)

THAL thalidomide, LEN lenalidomide, ISS International Staging System, POM pomalidomide, LoDEX low-dose dexamethasone, BTD bendamustine + thalidomide + dexamethasone

Table 4 Summary results (OS, PFS) from the POM + LoDEX vs. BTD comparison

Analysis (trial data used in the dataset)		OS		PFS, in months (95% CI)	
		Unadjusted	Adjusted	Unadjusted	Adjusted
Base-case (MM-002 [29], MUK-One [27], Gooding et al. [15], Tarant et al. [16])	Median, in months (95% CI)	POM + LoDEX: 16.5 (12.6–19.8) BTB: 8.1 (5.3–15.5)	POM + LoDEX: 16.6 (12.6–21.3) BTB: 10.5 (5.8–14.8)	POM + LoDEX: 4.2 (3.7–5.8) BTB: 3.3 (2.5–5.5)	POM + LoDEX: 4.7 (3.7–6.6) BTB: 3.7 (2.8–5.6)
	HR (95% CI)	0.55 (0.38–0.81)	0.58 (0.36–0.94)	0.76 (0.56–1.05)	0.79 (0.52–1.22)
	Sensitivity analysis 1 (MM-002 [29], MM-003 [28], MM-010 [30], MUK-One [27], Gooding et al. [15], Tarant et al. [16])	Median, in months (95% CI)	POM + LoDEX: 12.6 (11.6–13.8) BTB: 8.1 (5.3–15.5)	POM + LoDEX: 12.7 (11.9–13.9) BTB: 8.1 (6.1–12.4)	POM + LoDEX: 4.3 (3.9–4.7) BTB: 3.3 (2.5–5.5)
	HR (95% CI)	0.68 (0.51–0.92)	0.64 (0.45–0.91)	0.80 (0.62–1.03)	0.61 (0.45–0.84)
Sensitivity analysis 2 (MM-003 [28], MM-010 [30], MUK-One [27], Tarant et al. [16])	HR (95% CI)	NR	0.72 (0.47–1.11)	NR	0.62 (0.43–0.90)
Sensitivity analysis 3 (MM-003 [28], MM-010 [30], MUK-One [27], Tarant et al. [16])	HR (95% CI)	NR	0.82 (0.54–1.27)	NR	0.62 (0.43–0.89)

OS overall survival, PFS progression-free survival, POM pomalidomide, LoDEX low-dose dexamethasone, CI confidence interval, HR hazard ratio, BTB bendamustine + thalidomide + dexamethasone, NR not reached

Table 5 Summary results (OS, PFS) from the POM + LoDEX vs. PANO + BOR + DEX comparison

	POM + LoDEX	PANO + BOR + DEX
<i>OS</i>		
Median OS, months (95% CI)		
Unweighted	12.4 (11.1–13.4)	17.5 (10.8–22.22)
Weighted	13.4 (11.4–15.6)	NA
HR (95% CI)		
Unweighted		0.73 (0.52–1.02)
Weighted		0.78 (0.56–1.09)
<i>PFS</i>		
Median PFS, months (95% CI)		
Unweighted	4.1 (3.7–4.6)	5.3 (3.9–6.6)
Weighted	4.2 (3.7–4.8)	NA
HR (95% CI)		
Unweighted		1.12 (0.85–1.48)
Weighted		1.18 (0.89–1.56)

Source: Evidence Review Group report, Table 4.27 (page 85) [33]

POM pomalidomide, LoDEX low-dose dexamethasone, PANO panobinostat, BOR bortezomib, DEX dexamethasone, OS overall survival, PFS progression-free survival, HR hazard ratio, CI confidence interval, NA not available

progression but increased the risk of death compared with PANO + BOR + DEX; however, the company noted that the differences were not statistically significant and the

evidence for PANO + BOR + DEX was based on only 55 patients compared with 886 patients receiving POM + LoDEX.

3.1.3 End of Life

The company stated that POM + LoDEX is considered to meet the NICE end-of-life criteria in comparison with BEN and CC. They stated that “The estimated survival benefit compared to BEN and conventional chemotherapy is > 5 months in all comparisons (covariate adjusted and unadjusted, crossover adjusted and unadjusted). Modelled mean survival increase is 7–8 months”.

They further stated that in relation to PANO “Evidence for end of life is less compelling in the comparison to PANO + BOR + DEX as no improvement was demonstrated in median outcomes for OS; difficulties in comparing to PANO + BOR + DEX are, however, considerable given the limited evidence available and lack of patient level data to correct for differences in patient population”.

3.2 Critique of the Clinical Evidence

The systematic reviews conducted by the company (for both the clinical studies and the prognostic covariate selection procedure) were deemed appropriate to the scope

of the submission. Although the ERG identified a number of problems in relation to searching for studies of clinical effectiveness, it was satisfied that the evidence presented in the submission was the best available in this limited area. The ERG agreed that a meta-analysis could not have been conducted as only one trial was deemed directly relevant to the decision problem (MM-003).

Although MM-003, the main trial forming the direct evidence for pomalidomide, was a reasonably large, well-conducted, multicentre trial, the main comparator (HiDEX) is no longer optimal in current practice; therefore, the comparator can only be viewed as a proxy for CC. Additionally, over 50% of patients in the trial were aged 65 years or under, therefore may reflect a younger population than that typically seen in practice. The ERG noted an underrepresentation of non-White participants. Under 1% were of Asian origin and 1.5% were of Black or African American origin. The trial was in a heavily treated population who had received a median of five therapies (range 2–17). Results were presented for only 17 patients receiving two prior therapies, thus the trial is not representative of POM + LoDEX as a third-line therapy. Within these constraints, POM + LoDEX appears to extend OS and PFS in comparison with HiDEX in a heavily treated population who are refractory to BOR and LEN. The AE profile appears to be manageable with appropriate dose reductions and interruptions. However, the slightly higher incidence of serious AEs (grade 3 and 4) attributed to POM + LoDEX was drawn to the attention of the committee, along with the more frequent occurrence of neutropenia.

No studies directly compared POM + LoDEX with either BTD or PANO + BOR + DEX. In addition, the available studies did not include a common comparator that would permit an indirect treatment comparison or MTC. As a result, the company presented evidence based on indirect comparisons. The ERG noted that the covariate adjusted results were very similar to the unadjusted results in terms of both PFS and OS for the base-case and the sensitivity analysis of POM + LoDEX compared with BTD, indicating that the differences between studies in the selected covariates (patient characteristics) have relatively little impact on the outcomes observed. The selection of different datasets for POM + LoDEX alters the results for OS. Results suggested that the survival benefit of POM + LoDEX was less for patients in the MM-003 and MM-010 studies than for patients in the MM-002 study of POM. The ERG is doubtful about the exclusion of the MM-003 and MM-010 trials in the base-case analysis of the BTD versus POM + LoDEX comparison on the basis of higher LEN refractoriness (approximately 20% in BTD trials, 78% in MM-002 and 95% in MM-003 and MM-010).

For the comparison with PANO + BOR + DEX, as in the comparison with BTD, the matching adjustment did not

substantially alter the results. This implies that the differences between studies in the selected covariates have relatively little impact on the outcomes observed. The ERG noted that the number of patients receiving PANO + BOR + DEX was small ($n = 55$) and there were some differences between POM studies and PANORAMA-2 in terms of the number of prior lines of therapy.

The ERG recognised that the lack of all patient-level data might have excluded many of the standard alternatives/methodologies used in the indirect comparison of non-randomised data. However, the ERG considered that some of the choices (e.g. adjustment methods as well as datasets included) were rather arbitrary and the company could have followed a more systematic approach (e.g. following the recommendations in NICE DSU TSD 17 and TSD 18 [34]). Furthermore, there can be unmeasured confounding factors, which could have been quantitatively assessed.

In terms of the end-of-life criteria, the ERG agrees that the patient group, being at least at the third line of treatment for RRMM, have a short life expectancy, normally < 24 months. Hence, the first criterion for end of life has been met. As regard the second criteria, the ERG agrees that POM + LoDEX appear to meet end-of-life criteria of increasing survival in relation to BTD and DEX; however, the evidence suggests that POM + LoDEX does not meet this criterion compared with PANO + BOR + DEX. It is noted though that the evidence for PANO + BOR + DEX is based on a small number of patients ($n = 55$) and the analysis was limited by the lack of studies comparing these treatments.

3.3 Summary of the Cost-Effectiveness Evidence

The company performed an update of the systematic review of cost-effectiveness studies and evidence on healthcare resource use and HRQoL. Furthermore, the key issues raised by the committee from the previous appraisal (TA338) were extracted and are summarised in Table 6.

An electronic model was developed in Microsoft[®] Excel 2010 (Microsoft Corporation, Redmond, WA, USA) using a semi-Markov partitioned survival structure. The objective of the model was to estimate the cost effectiveness of POM + LoDEX when compared with BTD, PANO + BOR + DEX and CC for the treatment of patients with RRMM who were previously treated with LEN and BOR and whose disease progressed during the last therapy.

The model structure presented four health states: a pre-progressive state split into 'on-treatment' and 'off-treatment', a post-progression state (progressive disease), and death. The cycle length of the model was 1 week and the time horizon was 15 years (i.e. lifetime, since virtually every patient in the model died within 15 years). The

Table 6 List of issues raised by the committee in TA338 and the company's description on how these issues are addressed in the current company submission

	Issues in TA338	How the issue is addressed in the current submission, according to the company
Comparative effectiveness data	Very few data were identified for current care	Updated data [i.e. MUK-One study, Gooding et al. study and Tarant et al. study (to inform comparison with BTD) and PANORAMA-2 (to inform comparison with PANO + BOR + DEX)]
Assumptions regarding equivalence of comparators	Assumption of equal effectiveness among all comparators	Comparator-specific data provided to inform comparison with BTD and PANO + BOR + DEX
Relative benefit of current care vs. HiDEX	Concerns around the survival predictions for the current care (lower survival than for HiDEX, which was considered suboptimal treatment)	This unexpected result was no longer seen
Adjustment of trials to provide comparable estimates	Differences between the populations in the studies included for analysis in TA338	Adjustments for differences in patient characteristics were conducted based on a covariate selection process
Adverse events	Cut-off point to include disutility values only for AEs that occurred in more than 2% of patients	No change
Dosing	Exclusion of costs of unused tablets (not necessarily entire pack) of POM + LoDEX due to non-protocol interruptions	Only dose interruptions that would result in an entire pack not being used are assumed not to incur costs

POM pomalidomide, *LoDEX* low-dose dexamethasone, *PANO* panobinostat, *BOR* bortezomib, *DEX* dexamethasone, *BTD* bendamustine + thalidomide + dexamethasone, *AEs* adverse events, *HiDEX* high-dose dexamethasone

model considered an NHS and personal social services perspective, and costs and utilities were discounted using a yearly rate of 3.5%.

The transition probabilities between health states were estimated from the parametric survival functions fitted to OS, PFS and time to treatment discontinuation (TTD) data from the relevant data sources. The data sources and methodological approaches for covariate adjustment for each comparison are summarized in Table 7.

Seven parametric distributions (exponential, log-normal, log-logistic, Gompertz, gamma, extreme value and Weibull) were examined for each clinical outcome (OS, PFS and TTD). The fit of each covariate-adjusted parametric model to the Kaplan–Meier survival data was explored using visual inspection, log-cumulative hazard and Q–Q plots, Akaike/Bayesian information criterion (AIC/BIC) goodness-of-fit statistics and clinical plausibility. In Table 8, the survival analysis methods and the curves chosen for the base-case are summarized.

TEAEs of grade 3/4 were included in the economic model based on a 2% occurrence threshold in the MM-003 trial dataset. For BTD and PANO + BOR + DEX, in the base-case, the AE rates of POM + LoDEX were multiplied by risk ratios for treatment discontinuation due to a TEAE (under the comparator vs. under POM + LoDEX). In a scenario analysis, instead of risk ratios, safety scores provided by an advisory board were used.

Utilities for each health state were found using a regression model based on EQ-5D data from the MM-003 trial. While many covariates were assumed to be the same between treatments, treatment-specific utilities were obtained by using treatment-specific values for the following covariates: disease progression, best overall response (applied after week 12, maintained lifelong), hospitalisations and AEs. Furthermore, a utility decrement of 0.025 (based on two previous STAs on lung cancer treatments [36, 37]) was applied for patients receiving subcutaneous or intravenous treatments in the base-case.

The drug acquisition costs of BTD, PANO + BOR + DEX and CC treatments were based on list prices. For CC, cyclophosphamide in combination with THAL + DEX was assumed to reflect the UK practice. A PAS price discount agreement was available for POM + LoDEX. In addition, for pomalidomide, cost savings due to less drug use from dose interruptions longer than 28 days were incorporated. A questionnaire completed by six clinical specialists was used to collect data on monitoring, concomitant medication and AE costs. End-of-life cost estimates from a UK study were used [38]. Costs associated with intravenous and subcutaneous administration visits were based on the TA311 appraisal for BOR in first-line MM therapy [19]. In the base-case, no subsequent therapies were included following treatment discontinuation, but the impact of this assumption was explored in a scenario analysis based on

Table 7 Data sources and methodological approaches for covariate/crossover adjustments for each comparison

	POM + LoDEX vs. BTD	POM + LoDEX vs. PANO + BOR + DEX ^a	POM + LoDEX vs. CC
Data source used in covariate adjustment	IPD from MM-002, Gooding et al., Tarant et al. and MUK-One	IPD from a subset of MM-002, MM-003 and MM-010 Summary data and pseudo patient-level data from PANORAMA-2	MM-003 (HiDEX used as a proxy)
Covariates included for the adjustment	Age Prior lines of therapy Refractoriness to LEN Receipt of prior THAL	Age Prior lines of therapy Receipt of prior THAL ECOG status	Not required—within trial comparison
Adjustment method	CGP in the base-case and mean covariates methods for the scenario analysis, both methods described in Ghali et al. [35]	Matching-adjusted indirect comparison (MAIC)	Two-stage OS crossover adjustment for patients in the HiDEX arm who received POM + LoDEX after progression

POM pomalidomide, *LoDEX* low-dose dexamethasone, *PANO* panobinostat, *BOR* bortezomib, *DEX* dexamethasone, *OS* overall survival, *PFS* progression-free survival, *BTB* bendamustine + thalidomide + dexamethasone, *IPD* individual patient-level data, *THAL* thalidomide, *LEN* lenalidomide, *CGP* corrected group prognostics, *CC* conventional chemotherapy, *HR* hazard ratio, *ECOG* Eastern Cooperative Oncology Group, *MAIC* matching-adjusted indirect comparison, *HiDex* high-dose dexamethasone

^aThe PANO + BOR + DEX vs. POM + LoDEX PFS/OS HRs found from the MAIC method was applied on the POM + LoDEX curve that was derived from the CGP method conducted on MM-002, MM-003, MM-010 and all BTB trial data

Table 8 Survival analysis methods and curves chosen in the base-case for each comparison

	POM + LoDEX vs. BTB	POM + LoDEX vs. PANO + BOR + DEX	POM + LoDEX vs. CC
OS	Exponential distribution-based curves with baseline characteristics as covariates used for the joint modelling (i.e. for both POM + LoDEX and BTB) of OS data from the MM-002, MUK-ONE, Gooding et al. and Tarant et al. trials	For POM + LoDEX: generalised gamma distribution-based curve with baseline characteristics as covariates used for the extrapolation of OS data from the MM-002, MM-003, MM-010, MUK-ONE, Gooding et al. and Tarant et al. trials. Then, for PANO + BOR + DEX, the OS HR derived from MAIC was applied on that POM + LoDEX OS curve	Exponential distribution-based curves adjusted using two-stage Weibull approach for treatment switching were used for the joint modelling (i.e. for both POM + LoDEX and CC) of OS data from MM-003 trial
PFS	Generalized gamma distribution-based curves with baseline characteristics as covariates used for the joint modelling (i.e. for both POM + LoDEX and BTB) of PFS data from the MM-002, MUK-ONE, Gooding et al. and Tarant et al. trials	For POM + LoDEX: generalised gamma distribution-based curve with baseline characteristics as covariates used for the extrapolation of PFS data from the MM-002, MM-003, MM-010, MUK-ONE, Gooding et al. and Tarant et al. trials. Then, for PANO + BOR + DEX, the PFS HR derived from MAIC was applied on that POM + LoDEX PFS curve	Generalised gamma distribution-based curves were used for the joint modelling (i.e. for both POM + LoDEX and CC) of PFS data from the MM-003 trial
TTD	A common treatment effect based on the MM-002 and BTB trials applied to PFS generalized gamma curves	A common treatment effect based on the MM-002, MM-003, MM-010 and BTB trials was applied to the PFS generalized gamma curve for POM + LoDEX. No modelling for TTD of PANO + BOR + DEX treatment as they receive a fixed dose regimen	Extreme value distribution-based curves were used for the individual modelling (i.e. for POM + LoDEX and CC separately) of the TTD data from the MM-003 trial

POM pomalidomide, *LoDEX* low-dose dexamethasone, *PANO* panobinostat, *BOR* bortezomib, *DEX* dexamethasone, *OS* overall survival, *PFS* progression-free survival, *BTB* bendamustine + thalidomide + dexamethasone, *CC* conventional chemotherapy, *TTD* time to treatment discontinuation, *HR* hazard ratio, *MAIC* matching-adjusted indirect comparison

Haematological Malignancy Research Network registry data [4, 39]. All costs used in the model calculations were based on their 2015 (original or inflation-corrected) values.

The company did not provide a full incremental analysis including all comparators. Instead ICERs of POM + LoDEX versus each comparator were presented: £39,665 per quality-adjusted life-year (QALY) gained (vs. BTM), £141,793 savings per QALY lost—southwest quadrant (vs. PANO + BOR + DEX), and £44,811 per QALY gained (vs. CC).

Probabilistic analysis that included parametric uncertainty as well as structural uncertainty around the fitted curve distribution choice, generated similar mean ICERs and indicated the following probabilities of cost effectiveness of POM+LoDEX for each comparison: 92.8% versus BTM, 100% versus PANO+BOR+DEX (at list price for PANO), and 56.9% versus CC at a willingness-to-pay threshold of £50,000 per QALY gained.

The one-way sensitivity analyses showed that the parameters with the greatest impact on the outcomes were the coefficients used within the utility regression analysis for the comparison against BTM and CC, and the MAIC HRs (OS and PFS) for the comparison with PANO + BOR + DEX. The model is relatively insensitive to the majority of other parameters.

The scenario analyses demonstrated that for the comparison with BTM, time horizon, choice of dataset (e.g. MM-002, MM-003 and MM-010) and assumptions surrounding subsequent therapy, administration and AE costs had considerable impact on the ICER. For the PANO + BOR + DEX comparison, time horizon, OS/PFS curve distribution choices and the covariate adjustment method were the most influential on the ICER. Finally, for the comparison with CC, choice of the OS/PFS curve distribution and the method for the crossover adjustments had the biggest impact on the ICER.

3.4 Critique of the Cost-Effectiveness Evidence

The searches and economic model were considered by the ERG to meet the NICE reference case; however, in the final scope, NICE requested that at third line only PANO + BOR + DEX, and at later lines all comparators (BTM, PANO + BOR + DEX and CC), should have been considered. This line-based comparator stratification could not be achieved due to the lack of data.

In the CS base-case, a different dataset was used for each of the three comparisons, which implied that each comparison was based on a slightly different population. The ERG considered that an assessment based on a fully incremental analysis (e.g. an analysis that involves the calculation of incremental QALY gains and costs along all treatment options ranked by ascending cost), using the

same dataset (MM-002, MM-003, MM-010 and all BTM trials) for the POM + LoDEX PFS/OS estimates and applying the MAIC-based PFS/OS HRs for the PANO + BOR + DEX comparison and PFS HR and crossover-adjusted OS HR from the MM-003 trial for the CC comparison, might have been more consistent.

In general, the process for extrapolating survival curves was in line with the NICE DSU 14 guideline [40]. However, it was not clear how the visual fit has been assessed (i.e. it seemed as if the covariates of the parametric survival curves were adjusted according to the baseline of the pooled trials, whereas treatment-specific baselines might have been more appropriate for visual fit). Furthermore, the ERG noted that different baselines were assumed in the corrected group prognosis (CGP) method (baseline from the pooled trial datasets) and in the mean covariates method (baseline representing UK clinical practice).

Similarly, the ERG had some concerns related to the implementation of AEs. The followed approach would mirror the frequency order of the AEs of POM + LoDEX (MM-003 trial) for each of the comparators, in the same magnitude. The ERG considered this assumption not to be plausible because each drug has a different working mechanism and different safety profiles, and it is unlikely that the AE frequency order would be mirrored for other comparators in the same magnitude.

The ERG did not encounter major issues with the approach used to include quality of life in the model. Nevertheless, the data and the assumptions about the covariates that are included in the utility regression model (e.g. best overall response rate, hospitalisation and AEs) might lead to bias in utility estimates (e.g. assuming the same hospitalisation rate from the HiDEX arm of the MM-003 trial for all comparators, inconsistencies within the categorisation of best overall response rate among different comparators, and the uncertainty around the disutility associated with intravenous/subcutaneous treatments).

The current submission re-estimated some of the resource use inputs compared with TA338 (e.g. monitoring costs were based on an extensive questionnaire completed by six clinical experts). The ERG considered that the input parameters derived from the resource use questionnaire should be considered with care due to the low number of experts ($n = 6$) who completed the questionnaire, as well as the length and level of detail of the questionnaire, which might have provided a challenge to fill in all fields with the required attention. Furthermore, the model allowed for a decrease in drug acquisition costs based on treatment interruptions lasting longer than 28 days only for POM and panobinostat, whereas the dose interruptions of BOR (within PANO + BOR + DEX), BTM and CC were not taken into account at all, creating a potential inconsistency. Finally, regarding the costs of subsequent treatment, the

ERG does not agree with the base-case choice of excluding these costs. Since the effects of the subsequent treatments were implicitly incorporated into the OS results, it would be rational to also include the costs required to achieve those effects. The CS provided two cost estimates for the subsequent treatments to be used in scenario analyses. These estimates differed greatly, and the information on them was insufficient for the ERG to judge which estimate was better to use.

3.5 Additional Exploratory Analyses Conducted by the ERG

In the original CS model, the ERG identified some programming errors. After these errors were corrected, the base-case analyses of the company were repeated with the ERG-corrected model. The ICER results were approximately £45,000 per QALY gained for POM + LoDEX versus BTD, £143,000 savings per QALY lost for POM + LoDEX versus PANO + BOR + DEX—southwest quadrant—and £49,000 per QALY gained for POM + LoDEX versus CC.

The ERG then conducted a list of exploratory full-incremental analyses in which the effectiveness of BTD and POM + LoDEX treatments were based on the pooled dataset of MM-002, MM-003, MM-010 and all other BTD trials. In the first analysis, the CGP method was used for the covariate adjustment and POM + LoDEX and BTD OS and PFS curves were generated. For PANO + BOR + DEX, the HRs obtained from the MAIC were applied on top of the POM + LoDEX curves; and, similarly for CC, the HRs from the intention-to-treat analysis of the OS and PFS data from the MM-003 trial were applied on top of the POM + LoDEX curves. The second full-incremental analysis had the same assumptions as the first full incremental analysis, except the OS HR from the MM-003 trial that was applied on top of the POM + LoDEX curve was corrected for the treatment switching using a two-stage method. Finally, in the third full incremental analysis, for generating POM + LoDEX and BTD curves, the mean covariate adjustment method was used (not using the baseline of the pooled trials but a baseline reflecting real-world data from UK centres) instead of the CGP method. The results of these full incremental analyses (with list prices for panobinostat) are provided in Table 9.

In all of these exploratory analyses, CC, POM + LoDEX and PANO + BOR + DEX were on the cost-effectiveness frontier. POM + LoDEX seems to be cost effective in comparison with PANO + BOR + DEX, however the cost effectiveness of POM + LoDEX versus CC was dependent on underlying assumptions of the analyses.

Furthermore, in several scenario analyses, the ERG explored the uncertainty surrounding the assumptions on

dose interruptions, subsequent treatment costs, wastage drug costs, utility regression model inputs, AE rates and other utility estimates. The ICER results were more or less similar across these analyses. Finally, the ERG implemented the confidential PAS price for PANO + BOR + DEX, which decreased the ICER of PANO + BOR + DEX versus POM + LoDEX comparison (still in the southwest quadrant), however it still remained far above acceptable thresholds.

3.6 Conclusions of the ERG Report

Based on the MM-003 trial, POM appeared to extend OS and PFS in comparison with HiDEX in a heavily treated population who were refractory to BOR and LEN. The AE profile appears to be manageable with appropriate dose reductions and interruptions. The ERG drew attention to these AEs occurring more frequently in the POM arm, notably neutropenia.

For the comparison with BEN, the company used covariate adjustment methods to adjust for the differences between studies in patient characteristics. The covariate-adjusted results were very similar to the unadjusted results in terms of both PFS and OS, indicating that the differences between studies in the selected covariates (patient characteristics) have relatively little impact on the outcomes observed. However, the selection of different datasets for POM + LoDEX seemed to alter the results for OS and PFS.

For the comparison with PANO + BOR + DEX, the POM + LoDEX arms of a subpopulation from the MM-002, MM-003 and MM-010 studies were compared with the PANO + BOR + DEX arm of the PANORAMA-2 study using the MAIC method, which resulted in a 1-month increase in median OS for patients receiving POM + LoDEX compared with the unweighted analysis. In both cases, the median OS was shorter than those patients receiving PANO + BOR + DEX. The hazard of death seemed to be reduced by a similar amount for patients receiving PANO + BOR + DEX compared with POM + LoDEX (in both MAIC weighted and unweighted analysis, HR was approximately 0.75).

The economic model described in the CS was considered by the ERG to meet the NICE reference case to a reasonable extent and was mostly in line with the decision problem specified in the scope. The submitted model included some errors and the ERG corrected the model and conducted some exploratory analyses based on the pooled MM-002, MM-003, MM-010 and all BTD trials dataset. The full incremental exploratory analysis revealed that CC, POM + LoDEX and PANO + BOR + DEX were on the cost-effectiveness frontier and BTD was either dominated by CC or extendedly dominated by POM + LoDEX.

Table 9 Results from the full incremental exploratory analyses

	First full incremental analysis (CGP method, ITT OS HR from MM-003)	Second full incremental analysis (CGP method, two-stage crossover-adjusted OS HR from MM-003)	Third full incremental analysis (mean covariate method, two-stage crossover-adjusted OS HR from MM-003)
PANO + BOR + DEX	£142,930 per QALY gained	£142,930 per QALY gained	£146,307 per QALY gained
POM + LoDEX	£81,209 per QALY gained	£57,288 per QALY gained	£84,091 per QALY gained
BTD	Dominated by CC	Extendedly dominated by POM + LoDEX	Dominated by CC
CC	–	–	–

POM pomalidomide, *LoDEX* low-dose dexamethasone, *PANO* panobinostat, *BOR* bortezomib, *DEX* dexamethasone, *OS* overall survival, *BTD* bendamustine + thalidomide + dexamethasone, *CGP* corrected group prognostics, *CC* conventional chemotherapy, *HR* hazard ratio, *ITT* intention to treat, *QALY* quality-adjusted life-year

When POM + LoDEX was compared with CC and BTD, it can be considered as satisfying the end-of-life criteria, but when compared with PANO + BOR + DEX, the end-of-life criteria did not hold. The pairwise comparisons on the corrected company base-case revealed that the ICERs of POM + LoDEX versus CC and BTD were both below the £50,000 threshold. The ICER of POM + LoDEX versus PANO + BOR + DEX showed that POM + LoDEX was less effective but less costly, and the savings per QALY lost for POM + LoDEX suggested it was cost effective, without applying the end-of-life criteria threshold. Various scenario analyses revealed that the ICER is relatively robust against changes in input values and assumptions.

3.6.1 Key Methodological Issues

Due to the lack of randomized evidence, the company used different data sources and methodological approaches to assess the comparative effectiveness of POM + LoDEX versus BTD and PANO + BOR + DEX. Even though the chosen methods were deemed to be appropriate, the ERG noted that there was no formal method selection process. For instance, no justification was given by the company as to why MAIC was chosen instead of other alternatives such as simulated treatment comparison. Furthermore, the ERG was concerned that there was no assessment for possible unobserved confounding, which might have added additional uncertainty on the clinical and cost-effectiveness estimates. Finally, the ERG identified some inconsistencies in the implementation of these methods; for instance, CGP and the mean covariate methods assumed different baseline characteristics, or the indirect use of unrelated data from the BTD trials in the POM + LoDEX versus PANO + BOR + DEX comparison, etc. Therefore, the ERG has the opinion that a more systematic approach could have been followed in choosing and applying these methods for the

analysis of nonrandomized observational data (e.g. in line with the recommendations in NICE DSU TSD 17 and TSD 18 [34]).

The lack of a full incremental analysis including all comparators in the CS was another concern. Instead of a full incremental analysis, the company provided individual comparisons and each of the comparisons was conducted on a different dataset, which implied a slightly different population for each of the comparisons. The ERG was aware that a full incremental analysis would break the randomization link between POM + LoDEX and CC, however considered that the additional insights gained from full incremental analyses should not be dismissed.

4 National Institute for Health and Care Excellence Guidance

The committee concluded that the appropriate positioning of POM + LoDEX, in line with clinical practice and the evidence base, was after third or subsequent relapses, and the relevant comparators were PANO + BOR + DEX, BTD and CC. The committee acknowledged that the indirect comparisons were associated with considerable uncertainty, but recognised that the company had presented all the appropriate evidence available for its decision making.

The committee agreed that the model structure was appropriate. The ICERs varied based on the clinical datasets included, and using crossover adjustment and covariate adjustment methods. The committee understood the full incremental analyses provided by the ERG, but concluded that it would base its decisions on the company's base-case ICERs, corrected by the ERG for errors, for the sake of not breaking the randomisation between the POM + LoDEX and CC comparison.

The most plausible ICERs for POM + LoDEX versus CC and BTD were below £50,000 per QALY gained, and

the committee concluded that POM meets the end-of-life criteria compared with these two comparators. The end-of-life criterion for an additional 3 months survival gain was not met for the comparison with PANO + BOR + DEX, and, after incorporating all PAS prices, the ICERs reflected ‘savings per QALY lost’, i.e. POM + LoDEX was less effective but less costly. The committee concluded that the savings per QALY lost for POM + LoDEX compared with PANO + BOR + DEX were high enough for it to be considered a cost-effective use of NHS resources without applying the end-of-life criteria.

4.1 Final Guidance

POM + LoDEX was recommended as an option for treating multiple myeloma in adults at third or subsequent relapse, i.e. after three previous treatments, including both LEN and BOR, only when the company provided POM with the discount agreed in the PAS.

5 Conclusions

In the absence of randomized evidence, alternative data sources (including non-randomized evidence) can be used to assess comparative effectiveness and to inform cost-effectiveness analyses. However, adjustment methods are needed to account for the differences between patient populations at baseline. This STA showed that the selection of the method and its implementation can have quite an impact on the ICER; therefore, it is important to base all of these decisions on a systematic approach.

This STA also showed that a full incremental analysis based on a common dataset for all comparisons might provide additional insights (e.g. the treatments that were on the cost-effectiveness frontier) compared with the individual comparisons based on different datasets, which led to different cost and QALY outcomes for the same treatment (POM + LoDEX) in each comparison. However, the full incremental analysis in this STA required the breakdown of the randomisation between POM + LoDEX and CC, and the decision makers might prefer their decision to be based on direct head-to-head evidence rather than indirect evidence.

Another interesting outcome of this STA was the fact that an intervention might satisfy the end-of-life criteria for some of the comparators, whereas for others, it might not since it leads to a lower total QALY outcome. In such a situation, the decision maker can still recommend the treatment if the ‘costs per QALY gained’ were below the end-of-life threshold for the comparisons with less effective comparators, and if the ‘costs saved per QALY lost’

were higher than the standard threshold for the comparisons with more effective comparators.

Finally, this STA showed that a negative treatment recommendation from a previous appraisal can be changed in the presence of additional evidence, changes in the clinical landscape, and in the costs and drug prices.

Author contributions NB acted as a health economist on this assessment, critiqued the manufacturer’s economic evaluation, wrote the manuscript and acted as overall guarantor for the overall content of this article. SdG, FAdA and NA acted as health economists on this assessment, critiqued the manufacturer’s economic evaluation and contributed to the writing of the article. RW acted as project lead and systematic reviewer on this assessment and critiqued the clinical effectiveness methods and evidence. MA acted as health economic project lead and critiqued the manufacturer’s economic evaluation. DF acted as systematic reviewer and critiqued the clinical effectiveness methods and evidence. GW acted as statistician and critiqued the analyses in the manufacturer’s submission. LS critiqued the search methods in the submission. JK critiqued the manufacturer’s definition of the decision problem and their description of the underlying health problem and current service provision. All authors contributed to the conception and planning of the work, and critically revised and approved the final submitted version of the manuscript. This summary has not been externally reviewed by *PharmacoEconomics*.

Compliance with Ethical Standards

Funding This project was funded by the National Institute for Health Research (NIHR) Health Technology Assessment (HTA) Program (project number 13/104/01 STA) (see the HTA program website for further information—<http://www.hta.ac.uk>). This summary of the ERG report was compiled after the AC’s review. The views and opinions expressed are the authors and do not necessarily reflect those of the HTA Programme, NICE, NIHR, NHS, or the Department of Health.

Conflict of interest Nasuh C. Büyükkaramikli, Saskia de Groot, Debra Fayter, Robert Wolff, Nigel Armstrong, Lisa Stirk, Gill Worthy, Fernando Albuquerque de Almeida, Jos Kleijnen and Maiwenn J. Al have no conflicts of interest to declare that are relevant to the content of this review. We acknowledge Richard Birnie for his critique of the clinical-effectiveness evidence.

Open Access This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

1. National Institute for Health and Care Excellence. Guide to the processes of technology appraisal. London: NICE, 2014. <http://www.nice.org.uk/article/pmg19>. Accessed 2 Sep 14.
2. National Institute for Health and Care Excellence. Single technology appraisal: company evidence submission template. London: NICE, 2015. <http://www.nice.org.uk/Media/Default/About/what-we-do/NICE-guidance/NICE-technology-appraisals/specification->

- [for-company-submission-of-evidence-2015-version.docx](#). Accessed 11 Jun 15.
3. Hoffbrand A, Catovsky D, Tuddenham E, editors. Postgraduate haematology. 5th ed. Oxford: Blackwell; 2005.
 4. Haematological Malignancy Research Network (HMRN). Analysis of multiple myeloma treatment pathway provided to Celgene for sole purpose of NICE submission. Data rights belong to HMRN network and is shared as academic in confidence (**in print**).
 5. Howlader N, Noone A, Krapcho M, et al. SEER cancer statistics review, 1975–2013. 2016. http://seer.cancer.gov/csr/1975_2013/. Accessed 14 Jun 16.
 6. Cancer Research UK. Statistics and outlook for myeloma 2014. <http://www.cancerresearchuk.org/cancer-help/type/myeloma/treatment/statistics-and-outlook-for-myeloma#stage>. Accessed 11 Dec 13.
 7. Rajkumar SV, Harousseau JL, Durie B, Anderson KC, Dimopoulos M, Kyle R, et al. Consensus recommendations for the uniform reporting of clinical trials: report of the International Myeloma Workshop Consensus Panel 1. *Blood*. 2011;117(18):4691–5.
 8. Kyle RA, Gertz MA, Witzig TE, Lust JA, Lacy MQ, Dispenzieri A, et al. Review of 1027 patients with newly diagnosed multiple myeloma. *Mayo Clin Proc*. 2003;78(1):21–33.
 9. International Myeloma Foundation. Multiple myeloma: concise review of the disease and treatment options 2016. <http://myeloma.org/pdfs/ConciseReview.pdf>. Accessed 21 Jun 16.
 10. Borrello I. Can we change the disease biology of multiple myeloma? *Leuk Res*. 2012;36(Suppl 1):S3–12.
 11. British Committee for Standards in Haematology (BCSH). Guidelines for the diagnosis and management of multiple myeloma 2014. http://www.bcsghguidelines.com/4_HAEMATOLOGY_GUIDELINES.html?dtype=All&dstatus=All&dsdorder=&dstorder=&dmax=10&dsearch=multiple+myeloma+#gl. Accessed 11 Nov 13.
 12. Kumar SK, Therneau TM, Gertz MA, Lacy MQ, Dispenzieri A, Rajkumar SV, et al. Clinical course of patients with relapsed multiple myeloma. *Mayo Clin Proc*. 2004;79(7):867–74.
 13. Rawstron AC, Gregory WM, de Tute RM, Davies FE, Bell SE, Drayson MT, et al. Minimal residual disease in myeloma by flow cytometry: independent prediction of survival benefit per log reduction. *Blood*. 2015;125(12):1932–5.
 14. Streetly M. Clinical review of overall survival for myeloma patients progressing after both bortezomib and lenalidomide based therapy. 54th British Society for Haematology annual scientific meeting, April 2014; Birmingham; 2014. p 163.
 15. Gooding S, Lau IJ, Sheikh M, Roberts P, Wong J, Dickens E, et al. Double relapsed and/or refractory multiple myeloma: clinical outcomes and real world healthcare costs. *PLoS One*. 2015;10(9):e0136207.
 16. Tarant J, Ashcroft J, Feyler S, et al. Treatment patterns & survival in multiple myeloma patients sequentially exposed to thalidomide, bortezomib & lenalidomide in a UK single centre. *Blood*. 2013;122(21):5380.
 17. Lau IJ, Smith D, Aitchison R, Blesing N, Roberts P, Peniket A, et al. Bendamustine in combination with thalidomide and dexamethasone is a viable salvage option in myeloma relapsed and/or refractory to bortezomib and lenalidomide. *Ann Hematol*. 2015;94(4):643–9.
 18. Kumar SK, Lee JH, Lahuerta JJ, Morgan G, Richardson PG, Crowley J, et al. Risk of progression and survival in multiple myeloma relapsing after therapy with IMiDs and bortezomib: a multicenter International Myeloma Working Group study. *Leukemia*. 2012;26(1):149–57.
 19. National Institute for Health and Care Excellence. Bortezomib for induction therapy in multiple myeloma before high-dose chemotherapy and autologous stem cell transplantation. NICE technology appraisal guidance (TA311). London: NICE, 2014. <https://www.nice.org.uk/guidance/ta311>. Accessed 9 Jun 14.
 20. National Institute for Health and Clinical Excellence. Bortezomib and thalidomide for the first-line treatment of multiple myeloma. NICE technology appraisal guidance (TA228). London: NICE, 2011. <https://www.nice.org.uk/guidance/TA129>. Accessed 11 Dec 13.
 21. National Institute for Health and Clinical Excellence. Bortezomib monotherapy for relapsed multiple myeloma. NICE technology appraisal guidance (TA129). London: NICE; 2007. <https://www.nice.org.uk/guidance/TA129>. Accessed 11 Dec 13.
 22. National Institute for Health and Clinical Excellence. Lenalidomide for the treatment of multiple myeloma in people who have received at least one prior therapy. NICE technology appraisal guidance (TA171). London: NICE, 2009. <https://www.nice.org.uk/guidance/ta171>. Accessed 11 Dec 13.
 23. National Institute for Health and Care Excellence (NICE). Panobinostat for treating multiple myeloma after at least 2 previous treatments. NICE technology appraisal guidance (TA380) 2016 (June 2016). <https://www.nice.org.uk/guidance/ta380/>. Accessed 13 May 16.
 24. NHS England. National cancer drugs fund list Ver6.1. 2016. <https://www.england.nhs.uk/wp-content/uploads/2017/10/national-cdf-list-v1.47.pdf>. Accessed 13 May 16.
 25. Celgene Ltd. Summary of product characteristics (SPC). Imnovid® (pomalidomide). 6 June 2016. <http://www.medicines.org.uk/emc/medicine/28269/SPC/>. Accessed 14 Jun 16.
 26. National Institute for Health and Care Excellence. Pomalidomide with dexamethasone for treating relapsed and refractory multiple myeloma after at least two regimens including lenalidomide and bortezomib (review of TA338). Final scope. NICE, 2016. <https://www.nice.org.uk/guidance/ta427/documents/final-scope>. Accessed 18 Jul 16.
 27. Schey S, Brown SR, Tillotson AL, Yong K, Williams C, Davies F, et al. Bendamustine, thalidomide and dexamethasone combination therapy for relapsed/refractory myeloma patients: results of the MUKone randomized dose selection trial. *Br J Haematol*. 2015;170(3):336–48.
 28. San Miguel J, Weisel K, Moreau P, Lacy M, Song K, Delforge M, et al. Pomalidomide plus low-dose dexamethasone versus high-dose dexamethasone alone for patients with relapsed and refractory multiple myeloma (MM-003): a randomised, open-label, phase 3 trial. *Lancet Oncol*. 2013;14(11):1055–66.
 29. Richardson PG, Siegel DS, Vij R, Hofmeister CC, Baz R, Jagannath S, et al. Pomalidomide alone or in combination with low-dose dexamethasone in relapsed and refractory multiple myeloma: a randomized phase 2 study. *Blood*. 2014;123(12):1826–32.
 30. Dimopoulos MA, Sonneveld P, Leung N, Merlini G, Ludwig H, Kastiris E, et al. International Myeloma Working Group recommendations for the diagnosis and management of myeloma-related renal impairment. *J Clin Oncol*. 2016;34(13):1544–57.
 31. Richardson PG, Schlossman RL, Alsina M, Weber DM, Coutre SE, Gasparetto C, et al. PANORAMA 2: panobinostat in combination with bortezomib and dexamethasone in patients with relapsed and bortezomib-refractory myeloma. *Blood*. 2013;122(14):2331–7.
 32. Celgene Ltd. Pomalidomide for relapsed and refractory multiple myeloma previously treated with lenalidomide and bortezomib (review of TA338) (ID985): submission to National Institute of Health and Clinical Excellence. Single technology appraisal (STA). 2016. <https://www.nice.org.uk/guidance/ta427/documents/committee-papers>.
 33. Fayter F, Büyükkaramikli N, Birnie R, de Groot S, Wolff R, Armstrong N, et al. Pomalidomide with dexamethasone for

- treating relapsed and refractory multiple myeloma after at least two regimens including lenalidomide and bortezomib (review of TA338). 2016. <https://www.nice.org.uk/guidance/ta427/documents/committee-papers>.
34. Faria R, Hernandex Alava M, Manca A, Wailoo AJ. NICE DSU technical support document 17: the use of observational data to inform estimates of treatment effectiveness in technology appraisal: methods for comparative individual patient data. Sheffield: Decision Support Unit, ScHARR, University of Sheffield, 2015. <http://scharr.dept.shef.ac.uk/nicedsu/wp-content/uploads/sites/7/2016/03/TSD17-DSU-Observational-data-FINAL.pdf>. Accessed 6 Sep 16.
 35. Ghali WA, Quan H, Brant R, van Melle G, Norris CM, Faris PD, et al. Comparison of 2 methods for calculating adjusted survival curves from proportional hazards models. *JAMA*. 2001;286(12):1494–7.
 36. Roche. Tarceva® (erlotinib) NICE STA Submission: Achieving clinical excellence in the treatment of relapsed non-small cell lung cancer. 2006. <http://www.nice.org.uk/nicemedia/live/11714/37396/37396.pdf>. Accessed 3 Mar 13.
 37. AstraZeneca UK Ltd. Single technology appraisal (STA) for gefitinib for the first line treatment of locally advanced or metastatic non-small lung cancer. 2009. <https://www.nice.org.uk/guidance/ta192>. Accessed 3 Mar 14.
 38. Addicott R, Dewar S. Improving choice at end of life: a descriptive analysis of the impact and costs of the Marie Curie Delivering Choice Programme in Lincolnshire. 2008. <http://www.kingsfund.org.uk/sites/files/kf/improving-choice-end-of-life-descriptive-analysis-impact-costs-marie-curie-choice-programme-lincolnshire-rachael-addicot-steve-dewar-april-2008.pdf>. Accessed 2 Jul 16.
 39. Celgene Ltd. Multiple myeloma advisory board meeting, London. 23 Mar 2016 (**data on file**).
 40. Latimer N. NICE DSU technical support document 14: survival analysis for economic evaluations alongside clinical trials - extrapolation with patient-level data 2013. Sheffield: School of Health and Related Research, University of Sheffield, 2013. <http://www.nicedsu.org.uk/NICE%20DSU%20TSD%20Survival%20analysis.updated%20March%202013.v2.pdf>. Accessed 23 Oct 13.