Carfilzomib for previously treated multiple myeloma (part review of TA457) (ID1493)

Single Technology Assessment Report

Source of funding

This report was commissioned by the National Institute for Health Research Systematic Reviews Programme as project number 129575/T.

Title:	Carfilzomib for previously treated multiple myeloma (part review of TA457) (ID1493)
Produced by:	BMJ Technology Assessment Group (BMJ TAG)
Authors:	Steven J Edwards, Director of Health Technology Assessment, BMJ-TAG, London Samantha Barton, Principal Health Technology Assessment Analyst, BMJ-TAG, London Tracey Jhita, Health Economics Manager, BMJ-TAG, London Charlotta Karner, Health Technology Assessment Analysis Manager, BMJ-TAG, London Peter Cain, Senior Health Economist, BMJ-TAG, London
Correspondence to:	Steve Edwards, BMJ TAG, BMJ, BMA House, Tavistock Square, London, WC1H 9JR.
Date completed:	17/03/2020
Source of funding:	This report was commissioned by the National Institute for Health Research Systematic Reviews Programme as project number 129575/T.
Declared competing interests of the authors	No competing interests were declared which affect the impartiality of this report. BMJ Technology Assessment Group (BMJ-TAG) and the editorial team of The BMJ work independently to one another. The views and opinions expressed in this report are those of the BMJ-TAG.
Acknowledgments:	The ERG would like to thank Drs Kevin Boyd (Consultant Haematologist, Royal Marsden) and Majid Kazmi (Consultant Haematologist, Guy's and St Thomas' Hospital) for providing clinical advice throughout the project, and for providing feedback on the clinical sections of the report.
Rider on responsibility for report:	The views expressed in this report are those of the authors and not necessarily those of the NIHR Systematic Reviews Programme. Any errors are the responsibility of the authors.
Report reference:	Edwards SJ, Barton S, Jhita T, Karner C, Cain P. Carfilzomib for previously treated multiple myeloma (part review of TA457) (ID1493). A Single Technology Appraisal. BMJ Technology Assessment Group, 2019.

Contribution of authors:

Steve Edwards	Critical appraisal of the company's submission; validated the statistical analyses; provided feedback on all versions of the report. Guarantor of the report
Samantha Barton	Critical appraisal of the company's submission; critical appraisal of the clinical evidence; cross checking of company's search strategies; and drafted the summary, and clinical results sections
Charlotta Karner	Critical appraisal of the company's submission; critical appraisal of the clinical evidence; cross checking of company's search strategies; and assisted with drafting the background and clinical results sections
Tracey Jhita	Critical appraisal of the company's submission; critical appraisal of the economic model; cross checking of company's search strategies; critical appraisal of the economic evidence; carried out the economic analyses; and drafted the economic sections
Peter Cain	Critical appraisal of the company's submission; critical appraisal of the economic model; cross checking of company's search strategies; critical appraisal of the economic evidence; carried out the economic analyses; and drafted the economic sections

All authors read and commented on draft versions of the ERG report.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

2L	Second line
AE	Adverse event
AIC	Akaike Information Criterion
BSC	Best supportive care
C	Carfilzomib
CDF	Cancer Drugs Fund
CI	Confidence interval
CRd	Carfilzomib in combination with lenalidomide and dexamethasone
CS	Company's submission
d	Dexamethasone
DSU	Decision Support Unit
ECOG	Eastern Cooperative Oncology Group
EMA	European Medicines Agency
EORTC	European Organization for Research and Treatment of Cancer
EQ-5D	EuroQol-5-dimension Questionnaire
ERG	Evidence Review Group
HR	Hazard ratio
HR-QoL	Health-related quality of life
ICER	Incremental cost-effectiveness ratio
IMiD	Immunomodulatory drug
IRC	Independent Review Committee
ITT	Intention-to-treat
IV	Intravenous
KM	Kaplan–Meier
MID	Minimal important difference
MM	Multiple myeloma
NHS	National Health Service
NICE	National Institute of Health and Care Excellence
OR	Odds ratio
ORR	Overall response rate
OS	Overall survival
PFS	Progression-free survival
PSS	Personal Social Services
QALY	Quality adjusted life year
R	Lenalidomide
RCT	Randomised controlled trial
SCT	Stem cell transplantation

SD	Standard deviation
SLR	Systematic literature review
SmPC	Summary of Product Characteristics
STA	Single Technology Appraisal
TEAE	Treatment-emergent adverse effect
TTD	Time-to-treatment discontinuation
TTP	Time to progression

1 Executive summary

1.1 Critique of the decision problem in the company's submission

Evidence on the clinical and cost effectiveness for carfilzomib in a doublet (Cd) and a triplet (CRd) regimen in the management of multiple myeloma has previously been reviewed as part of the Technology Appraisal process (TA457), with the committee recommending:

- Cd as an option for treating multiple myeloma in adults, only if:
    - people have had only 1 previous therapy, which did not include bortezomib; and
    - the company provides carfilzomib with the discount agreed in the patient access scheme.

The triplet combination of CRd at third line (3L) was considered but not recommended. The clinical and cost effectiveness of CRd at second line (2L) was not discussed as part of TA457.

For the decision problem that is the focus of this STA, which is a part review of TA457, the company submitted evidence on the clinical effectiveness of CRd as a 2L treatment for those with multiple myeloma, and specifically those who have undergone prior treatment with a bortezomib-based regimen (2L prior bortezomib). Thus, the company's submission is narrower than the final scope issued by the National Institute for Health and Care Excellence (NICE), which specified the population to be adults with multiple myeloma who had received at least one prior therapy. As a consequence of the restriction of the population to those receiving CRd at 2L after a bortezomib-based regimen, the sole relevant comparator of interest available through routine commissioning becomes Rd.

Evidence in support of the clinical effectiveness of CRd in the management of multiple myeloma at 2L is derived from ASPIRE, a randomised controlled trial enrolling adults with multiple myeloma who had received one or more previous lines of therapy, which was reviewed in TA457. Revised estimates of comparative clinical effectiveness for CRd versus Rd at 2L based on more mature data are available for only PFS and OS (cut-off date of December 2017) compared with data presented in TA457. Analysis of response rates by the Independent Review Committee (IRC) and capture of health-related quality of life (HRQoL) outcomes ceased on demonstration of a benefit in PFS and, thus, results for those outcomes are based on data from the interim analysis (June 2014).

1.2 Summary of the key issues in the clinical effectiveness evidence

Considering the evidence informing estimates of effect for CRd versus Rd in the 2L setting, the Evidence Review Group's (ERG's) key reservations around the data are:

- estimates of effect are derived from *post hoc* subgroups from the ASPIRE trial, which was reviewed as part of TA457;
    - estimates derived from *post hoc* subgroups are at a higher risk of bias than those reported for the full trial population;
- the company provided data for a subgroup in which a proportion of people had not received bortezomib as part of their last regimen, and a proportion of people who had undergone treatment with lenalidomide (2L prior bortezomib), which does not reflect NICE approved first-line treatment for multiple myeloma. The ERG considers the subgroup in which all people had received one line of prior treatment that included bortezomib and no lenalidomide (2L prior bortezomib/no prior lenalidomide), to be more relevant to the decision problem and requested characteristics and results for this subgroup from the company at the clarification stage.
- as would be expected, imbalances were noted in some baseline characteristics between those given CRd and those administered Rd in the *post hoc* subgroups. The direction of bias introduced by the differences in baseline characteristics, and the impact on estimates of relative treatment effect, is unclear.
- updated estimate of PFS informing this STA is based on assessment by investigator as IRC ceased assessing results after demonstration of benefit in PFS at interim analysis;
    - ASPIRE is an open label trial and assessment of PFS is potentially at risk of bias.

To mitigate against the imbalances in baseline characteristics, and to address the limitations associated with use of data derived from a *post hoc* subgroup, the company carried out an inverse probability weighted (IPW) analysis to generate estimates for PFS and OS for CRd versus Rd in the *post hoc* subgroups. In TA457, results from subgroup analyses adjusted to account for imbalances in baseline characteristics arising from non-randomised groups were accepted by the committee. The ERG considers that the company's IPW analysis to adjust subgroup data for imbalances can be considered appropriate for decision-making. Additionally, the company highlights that for PFS, "*there is a consistent treatment effect across baseline covariate subgroups*". As hazard ratios (HRs) derived from an ITT population of an RCT are, by their nature, more robust than those generated from a subgroup analysis, the ERG considers that the results from the ITT population are relevant to

the STA. A summary of PFS and OS for the ITT population, and the 2L prior bortezomib/no prior lenalidomide and 2L prior bortezomib subgroups, is presented in Table A.

Table A. Summary of PFS and OS for CRd versus Rd for the ITT population of ASPIRE and the two post hoc subgroups evaluating CRd as a second-line treatment

Outcome	ITT[a]	2L prior bortezomib/no prior lenalidomide[b]	2L prior bortezomib[b]
PFS	0.659 (0.553 to 0.784)	[redacted]	[redacted]
OS	0.794 (0.667 to 0.945)	[redacted]	[redacted]

Results are presented as Hazard ratio with accompanying 95% confidence interval.
[a] Unadjusted analysis.
[b] Results of inverse probability weighted analysis, adjusted for covariates selected using stepwise logistic regression.
Abbreviations: CRd, carfilzomib, lenalidomide and dexamethasone; ERG, evidence review group; ITT, intention to treat; OS, overall survival; PFS, progression-free survival; Rd, lenalidomide and dexamethasone.

Although the ERG predominantly considers the company's approach to identification of relevant covariates for the IPW appropriate, the ERG considers it important to highlight that the regression analyses [redacted] for specific individual covariates. The ERG considers that the results could suggest that the identified covariates are potential treatment effect modifiers. In particular, adjustment for prior SCT and for β2-microglobulin level suggest that, compared with Rd, treatment with CRd is associated with a [redacted]. As data are derived from *post hoc* subgroup analyses, the ERG emphasises that any inferences from the results are hypothesis generating.

1.3 Summary of the key issues in the cost effectiveness evidence

The ERG considers the key issues with the cost-effectiveness analysis are as follows:

- As mentioned in Section 1.2, the company's subgroup of 2L prior bortezomib that is used for the base-case analysis includes a proportion of patients that received lenalidomide (Section 4.2.2). In England, bortezomib in combination with lenalidomide is not an approved regimen. In response to ERG clarification questions, the company provided scenario analysis for the 2L prior bortezomib/no prior lenalidomide subgroup, which the ERG deems more appropriate for the analysis and is used for the ERG base-case analysis.

- The company's approach to estimate OS for the Rd arm is based on a hybrid of extrapolated ASPIRE IPW OS data and real-world evidence from a French registry of multiple myeloma patients, MyelomaToul.
    - For the CRd arm, OS is also based on extrapolated ASPIRE IPW OS data and MyelomaToul data adjusted using the IPW OS hazard ratio (HR) from ASPIRE (Section 4.2.5). The company chose this approach as they deemed the survival estimates based solely on ASPIRE using the Weibull distribution, which they deemed the best-fitting distribution to the observed data, produced pessimistic results for the Rd arm.
    - The ERG consulted its clinical experts who confirmed that longer-term survival estimates for Rd patients based on ASPIRE are conservative. However, the ERG considers that the company's adjustment of Rd survival results in survival that is inflated for CRd compared with the extrapolated estimates based on IPW ASPIRE data.
    - As such, the ERG considers that the company could have chosen a more clinically plausible extrapolation of the ASPIRE data to use for the base-case. The company confirmed that if they used MyelomaToul to validate their extrapolations, the exponential distribution would have been appropriate to estimate OS. The ERG considers that the exponential distribution produced similar survival estimates for Rd compared with company's base-case estimates.
    - Furthermore, the CRd OS survival estimates are based entirely on mature ASPIRE OS data, which the ERG deems is appropriate and reduces the uncertainty in the analysis.
- As an illustrative scenario, the ERG tested the impact of utilising ITT hazard ratios (HRs) for PFS and OS for the reasons highlighted in Section 1.2.
- Pre-progression utility values in the model capture both mean increase in utility from baseline for both treatment arms as well as treatment-specific increase in utility if a patient is on CRd (Section 4.2.7.1). Change from baseline was the outcome of the utility model so the mean change from baseline is estimated from the individual effects of each covariate that is adjusted for. However mean change in utility over time was [redacted] for CRd than the Rd, even though all patients have progression-free disease. Furthermore, clinical expert advice sought by the ERG suggests that there is no clinical reason for there to be a treatment-specific utility benefit in addition to the benefit provided by any gains in

progression-free survival. Thus, the ERG considers that it is more appropriate for pre-progression utility values for both treatment arms to be equal and that difference in pre-progression quality-adjusted life-years (QALYs) should be determined by length of time spent in the progression-free health state.

- Other issues in the cost-effectiveness analysis that were investigated but found to have minimal impact on the ICER were alternative modelling of time-to-treatment discontinuation for CRd (Section 4.2.5.1), changes to assumptions for adverse events (Section 4.2.6.1), use of investigational drugs for subsequent treatment in ASPIRE (Section 4.2.8.8), alternative weighting of subsequent treatment costs and uncertainty around monitoring costs (Section 4.2.8.8).

1.4 Summary of the ERG's preferred assumptions and resulting ICER

The ERG's preferred assumptions for the cost-effectiveness analysis of CRd compared with Rd are as follows:

- 2L prior bortezomib/no prior lenalidomide subgroup – Section 4.2.5.1 & 6.2;
- Jointly fitted exponential distribution for OS – ASPIRE only – Section 4.2.5.1;
- Removal of treatment effect and average increase in utility for cycle 3 onwards for pre-progression health state utility value – Section 4.2.7.3.

Results of the ERG preferred base-case deterministic ICER compared with the company base-case deterministic ICER, including the confidential patient access scheme (PAS) of [redacted] for carfilzomib, are presented in Table B. The PSA ICER for the ERG preferred base-case is £55,530. A confidential appendix is supplied alongside this report with the confidential PAS's for the comparator lenalidomide and the subsequent therapies panobinostat, pomalidomide and bortezomib applied.

Table B. Deterministic cost-effectiveness results – company vs ERG base-case

Intervention	Total costs	Total QALYs	Δ costs	Δ QALYs	ICER £/QALY
Corrected company base-case					
Rd	[redacted]	2.58	-	-	-
CRd	[redacted]	3.96	60,467	1.38	**43,952**
ERG preferred base-case					
Rd	[redacted]	2.40			
CRd	[redacted]	3.44	53,017	1.04	**50,960**

Abbreviations: CRd, carfilzomib, lenalidomide and dexamethasone; ERG, evidence review group; ICER, incremental cost-effectiveness ratio; QALY, quality-adjusted life-year; Rd, lenalidomide and dexamethasone.

1.5 Summary of exploratory and sensitivity analyses undertaken by the ERG

Table C presents the ERG's exploratory analysis for the cost-effectiveness of CRd compared with Rd.

Table C. ERG exploratory analysis

Scenario	Section in ERG report	CRd		Rd		ICER £/QALY
		Costs (£)	QALYs	Costs (£)	QALYs	
Corrected company base-case	*6.1*	[redacted]	3.96	[redacted]	2.58	43,952
Corrected company scenario for the 2L prior bortezomib/no prior lenalidomide subgroup	*6.2*	[redacted]	3.94	[redacted]	2.58	40,335
Jointly fitted exponential distribution for OS – ASPIRE only	*4.2.5.1*	[redacted]	3.68	[redacted]	2.52	45,919
PFS and OS CRd curves using ITT PFS and OS HR applied to company scenario PFS and OS	*4.2.5.1*	[redacted]	3.26	[redacted]	2.58	76,716
PFS and OS CRd curves using ITT PFS HR applied to company scenario Rd PFS curve and ITT OS HR applied to ERG preferred Rd OS curve	*4.2.5.1*	[redacted]	3.16	[redacted]	2.52	81,593
Weibull distribution for CRd TTD	*4.2.5.1*	[redacted]	3.94	[redacted]	2.58	40,552
No treatment effect applied for pre-progression health state utility value	*4.2.7.3*	[redacted]	3.96	[redacted]	2.64	41,303
No average increase in baseline utility from cycle 3 onwards	*4.2.7.3*	[redacted]	3.68	[redacted]	2.43	43,583
Subsequent therapy based on ASPIRE and inclusion of investigational drugs cost for subsequent therapy	*4.2.8.8*	[redacted]	3.94	[redacted]	2.58	42,657
Assuming a 50% increase in costs for routine monitoring in the PFS health state	*4.2.8.8*	[redacted]	3.94	[redacted]	2.58	40,903
Alternative weighting of subsequent treatment costs	*4.2.8.8*	[redacted]	3.94	[redacted]	2.58	40,253

Abbreviations: 2L, second-line; CRd, carfilzomib plus lenalidomide and dexamethasone; HR, hazard ratio; ICER, incremental cost-effectiveness ratio; ITT, intention-to-treat; OS, overall survival; PFS, progression-free survival; QALY, quality-adjusted life-year; Rd, lenalidomide plus dexamethasone; TTD, time-to-treatment discontinuation

2 Introduction and background

2.1 Introduction

The company producing carfilzomib (Kyprolis®; Amgen) submitted to the National Institute for Health and Care Excellence (NICE) clinical and economic evidence in support of the effectiveness of carfilzomib (C) in combination with lenalidomide (R) and dexamethasone (d) for the treatment of adults with multiple myeloma. Specifically, the company presents evidence on comparative clinical effectiveness of CRd versus Rd for those who have received only one prior bortezomib-based therapy, which is narrower than the final scope issued by NICE.[1] Herein is a critique of the company's submission (CS) to the Single Technology Appraisal (STA), together with supplementary information, where necessary, provided by the company during the clarification process.

2.2 Background

Within Section B.1 of the CS, the company provides an overview of:

- carfilzomib, including its mode of action, dose and method of administration (Section B.1.2);
- multiple myeloma, including prevalence, prognosis and disease management (Section B.1.3).

The Evidence Review Group (ERG) considers the CS to present an accurate overview of carfilzomib.

The current treatment pathway for multiple myeloma is complex and rapidly changing, with multiple treatments approved at some lines of therapy and a lack of options at other lines. Given that the company is proposing restricting use of CRd to the second-line setting and after treatment with a bortezomib-based regimen, a decision with which the ERG's clinical experts agree (discussed in greater detail in Section 2.3.1), the ERG considers it would be beneficial to simplify the company's overview of the treatment pathway to focus on treatment options available at second line in UK clinical practice (Figure 1).

As the company highlights, various factors are considered when deciding on treatment, including comorbidities, age, general health status, and prior myeloma treatment. The preferred first-line treatment for patients younger than 65 years and who are physically fit is high-dose chemotherapy with stem cell transplant (SCT). Patients deemed eligible for SCT initially undergo induction therapy with a bortezomib-based regimen to reduce the number of myeloma cells in the bone marrow (Figure 1).[2] However, many patients are not suitable for SCT and will be treated with pharmacotherapy alone.[3] NICE recommends a thalidomide-based regimen as a first-line treatment

for patients who are ineligible for SCT.[4] Alternative options for those who are contraindicated or unable to tolerate thalidomide are Rd or a bortezomib-based therapy (Figure 1).[4, 5]

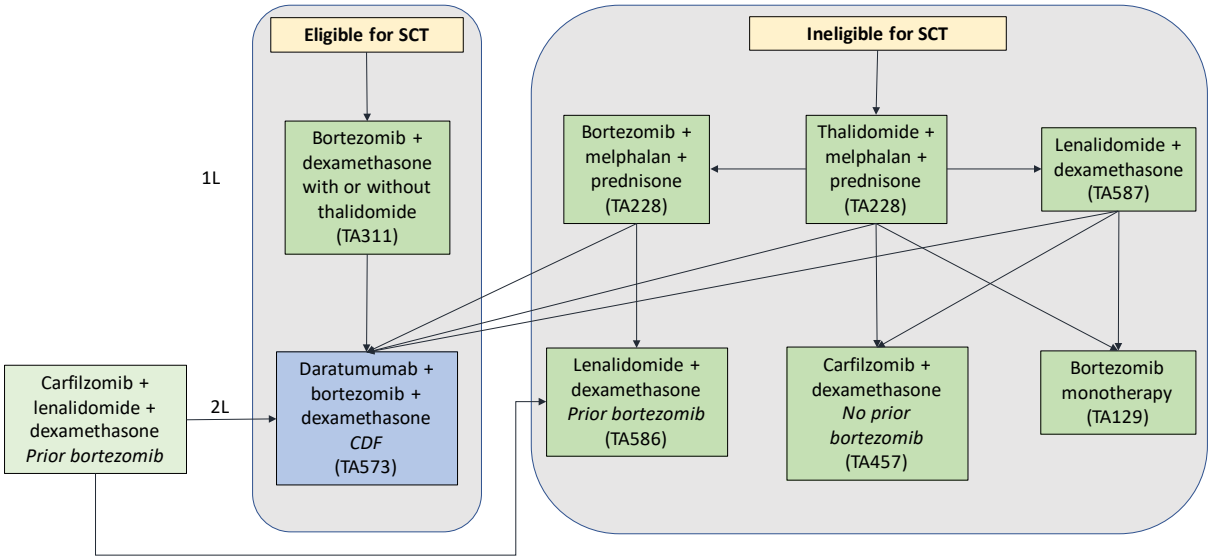
Based on the company's interpretation of the treatment pathway for multiple myeloma, the ERG considers that the company is positioning CRd as a second-line treatment for people who have had prior bortezomib-based treatment, irrespective of eligibility for SCT. However, based on NICE guidance, whether a person undergoes SCT influences the treatment options available at second line (Figure 1). At the time of writing, for those who receive SCT, no treatment option is available in the second-line setting as part of routine commissioning. Daratumumab (D) in combination with bortezomib (V) and dexamethasone (d; DVd) for second-line treatment after SCT became available in April 2019 through the Cancer Drugs Fund (CDF).[6] DVd was recommended with no stipulation on eligibility for SCT or type of prior therapy, and so is also an option for those deemed to be ineligible for SCT. Remaining treatment options at second line for those who have not undergone SCT at that time are:

- Rd;[7]
- Cd;[8]
- bortezomib monotherapy.

However, Rd is recommended after prior bortezomib-based treatment, and Cd is available to those who have not received prior bortezomib (Figure 1).

The ERG's clinical experts fed back that bortezomib monotherapy is rarely given as more effective treatment options are available at second line: relevant comparators for CRd are discussed in greater detail in Section 2.3.3.

Figure 1. Current pathway for first- and second-line treatment of multiple myeloma based on NICE guidance, and the proposed position of CRd



Abbreviations: 1L, first line; 2L, second line; C, carfilzomib; CDF, Cancer Drugs Fund; d, dexamethasone; NICE, National Institute for Health and Care Excellence; R, lenalidomide; SCT, stem cell transplant; TA, Technology Appraisal.

2.3 Critique of the company's definition of the decision problem

The company provided a summary of the final scope issued by the NICE, together with their rationale for any deviation from the final scope (Table 1).[1] The company highlights that the submission differs from the final scope primarily in terms of the population of interest to the decision problem (Table 1 and Table 2). The differences between the decision problem addressed in the company submission (CS) and the scope are discussed in greater detail in the sections that follow.

Table 1. Summary of decision problem (adapted from Table 1 in Document B, pages 9–12)

	Final scope issued by NICE	Decision problem addressed in the submission
Population	Adults with multiple myeloma who have had at least 1 previous therapy	Adults with multiple myeloma who have received only one prior therapy with bortezomib
Intervention	Carfilzomib plus lenalidomide and dexamethasone	Per final scope
Comparator(s)	For people who have had 1 previous therapy: • carfilzomib plus dexamethasone; • lenalidomide plus dexamethasone; • bortezomib. For people who have had 2 previous therapies: • lenalidomide plus dexamethasone; • panobinostat plus bortezomib and dexamethasone. For people who have had 3 or more previous therapies: • lenalidomide plus dexamethasone • panobinostat plus bortezomib and dexamethasone; • pomalidomide plus dexamethasone.	For people who have received one prior therapy with bortezomib: • lenalidomide plus dexamethasone. An additional analysis is also presented versus DVd which is currently recommended for use within the Cancer Drugs Fund as a treatment option for adults who have had one prior therapy.
Outcomes	The outcome measures to be considered include: • progression-free survival; • overall survival; • response rates (for example complete response); • time to next treatment; • adverse effects of treatment; • health-related quality of life.	Per final scope
Economic analysis	The reference case stipulates that the cost effectiveness of treatments should be expressed in terms of incremental cost per quality-adjusted life year. The reference case stipulates that the time horizon for estimating clinical and cost effectiveness should be sufficiently long to reflect any differences in costs or outcomes between the technologies being compared. Costs will be considered from an NHS and Personal Social Services perspective.	Per final scope

	The availability of any patient access schemes for the intervention or comparator technologies will be taken into account.	
Subgroups to be considered	If the evidence allows, subgroup analyses based on type and number of lines of previous therapy will be considered	Patients who have received one prior therapy with bortezomib
Special considerations, including issues related to equity or equality	None included	None included

Abbreviations: CDF, Cancer Drugs Fund; CRd, Carfilzomib in combination with lenalidomide and dexamethasone; CS, company submission; DVd, Daratumumab in combination with bortezomib and dexamethasone; ERG, Evidence Review Group; NHS, National Health Service; N/A, not applicable; NICE, National Institute for Health and Care Excellence; Rd, Lenalidomide and dexamethasone; TA, Technology Appraisal.

Table 2. Rationale for deviation from decision problem (adapted from Table 1 in Document B, pages 9–12)

	Company's rationale if different from the scope	ERG comment
Population	CRd is not positioned for use in patients who have received more than one prior therapy as it is anticipated to be used earlier in the treatment pathway in clinical practice	Based on feedback from the ERG's clinical experts, the ERG considers the company's rationale for focusing on those receiving CRd at second line and after prior bortezomib at first line to be appropriate (discussed in greater detail in Section 2.3.1). As part of the clarification process, the ERG requested that the company generate subgroups for CRd and Rd as per the population of interest to the STA (discussed in greater detail in Section 2.3.1).
Intervention	N/A	Schedule of CRd assessed in key RCT on clinical effectiveness of triplet combination (ASPIRE[9]) restricts use of carfilzomib to 18 cycles, whereas carfilzomib could be given for more cycles in UK clinical practice (dosing schedule reported in Section 2.3.2). Maximum of 18 cycles implemented in economic evaluation.
Comparator(s)	People who have received one prior therapy: • Amgen proposes that CRd will be used primarily as an alternative treatment option to Rd in patients who have received one prior therapy with bortezomib. This positioning is aligned with clinical experts' opinion on appropriate use of CRd in UK clinical practice, the primary evidence base underlining this appraisal, the reimbursed population of the primary comparator, and where CRd is likely to derive the most benefit for patients;	Restriction of the population of interest to CRd at second line after prior bortezomib results in narrowing of the relevant comparators for CRd to Rd and bortezomib monotherapy, based on the final scope issued by NICE. In the CS, the company focuses on comparison of CRd with Rd. The ERG agrees with the company's rationale for not considering re-challenge with bortezomib (discussed in greater detail in Section 2.3.3). At the time of writing, DVd is recommended only for use within the Cancer Drugs Fund, and, therefore, is outside of the remit of the STA process and is not assessed further by the ERG.

	• In addition, Amgen proposes that a comparison versus DVd remains informative to the decision problem given the high expected uptake of DVd in clinical practice following the CDF recommendation; • Amgen does not propose that CRd will be used as an alternative treatment to bortezomib re-challenge as it is anticipated that use bortezomib will be limited in this population, due to the availability of superior regimens with alternative mechanisms of action and the standard clinical practice of switching between drug classes with different mechanisms of action. This position is aligned with the recent conclusion of the NICE Committee during TA586 where treatment re-challenge with bortezomib was not considered to be an appropriate comparator to lenalidomide plus dexamethasone in the population under consideration. As such, bortezomib is not considered to be a relevant comparator within this appraisal. People who have received at least two prior therapies: • As outlined above, Amgen does not propose that CRd will be used in patients who have received at least two prior therapies; • CRd was previously appraised as a 3rd-line treatment option (NICE TA457) and was not recommended for use in this setting.	
Outcomes	N/A	In the CS, for those receiving CRd at second line after prior bortezomib, unadjusted and adjusted analyses are reported for only progression-free survival and overall survival for CRd versus Rd, with analyses based on more mature data than presented in an earlier TA evaluating carfilzomib in the treatment of multiple myeloma (TA457[1]). Results for response rate, time to next treatment, health-related quality of life and adverse effects are presented for the full population of the ASPIRE RCT.[9] Analysis of response to treatment by the Independent Review Committee and capture of health-related quality of life outcomes ceased on demonstration of a benefit in PFS and, thus, results are based on data from the interim analysis (June 2014). Given that progression-free survival and overall survival are the only clinical outcomes informing the economic analysis, the ERG considers that no

		clinically important estimates of comparative effectiveness for the subgroup of interest have been omitted from the CS. As part of the clarification process, the company provided estimates of relative treatment effect for PFS and OS based on a revised subgroup requested by the ERG (discussed in Section 2.3.1)
Economic analysis	N/A	N/A
Subgroups to be considered	Amgen propose to consider a subgroup of the marketing authorisation as the primary population of interest in this appraisal. Specifically, patients who have received prior bortezomib are the most appropriate population for consideration given: • this positioning is aligned with clinical expert opinion on the optimal use of CRd in UK clinical practice; • the most relevant comparator, Rd, is recommended by NICE in this subgroup and a comparison is supported by robust head-to-head evidence; and • in this position CRd is likely to derive the most benefit for patients.	The ERG considers it appropriate to present the *post hoc* subgroup as the primary population for the decision problem that is the focus of this STA (discussed in greater detail in Section 2.3.1).
Special considerations, including issues related to equity or equality	N/A	N/A

Abbreviations: CDF, Cancer Drugs Fund; CRd, Carfilzomib in combination with lenalidomide and dexamethasone; CS, company submission; DVd, Daratumumab in combination with bortezomib and dexamethasone; ERG, Evidence Review Group; NHS, National Health Service; N/A, not applicable; NICE, National Institute for Health and Care Excellence; Rd, Lenalidomide and dexamethasone; TA, Technology Appraisal.

2.3.1 Population

In 2016, the marketing authorisation for carfilzomib was extended as follows, "*Kyprolis [carfilzomib] in combination with either lenalidomide and dexamethasone or dexamethasone alone is indicated for the treatment of adult patients with multiple myeloma who have received at least one prior therapy*".[10] In line with the marketing authorisation, the final scope issued by NICE specifies the population of interest for this part review of a previous technology appraisal (TA; TA457[8]) to be adults with multiple myeloma who had received at least one prior therapy, with no restriction to a particular line of treatment (Table 1).[1]

In TA457, the company submitted evidence on the clinical and cost effectiveness for carfilzomib in a doublet (Cd) and a triplet (CRd) regimen in the management of multiple myeloma at specific lines within the treatment pathway:[11]

- Cd at second line;
- CRd at second line (prior therapy comprised bortezomib);
- CRd at third line (prior therapy did not include lenalidomide or carfilzomib).

After reviewing the evidence, NICE recommended:[8]

- Cd as an option for treating multiple myeloma in adults, only if:
    - people have had only 1 previous therapy, which did not include bortezomib; and
    - the company provides carfilzomib with the discount agreed in the patient access scheme.

NICE did not recommend the triplet combination of CRd at 3L, citing that overall survival (OS) data were immature, the life expectancy criterion for the end of life consideration was not met and the incremental cost-effectiveness ratios (ICERs) were higher than normally accepted as a cost-effective use of NHS resources.[8]

During the decision-making process, clinical experts present at the Committee meeting fed back that consideration of Cd and CRd and second and third line settings, respectively, was appropriate.[12] Thus, additional details on deliberation on the clinical and cost effectiveness of CRd at second line as part of TA457 are not available in the committee papers.

In the part review reported here, the company presents evidence on CRd at only second line after prior bortezomib in the first-line setting. The company's reasons for focusing on use of CRd at in this setting are:

- the clear unmet need for triplet therapies that target multiple pathways and enable deeper and more durable responses, as well as improved survival outcomes, earlier in the pathway;
- feedback from clinical experts that CRd will offer the greatest benefit to patients in the second-line setting;
    - in the pivotal ASPIRE trial, patients at second line demonstrated improved clinical outcomes compared with later lines, which supports the value of CRd being used early in the pathway;
- an alignment with the reimbursement criteria of the most relevant comparator (Rd), which is supported by a phase 3 randomised comparison;
- the subgroup for which CRd offers the greatest economic value given the substantial clinical benefit observed in this population.

The ERG's clinical experts agree with the company that there remains an unmet need for clinically effective treatments at second line for the management of multiple myeloma, and that CRd is likely to offer the most benefit at the proposed position.

In the CS, the company highlights that the Appraisal Committee for TA457 determined the evidence presented on use of CRd at 3L to be insufficient to establish cost-effectiveness in that setting, in part due to uncertainty arising from immature OS data from ASPIRE. Despite there now being more mature OS data for CRd at 3L, given that the triplet combination was not recommended as an option at this position in the treatment algorithm, the ERG considers it a pragmatic decision for the company to no longer pursue use of CRd at 3L in multiple myeloma.

Given that Cd has been recommended as a second-line treatment for those who received a regimen not including bortezomib, the ERG's clinical experts agree with the company's restriction of use of CRd to second line after prior bortezomib-based therapy.

In support of the proposed positioning of CRd, the company presented estimates of progression-free survival (PFS) and OS derived from a subgroup described as having received one prior therapy with bortezomib. The ERG noted that the subgroup included a proportion of people who had not received bortezomib as part of their last round of therapy ([redacted]), as well as people who had

undergone treatment with lenalidomide in their last regimen ([redacted]). Given that the company is positioning CRd at the second-line setting and after prior bortezomib, as part of the clarification process, the ERG requested that the company generate a new subgroup comprising people who had undergone only one round of therapy that was bortezomib-based and who had not received prior lenalidomide, and to provide revised estimates of PFS and OS for the new subgroup (discussed in greater detail in Section 3.3). The ERG's requested exclusion of those who had received lenalidomide as part of their first-line treatment regimen because no lenalidomide-based regimen is recommended by NICE as a treatment option in this setting. As part of the clarification process, the company highlighted that the subgroup presented in the CS comprised those who had received one prior regimen and had received prior bortezomib. The company commented that the inclusion of those who had not received bortezomib in their last regimen is a consequence of the definition of "last regimen" implemented in ASPIRE. The ERG could not locate a definition for "last regimen" in the CS or CSR. The ERG agrees with the company that people in England could receive a lenalidomide-based regimen at first line as per NICE guidance, if they are judged to be ineligible for SCT, but the available combination does not include bortezomib. Therefore, given the proposed position of CRd in the treatment pathway, the ERG maintains that its requested subgroup more closely reflects the characteristics of people who would likely be eligible for treatment with CRd in clinical practice in England.

In terms of the relevant comparators for CRd at second line (discussed in greater detail in Section 2.3.3), the ERG notes that applying recommendations from TA586[7] on use of Rd at second line could further confine the population who would be eligible for CRd to those who are deemed to be ineligible for SCT at the time of assessment or who cannot tolerate thalidomide. The ERG notes the two studies that informed TA586 included people who had undergone SCT and who had received prior thalidomide.[13, 14] The evidence informing the TA586 was derived from the full trial populations and not from the subgroup of those who were ineligible for SCT or who could not tolerate thalidomide. Additionally, eligibility for SCT or whether a person could tolerate thalidomide were not inclusion criteria for ASPIRE,[9] and, furthermore, have not been specified as baseline characteristics of the subgroup requested by the ERG.

2.3.2 Intervention

The dosing schedule for each drug comprising the triplet regimen of CRd as administered in ASPIRE,[9] and as reported in the Summary of Product Characteristics (SmPC) for carfilzomib,[15] is presented in Table 3. As noted in Table 1, in ASPIRE, use of carfilzomib, but not lenalidomide or dexamethasone,

was restricted to 18 cycles, which is not in line with the marketing authorisation. The ERG's clinical experts fed back that they might consider continuing beyond 18 cycles, if available as an option, for some patients after carrying out a benefit–risk assessment, as advised in the SmPC: [8, 15] data are limited on the tolerability and toxicity of carfilzomib beyond 18 cycles.[8, 15]

The company reports that, in ASPIRE, carfilzomib was administered for a median of 18 cycles (range: 1 to 18 cycles) and a median duration of 72 weeks (range: 1 to 93.1 weeks), which corresponded to the maximum protocol-defined carfilzomib treatment duration.[11] The median relative dose intensity of carfilzomib was 93.7%.

Table 3. Dose and schedule of treatment for carfilzomib, lenalidomide and dexamethasone

Treatment	Route of administration	Dose	Regimen	Treatment duration
Carfilzomib	IV (10 minute infusion)	Starting dose of 20 mg/m^2 on days 1 and 2 of cycle 1 (maximum dose 44 mg). If tolerated, dose should be increased to target dose of 27 mg/m^2 (maximum dose 60 mg).	**Cycles 1–12**: Given on days 1, 2, 8, 9, 15 and 16 of each 28-day treatment cycle.[a] **Cycles 13–18**: Given on days 1, 2, 15 and 16 of each 28-day cycle.	ASPIRE:[9] given for a maximum of 18 cycles, unless discontinued early for disease progression or unacceptable toxicity Median treatment duration: 72 weeks
Lenalidomide	Oral	25 mg	Daily on days 1–21 of each 28-day cycle.	ASPIRE:[9] could be continued after 18 cycles until treatment until progression of disease or unacceptable toxicity. Median treatment duration: • CRd group: 85 weeks; • Rd group: 57 weeks
Dexamethasone[b]	Oral or IV	40 mg	Days 1, 8, 15, and 22 of each 28-day cycle.	ASPIRE:[9] could be continued after 18 cycles until treatment until progression of disease or unacceptable toxicity. Median treatment duration: • CRd group: 80 weeks; • Rd group: 49 weeks

[a] Each 28-day cycle is considered one treatment cycle.
[b] Dexamethasone should be administered 30 minutes to 4 hours before carfilzomib.
Abbreviations: IV, intravenous; m^2, metre-squared; mg, milligram.

2.3.3 Comparators

In the CS, the company presents a matching adjusted indirect comparison on comparative clinical effectiveness of CRd versus daratumumab (D) in combination with bortezomib (V) and dexamethasone (d; DVd) at second line after prior bortezomib. At the time of writing, DVd is recommended only for use within the Cancer Drugs Fund,[6] and, therefore, is outside of the remit of the STA process and is not assessed further by the ERG.

As per the final scope issued by NICE, the relevant comparators for management of multiple myeloma at second line are:[1]

- Cd for those who have received only one previous therapy that did not include bortezomib;
- Rd for those who have received only one prior therapy that included bortezomib;
- bortezomib monotherapy for those who are at first relapse and who have undergone, or are unsuitable for, bone marrow transplantation.

In the CS, the company presents evidence on the comparative clinical effectiveness of only CRd versus Rd for the subgroup of interest, which is derived from the key RCT, ASPIRE, comparing the two treatment regimens. With a focus on implementation of CRd at second line for those whose regimen at first line included bortezomib, Cd is no longer a relevant comparator.

TA586[7] deemed that re-challenge with bortezomib-based therapy was not an appropriate comparator for Rd in the population under consideration. Thus, bortezomib monotherapy is not considered to be a relevant comparator for the population that is the focus of this STA. The ERG's clinical experts agree that people receiving a bortezomib-based regimen at first line would not undergo subsequent re-challenge with bortezomib monotherapy.

The ERG agrees that Rd is the only relevant comparator for CRd at second line after prior bortezomib and, to avoid confusion, reiterates that recommendations from TA586 specify that the population of interest for use of Rd in this setting is limited to those who cannot have a SCT (at the time of assessment) or cannot tolerate thalidomide, and who have already had bortezomib.[7] The ERG also emphasises that the evidence informing TA586 was derived from the full trial population of two studies that included people who had undergone SCT and treatment with thalidomide.[13, 14] The limitation on use of Rd at second line, as in whether Rd is considered only for those ineligible for SCT or who cannot tolerate thalidomide or also includes those who undergo SCT, is not clear from the final scope issued by NICE for the decision problem that is the focus of this STA.[1] The ERG highlights

that, at the time of writing, the NICE pathway for management of multiple myeloma lists no treatment option available through routine commissioning at second line for those who undergo SCT.[16]

3 Clinical effectiveness

The sections below discuss the evidence submitted by the company in support of the clinical effectiveness of carfilzomib (C) in combination with lenalidomide (R) and dexamethasone (d) as a second-line treatment for adults with multiple myeloma whose previous therapy included bortezomib (rationale for narrowing of population outlined in Section 2.3). The Evidence Review Group (ERG) has critiqued the details provided on:

- methods implemented to identify, screen and data extract relevant evidence;
- clinical efficacy of CRd in the subgroup of interest;
- assessment of comparative clinical effectiveness of CRd against relevant comparators in the subgroup of interest;
- safety profile of CRd.

A detailed description of an aspect of the company submission (CS) is provided only when the ERG disagrees with the company's assessment or proposal, or where the ERG has identified a potential area of concern that the ERG considers necessary to highlight for the Committee.

3.1 Critique of the methods review

The company undertook a broad systematic literature review (SLR) with the objective of identifying randomised controlled trials (RCTs) assessing the clinical efficacy and safety of carfilzomib and other therapies in the treatment of multiple myeloma. The company's SLR was conducted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines and The Cochrane Collaboration.[17, 18] Full methods and results of the SLR are reported in Appendix D of the CS and a summary of the methods with comments from the ERG about the appropriateness of the methods adopted are presented in Table 4.

The purpose of the SLR was to identify all relevant studies that could inform the comparison of CRd with other interventions for multiple myeloma. As stated in earlier sections, Rd is the only comparator relevant to this appraisal as daratumumab with bortezomib and dexamethasone (DVd), the only other treatment option at second line for patients previously treated with bortezomib, is currently available through the Cancer Drugs Fund (CDF) and not through routine commissioning. Relevant studies identified in the SLR are therefore limited to those of CRd and Rd.

Sixty three studies reported across 397 publications were identified for inclusion in the SLR, however, these included studies assessing any of the broad list of interventions specified in the inclusion criteria (CS, Appendix D, Table 17). One study relevant to the decision problem was identified (ASPIRE),[9] providing direct evidence on the clinical effectiveness of CRd versus Rd. All other studies were not described or discussed further in the CS.

Overall, the ERG found the company's SLR to be of reasonable quality and likely to have identified all studies relevant to the decision problem, despite limiting inclusion to English-language publications.

Table 4. Summary of ERG's critique of the methods implemented by the company to identify evidence relevant to the decision problem

Systematic review step	Section of CS in which methods are reported	ERG assessment of robustness of methods
Data sources	Appendix D.1.1	**The ERG considers the sources and dates searched appropriate.** MEDLINE, EMBASE, PubMed, The Cochrane Library, latest search date: 11 August 2019. Trial registries (ISRCTN registry, WHO ICTRP, clinicaltrials.gov), conference proceedings (ASH, ASCO, ESMO, EHA, IMW), regulatory bodies (EMA, FDA), HTA agencies (NICE, CADTH, SMC, AWMSG), reference lists of reviews. Latest search update: August 2019
Literature searches	Appendix D.1.1, Tables 1–16	**The ERG is satisfied that searches would have identified all evidence relevant to the decision problem.** Search strategies combined comprehensive terms for the population and interventions, medical subject headings, and study design filters
Inclusion criteria	Appendix D.1.1, Table 17	**The ERG considers it likely that no relevant evidence was excluded based on the eligibility criteria used.** Inclusion criteria were broader than the NICE final scope, especially listed interventions of interest, which were considerably broader than the scope and the company's positioning of CRd. No explanation was provided for the rationale for the broad inclusion criteria, or for the subsequent exclusion of the majority of studies. The ERG assumes studies were excluded because they were not relevant to the decision problem. Limited to English-language publications.

Screening and data extraction	Appendix D.1.1, Figure 1	**The ERG considers the methods for screening described to be robust. Details on data extraction were not reported.** Independent duplicate screening and data extraction by two reviewers against predefined criteria; discrepancies resolved by consensus/with a third reviewer, screening results summarised in a PRISMA diagram.
Tool for quality assessment of included study or studies	Appendix D.3, Table 18	**The ERG agrees with the quality assessment tool used for the key trial, ASPIRE, but the company's assessment lacks details to support the assessment.** It is unclear if quality assessment was done by one or two reviewers and, if so, whether the assessments were done independently. ASPIRE was assessed based on the NICE guidance for companies. Limited details were provided in the CS for the judgement on each of the questions. However, the ERG notes that greater detail on the quality assessment is available in TA457.[8]

Abbreviations: ASCO, American Society of Clinical Oncology; ASH, American Society of Hematology; AWMSG, All Wales Medicines Strategy Group; C, carfilzomib; CADTH, Canadian Agency for Drugs and Technologies; CS, company submission; d, dexamethasone; EMA, European Medicines Agency; ERG, Evidence Review Group; ESMO, European Society for Medical Oncology; FDA, Food and Drug Administration; HTA, Health Technology Assessment; ICTRP, International Clinical Trials Registry Platform; ISRCTN, International Standard Randomised Controlled Trials Number; NICE, National Institute for Health and Care Excellence; R, lenalidomide; SMC, Scottish Medicine Consortium; WHO, World Health Organisation.

3.2 Critique of trials of the technology of interest, the company's analysis and interpretation

The ERG reiterates that the population relevant to the decision problem is a subgroup of those enrolled in ASPIRE,[9] and, moreover, is not a pre-specified subgroup. As a subgroup, and, in particular a *post hoc* subgroup, relevant estimates of comparative clinical effectiveness for CRd versus Rd are at a higher risk of bias than those reported for the full trial population. Finally, ASPIRE was not powered to detect a statistically significant difference in clinically relevant outcomes in the subgroup of interest to the decision problem.

In subsequent sections, the ERG focuses on aspects of trial design, conduct and external validity of ASPIRE that are of import to this STA because the listed areas have previously been covered in greater depth in TA457,[8] the original TA evaluating CRd versus Rd for the management of multiple myeloma. The ERG's critique of the internal validity of ASPIRE is available in Table 5. The ERG agrees with the company's assessment of ASPIRE as being at overall low risk of bias, based on the trial conduct and analyses for the full trial population.

Considering the *post hoc* subgroup that forms the basis of the CS, as noted in Section 2.3.1, data are derived from a subgroup in which a proportion of people have not received bortezomib, and others have undergone treatment with lenalidomide, as part of their last treatment regimen (hereafter referred to as 2L prior bortezomib). Estimates of PFS and OS for CRd versus Rd for those forming the 2L prior bortezomib inform the company's base-case analysis of cost effectiveness of CRd. For reasons outlined in 2.3.1, the ERG's preferred subgroup is that comprising people who received

carfilzomib at 2L after one line of prior bortezomib-based therapy and no lenalidomide (2L after prior bortezomib and no lenalidomide). The ERG notes that the two *post hoc* subgroups have similar baseline characteristics for the CRd and Rd treatment groups, and also comparable hazard ratios (HRs) are derived for PFS and OS for CRd versus Rd (discussed in greater detail in relevant sections). Hereafter, the ERG focuses its critique on data and results derived from the subgroup receiving CRd and Rd 2L after prior bortezomib and no lenalidomide. For comparative purposes, data for the 2L prior bortezomib subgroup that informs the company's base case are also presented. Key differences between the two *post hoc* subgroups are highlighted where applicable.

Table 5. Summary of ERG's critique of the design and conduct of ASPIRE, the trial evaluating the technology of interest to the decision problem

Aspect of trial design or conduct	Section of CS in which characteristic is reported	ERG's critique
Trial conduct[9]		
Randomisation	Section B.2.5 (page 31)	Appropriate. Randomisation carried out by IVRS. People randomised 1:1 to CRd versus Rd. Randomisation stratified by: • β2-microglobulin level (<2.5 vs ≥2.5 mg/L); • previous bortezomib therapy (no vs yes); • previous lenalidomide therapy (no vs yes).
Concealment of treatment allocation	Section B.2.5 (page 31)	Appropriate. Treatment allocation concealed through use of IVRS at randomisation.
Baseline characteristics	Section B.2.5 (page 31)	Baseline characteristics were well balanced between CRd and Rd groups in the ITT population. Imbalances in baseline characteristics were noted in the 2L prior bortezomib subgroup, as expected.
Masking appropriate	Section B.2.5 (page 31)	Open label design. However, primary analyses for disease progression-related outcomes (e.g., PFS and ORR) were based on assessment by a blinded IRC, including the primary outcome of PFS.
No difference between groups in treatments given, other than intervention versus control	Section B.2.5 (page 31)	No evidence to suggest that standard of care differed between treatment groups
Dropouts (high drop out and any unexpected imbalance between groups)	Section B.2.5 (page 31)	Low rate of loss to follow-up (1 person lost to follow-up from Rd group).
Outcomes assessed	Section B.2.5 (page 31)	No evidence to suggest that additional outcomes were assessed and not reported. All clinically relevant outcomes reported. Primary outcome was PFS as determined by IRC. Investigator-assessed PFS reported as a secondary outcome. Other secondary outcomes included: • OS; • ORR; • Time to response;

		• Best response; • Disease control rate; • Duration of disease control; • HRQoL (as assessed by EORTC QLQ-C30 GHS/QoL).
ITT analysis carried out	Section B.2.5 (page 31)	Yes ITT population formed the basis for analyses of efficacy and a PP population (all patients who received ≥1 dose of study drug) informed analyses of safety and tolerability.
Subgroup analyses	Section B.2.5 (page 31)	*Post hoc* subgroup analysis forms the basis of the submission that is the focus of this STA.
Statistical analysis plan		
Sample size	Section B.2.4 (page 30)	Calculation informed by median PFS for Rd (high-dose dexamethasone) derived from a phase III study. Sample size of 526 PFS events required at the time of the final analysis to give the desired power.
Power	Section B.2.4 (page 30)	Calculated sample size gives the study 90% power to detect a 33% increase in median PFS associated with CRd compared with Rd (14.9 months with CRd vs 11.2 months with Rd). A 33% increase in median PFS for CRd corresponds to a 25% decrease in risk of progression compared with Rd (i.e., HR 0.75) at a one-sided significance level of 0.025.
Analysis for estimate of effect	Section B.2.4 (page 30) Section B.2.7.7 (pages 47–49)	An interim analysis was performed after approximately 420 events had occurred (80% of the planned total). If there was a significant between-group difference in PFS at the interim analysis, secondary end points would be sequentially tested in the order of OS, ORR, and HRQoL, each at a one-sided significance level of 0.025. PFS and OS were compared between treatment groups with the use of a log-rank test stratified according to the factors used for randomisation. Hazard ratios were estimated by means of a stratified Cox proportional-hazards model. Distributions were summarized with the use of the Kaplan–Meier method. For the *post hoc* subgroup analyses, the company presents an IPW treatment effect: methods implemented in the IPW analysis are discussed in more detail in Section 3.3.1.1.

Abbreviations: C, carfilzomib; CI, confidence interval; CS, company submission; d, dexamethasone; EORTC QLQ-C30 GHS/QoL, European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire Core 30 Global Health Status/Quality of Life; ERG, Evidence Review Group; IPW, inverse probability weighted; IRC, Independent Review Committee; ITT, intention to treat; IVRS, interactive voice response system; mg, milligram; PFS, progression-free survival; ORR, overall response rate; OS, overall survival; PP, per protocol; R, lenalidomide; STA, Single Technology Appraisal.

3.2.1 Baseline characteristics

Baseline characteristics for the subgroup comprising people receiving carfilzomib 2L after prior bortezomib and with no lenalidomide, as provided by the company during clarification at the request of the ERG, are available in Appendix 9.1 (Table 37): baseline characteristics of the full trial population of ASPIRE (Table 38) and the subgroup of 2L prior bortezomib (Table 39) are also presented in Appendix 9.1.

As noted in TA457,[8] baseline characteristics for the full trial population of ASPIRE were well balanced between the treatment groups (Appendix 9.1; Table 38). The ERG's clinical experts highlighted that, as would be expected in a clinical trial, the enrolled population is slightly younger and has a better performance status (i.e., lower ECOG scores) than people typically presenting with multiple myeloma in clinical practice. After discussion, the committee for TA457 concluded that the population forming ASPIRE was generalisable to the UK population likely to be eligible for treatment with carfilzomib.[12]

As would be expected for non-randomised, *post hoc* subgroups, the ERG notes imbalances in some baseline characteristics between those given CRd and those administered Rd after a bortezomib-based regimen and no prior lenalidomide (Appendix 9.1; Table 37). Imbalances that require particular consideration are those characteristics that are considered to be factors that would influence prognosis, and those that are potential modifiers of treatment effect. The company consulted with clinical experts to identify key prognostics factors, a list of which is provided in Section 3.3.1.1: the ERG's advisors agreed with the factors identified by the company as impacting prognosis. One example of a marked imbalance in key characteristics is mean time since initial diagnosis, which is [redacted] for those allocated to Rd compared with those receiving CRd ([redacted] months with CRd vs [redacted] with Rd), which could introduce bias [redacted] of treatment with Rd. However, the ERG also notes that the standard deviation accompanying mean time since initial diagnosis is also [redacted] for the Rd group ([redacted] [95% CI: [redacted]] with CRd vs [redacted] [95% CI: [redacted]] with Rd), suggesting a [redacted] in time since initial diagnosis in those forming the Rd group compared with the CRd arm (Appendix 9.1; Table 37): 95% CI calculated by the ERG. By contrast, in the 2L prior bortezomib subgroup, median time since diagnosis is [redacted] in the CRd and Rd treatment groups at [redacted] and [redacted] years, respectively, but with a marked difference in the maximum time since diagnosis between groups ([redacted] years with CRd vs [redacted] years with Rd; Table 39).

In the 2L prior bortezomib no lenalidomide subgroup, differences between treatment groups were observed in the proportion of people refractory to prior bortezomib ([redacted]/74 [redacted] with CRd vs [redacted]/66 [redacted] with Rd), prior SCT ([redacted]/74 [redacted] with CRd vs [redacted]/66 [redacted] with Rd) and age of 75 years and older ([redacted]/74 [redacted] with CRd vs [redacted]/66 [redacted] with Rd). Similar differences were noted for the 2L prior bortezomib subgroup (Table 39).

The direction of bias introduced by the differences in baseline characteristics, and the impact on estimates of relative treatment effect, is unclear. To account for the imbalances between treatment groups, the company carried out an inverse probability weighted (IPW) analysis to adjust patient-level data for covariates identified by clinical experts as prognostic factors. The ERG's critique of the methods implemented by the company to carry out the IPW analyses is available in Section 3.3.1.1.

3.2.2 Outcome assessment

As noted in Table 1, revised estimates of comparative clinical effectiveness for CRd versus Rd that are based on more mature data are available in the CS for only PFS and OS, for both the ITT population of ASPIRE and the subgroup of those receiving treatment at 2L after prior bortezomib. Analysis of response rates by the Independent Review Committee (IRC) and capture of health-related quality of life (HRQoL) outcomes ceased on demonstration of a benefit in PFS and, thus, results for those outcomes are based on data from the interim analysis (June 2014).

Estimates of comparative clinical effectiveness for clinical outcomes at the interim analysis have been reported and critiqued as part of TA457.[8] Here, the ERG focuses on the robustness of the effect estimates generated for PFS and OS for the subgroup of interest to the STA.

For completeness, the ERG provides a brief summary of response rates, HRQoL, time to next treatment (TTNT) and adverse effects in the ITT population of ASPIRE. As noted in TA457, statistical significance of the difference between groups in secondary outcomes was only to be tested in a fixed sequence if the null hypothesis for the primary outcome of PFS (interim or final) was rejected. At the interim analysis, a statistically significant difference was found between treatment groups in PFS and, thus, significance of difference between groups for other outcomes was tested, starting with OS. At the interim analysis, the p-value boundary for OS to trigger testing of the next outcome in the sequence was not met and so formal statistical testing for the remaining secondary endpoints was precluded. Thus, any reported p-values for ORR, HRQoL and TTNT are descriptive in nature.

3.3 Clinical effectiveness results

As noted earlier, the ERG's preferred subgroup is that in which people received CRd 2L after prior bortezomib and with no prior treatment with lenalidomide. The company's base case is based on their preferred subgroup for CRd at 2L, which includes a proportion of people who received prior lenalidomide. For reference purposes, unadjusted and IPW-adjusted estimates of PFS and OS for the subgroup preferred by the company (received CRd at 2L after bortezomib) are also presented.

Estimates of comparative treatment effectiveness for PFS and OS reported in the CS are based on an additional 3 years of follow up compared with the data presented in TA457. Median follow-up at the time of the interim analyses (June 2014), which were evaluated in TA457, and at the time of the primary OS analysis (data cut-off of April 2017), which are reported here, are available in Table 6. Event rates for the full trial population at the time of the analysis are also provided in Table 6. The ERG notes that the sample size required at the time of the final analysis to give the desired power was 526 PFS events (Table 5). At the time of the primary OS analysis, 516 PFS events had occurred. As a statistically significant result was identified at the interim analysis of PFS, and also at the primary OS analysis, the ERG considers the results in the ITT population to be robust.

In TA457, the Committee recognised the limitations and uncertain outcomes associated with using data derived from subgroups that were not prespecified. The Committee also acknowledged that the company had attempted to mitigate against the uncertainty from using *post hoc* data through identifying additional covariates through a Cox proportional hazards model, and adjusting imbalances in baseline characteristics accordingly to provide estimates of efficacy for carfilzomib and its comparators. For consistency, here, the ERG focuses on the estimates of PFS and OS for CRd versus Rd generated from IPW-adjusted analyses based on data from the ERG's preferred subgroup. However, the ERG recognises that generation of *post hoc* subgroups renders the data produced to be observational in nature, and that any analyses derived from *post hoc* subgroups are considered to be hypothesis generating. In the CS, for the ITT population of ASPIRE, the company comments that, "*based on stepwise Cox regression modelling, there was a lack of evidence of treatment-covariate interactions for PFS suggesting an overall consistent treatment effect across the baseline covariate subgroups*". As the relative treatment effect for CRd versus Rd is consistent, irrespective of subgroup, to mitigate against uncertainty associated with *post hoc* subgroup analyses reported here, the ERG considers HRs derived from the ITT population of ASPIRE are informative. For comparison purposes, PFS and OS results for the ITT population in ASPIRE are presented alongside those for those receiving CRd at 2L after prior bortezomib, but with no lenalidomide (Sections 3.3.1.2 and

3.3.1.3 for PFS and OS, respectively). The ERG notes that, although the relative treatment effect of CRd versus Rd is constant irrespective of treatment group, there could be differences between treatment groups in absolute gain or loss of time to progression or death in the individual subgroups.

Table 6. Summary of median follow-up times and number of events on which PFS and OS analyses are based for the ITT population of ASPIRE

	PFS		OS	
Data cut-off	**CRd**	**Rd**	**CRd**	**Rd**
Interim (June 2014), follow-up months (95% CI)	31.4 (30.7 to 31.9)	30.1 (28.8 to 31.4)	32.3 (31.7 to 33.2)	31.5 (30.8 to 32.5)
Number of events	431		305	
Primary OS (April 2017), follow-up months	48.8 ([redacted]	[redacted]	[redacted]	[redacted]
Number of events	516		513	

Abbreviations: C, carfilzomib; CI, confidence interval; CS, company submission; d, dexamethasone; PFS, progression-free survival; OS, overall survival; R, lenalidomide.

3.3.1 *Progression-free survival and overall survival in* post hoc *subgroup*

3.3.1.1 *Inverse probability weighted analysis to derive effect estimates for relevant subgroup*

To account for imbalances in baseline characteristics, and to address the limitations associated with use of data derived from a *post hoc* subgroup, the company carried out an IPW to generate estimates for PFS and OS for CRd versus Rd in the subgroup of interest to the STA: details of methodology followed are available in Section B.2.7.2 (page 47) of the CS. The ERG agrees with the company's approach to mitigate against the issues arising from use of a *post hoc* subgroup.

Based on details available in the CS, the ERG had reservations on two aspects of the IPW:

- covariates accounted for in the IPW (discussed in subsequent section);
- use of Cox regression model to select covariates for the IPW (discussed in subsequent section).

Covariates accounted for in the IPW

In the CS, the company reports that they implemented a stepwise (backwards and forwards) Cox regression model to select covariates that should be accounted for in the analyses from a list of characteristics identified by clinical experts as being prognostic of outcomes in multiple myeloma.

In terms of the covariates assessed for inclusion in the regression analyses, the company reports that clinical experts identified the characteristics below as influencing prognosis:

- number of prior lines of therapy;
- prior exposure to lenalidomide or bortezomib;
- age (<65 vs ≥65 years);
- Eastern Cooperative Oncology Group performance status score (0 vs 1 or 2);
- creatinine clearance (<50, 50–80, or ≥80 mL/min);
- time since diagnosis;
- time since last relapse;
- International Staging System stage (I vs II or III);
- prior SCT;
- β2-microglobulin (<3.5 vs ≥3.5 mg/L);
- refractory to last prior treatment;
- cytogenetic risk status (high, standard, or unknown/missing).

As highlighted earlier, the ERG's clinical experts agreed that the characteristics listed are those likely to influence prognosis, and went on to comment that the extent of impact on prognosis will differ across the characteristics.

The ERG recognises that the most relevant characteristics have been considered by the company but comments that it is unclear from the CS what criterion has been applied to add or remove a covariate from the model for IPW analysis based on the company's preferred subgroup. The ERG notes that similar issues were raised in TA457, with the ERG commenting, "*The ERG has concerns about the lack of justification and use of a large number of covariates in the Cox proportional hazards models to estimate the efficacy of Cd and CRd in these* post hoc *subgroups of ENDEAVOR and ASPIRE*".[12]

During clarification, the company helpfully outlined the approach used to identify covariates to be included in the IPW estimate of comparative treatment effectiveness for PFS and OS, and reported the HR and 95% CI for CRd versus Rd after adjustment for an individual covariate, as well as an estimate of effect adjusted for all retained covariates. Details of estimates of effect for CRd versus Rd for included covariates are presented in Table 7.

Although the ERG predominantly considers the company's approach to identification of relevant covariates appropriate, the ERG considers it important to highlight that the regression analyses [redacted] for some covariates. The ERG considers that the results could suggest that the characteristics are potential treatment effect modifiers. In particular, adjustment for prior SCT and for β2-microglobulin level suggest that, compared with Rd, treatment with CRd is associated with a [redacted] (Table 7): [redacted] (Table 7). Similar [redacted] was noted in the IPW analyses presented in the CS relating to the subgroup preferred by the company. During clarification, the ERG queried whether the reported effect estimates were for, [redacted] irrespective of treatment received. The company confirmed that the estimates presented were for CRd versus Rd after adjustment for the individual covariate. The company commented that [redacted]. The ERG agrees that [redacted] are likely to have different characteristics [redacted] but considers that there is no clear clinical rationale [redacted]. As data are derived from *post hoc* subgroup analyses, the ERG emphasises that any inferences from the results are hypothesis generating.

Table 7. Results generated using covariates selected through the Cox proportional hazards regression model based on December 2017 data cut-off (adapted from Table 2 of the company's response to clarification)

Covariate	CRd versus Rd, 2L after bortezomib and no prior lenalidomide	
	PFS HR (95% CI)	OS HR (95% CI)
Treatment (CRd vs Rd)	[redacted]	[redacted]
Prior stem cell transplantation (yes vs no)	[redacted]	[redacted]
Age (≥65 vs <65)	[redacted]	[redacted]
ECOG status (1-2 vs 0)	[redacted]	[redacted]

Creatinine clearance (≥50 - <80 vs other)		
Creatinine clearance (≥80 vs other)		
Time from diagnosis		
Time from last relapse		
ISS stage (II-III vs I)		
β2-microglobulin (≥3.5 vs <3.5 mg/L)		
Refractory to last prior treatment (yes vs no)		

Abbreviations: 2L, second line; C, carfilzomib; CI, confidence interval; d, dexamethasone; ECOG, Eastern Cooperative Oncology Group; HR, hazard ratio; ISS, International Staging System; OS, overall survival; PFS, progression-free survival; R, lenalidomide.

Use of Cox regression model to select covariates for the IPW

IPW utilises a logistic regression model to estimate the probability (propensity score) for a particular person of receiving a specific treatment (e.g., CRd or Rd) given confounding variables (covariates) of the patient.[19] The inverse of the estimated probabilities is applied to reweight the population and adjust for imbalances in the included covariates. A key assumption is that all confounders have been measured and properly modelled in the regression model. The ERG has reservations around the use of the Cox regression model to select the variables that are subsequently modelled by logistic regression. The ERG considers it could be more appropriate to select covariates using the same regression method applied to generate the IPW.

The ERG is unaware of formal guidance on how adjusted survival estimates in oncology should be generated when there are imbalances in baseline characteristics between treatment groups. Technical Support Document (TSD) 14 produced by the NICE Decision Support Unit (DSU) outlines various methodologies to survival analysis and provides guidance on assessing suitability of each method for a particular case, but methods to account for analyses based on adjusting survival for imbalances arising from use of *post hoc* subgroup analyses are not covered in this TSD.[20] Additionally, TSD17 outlines use of observational data to inform estimates of treatment effectiveness and covers methods on how to adjust for confounders, including IPW. Guidance in TSD17 highlights that the utility of IPW depends on how well the model for the propensity score predicts the probability of treatment, and that the propensity score should be sufficiently flexible, which can be achieved using a parametric model (e.g., probit or logit): the choice of parametric model can have an impact on the results.[21]

During clarification, the company gave a detailed description of the methods followed to generate the IPW-adjusted estimates of PFS and OS, which were as follows:

- the treatment indicator and the clinician-identified covariates were considered in a Cox proportional hazards model, and an automated stepwise variable selection procedure was performed using the stepAIC function in R, which minimises the AIC. Treatment-covariate interactions were not tested due to constraints related to sample size.
- a logistic regression model was subsequently conducted in which the treatment indicator was defined as the dependent variable and the covariates identified in the stepwise selection Cox model were used as independent variables. The retained variables for PFS and OS are summarised in Table 10.
- survival analyses were conducted on the weighted dataset.

The ERG highlights that, as requested during the clarification process, the company additionally used logistic regression to select covariates for which to retain in the IPW analyses, and provided results for IPW- adjusted PFS and OS.

In support of the use of the Cox regression model for covariate selection, as part of the clarification process, the company stated that, "*the logistic regression model can be interpreted as an approach where one searches and adjusts for covariates that are strongly related to the treatment received. However, the subgroup data is coming from a well-conducted randomised clinical trial where patients were randomly assigned to treatments. Therefore, in our view a more appropriate approach is to identify which covariates are strongly related to the outcome and adjust for imbalances in these covariates*".

The company also commented that resulting AIC for the model based on the stepwise selection within the logistic model were higher than those obtained when applying the stepwise selection within the Cox proportional hazard model, suggesting that selection of covariates using the Cox proportional hazards might provide a better representation of the data. The ERG disagrees with the company that the lower AIC associated with the final Cox proportional hazards regression implies that the model is a better representation of the data. AIC estimates the quality of each model relative to each of the other models in that analysis, and the lowest score identifies the best fitting model for that data set. The ERG considers that using different regression techniques to identify the covariates generates different data sets and thus the AIC scores are not directly comparable.

The ERG notes that similar estimates of comparative treatment effectiveness for PFS and OS are generated from IPW analyses adjusted for covariates identified using the Cox proportional hazards regression and the logistic regression. In their base case, the company utilises estimates derived

from covariates selected with the Cox proportional hazards regression model. For completeness, the ERG presents results from both analyses (Sections 3.3.1.2 and 3.3.1.3).

3.3.1.2 Progression-free survival

The ERG notes that PFS assessed by the IRC was the primary outcome in ASPIRE, and was met at the time of the interim analysis (data cut-off June 2014). Thus, as highlighted by the company, assessments of PFS after June 2014 are based on determinations by investigators and are not supported by determinations from an IRC. Given that ASPIRE was an open label study, the ERG notes that the determination of progression subsequent to June 2014 is at increased risk of bias.

At a cut-off date of December 2017, PFS for the ERG's preferred subgroup was based on a median follow-up of [redacted] and [redacted] months in the CRd and Rd groups, respectively (Table 8), which is [redacted] to that of the ITT population of ASPIRE (cut-off date of April 2017): Kaplan–Meier (KM) plot for PFS for the ITT population of ASPIRE is presented in Figure 2.

Without adjustment for imbalances in baseline characteristics, CRd was associated with an absolute increase in median PFS of [redacted] months compared with Rd (median PFS [months]: [redacted] [95% CI: [redacted] to [redacted]] with CRd vs [redacted] [95% CI: [redacted] to [redacted]] with Rd). Treatment with CRd was associated with a [redacted] in risk of progression or death (PFS), with the difference between CRd and Rd [redacted] in PFS compared with Rd (HR [redacted]; 95% CI: [redacted] to [redacted]; Table 8 and Figure 3).

After IPW adjustment for imbalances in key baseline characteristics (either method for stepwise selection of covariates), the [redacted] associated with treatment with CRd in the unadjusted analysis [redacted]. The difference between CRd and Rd in PFS [redacted], with [redacted] in risk of progression or death from [redacted] to [redacted] or to [redacted] for Cox proportional hazards and logistic regression analyses, respectively (Table 8 and Figure 4).

The ERG notes that the absolute difference in PFS is [redacted] when the restricted mean value is considered, with CRd associated with an improvement in PFS of [redacted] months [redacted] Rd (Table 8). The [redacted] in the median and mean values of PFS [redacted].

Table 8. Estimates of effect for progression-free survival for ITT population of ASPIRE and subgroup of those receiving CRd at 2L after prior bortezomib and no prior lenalidomide (ERG favoured subgroup) (adapted from Table 12 of the CS and Tables 4 and 5 of the company's response to clarification)

	ASPIRE ITT population April 2017 cut off		2L prior bortezomib/no prior lenalidomide December 2017 cut off PFS determined by investigators				2L prior bortezomib December 2017 cut off PFS determined by investigators			
			Unadjusted		IPW adjusted		Unadjusted		IPW adjusted	
	CRd (N = 396)	Rd (N = 396)	CRd (N = 74)	Rd (N = 66)	CRd (N = 68)	Rd (N = 69)	CRd (N = 93)	Rd (N = 73)	CRd (N = 82)	Rd (N = 81)
Total number of events, n (%)	244 (61.6%)	272 (68.7%)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
• Progression	[redacted]	[redacted]	[redacted]	[redacted]	NR	NR	[redacted]	[redacted]	NR	NR
• Death	[redacted]	[redacted]	[redacted]	[redacted]	NR	NR	[redacted]	[redacted]	NR	NR
Median PFS (95% CI), months	26.1 (23.2 to 30.3)	16.6 (14.5 to 19.4)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Restricted mean PFS time (95% CI) [SE]	NR	NR	[redacted]	[redacted]	NR	NR	[redacted]	[redacted]	NR	NR
Median follow-up (95% CI), months	48.8	48.0	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Mean follow-up (95% CI), months	NR	NR	[redacted]	[redacted]	NR	NR	[redacted]	[redacted]	NR	NR
HR CRd vs Rd (95% CI) unadjusted	0.659 (0.553 to 0.784)		[redacted]		NA		[redacted]		NA	

HR CRd vs Rd (95% CI) adjusted for stratification variables[a]	NR	[redacted]	NA	[redacted]	NA
HR CRd vs Rd (95% CI) IPW-adjusted (stepwise selection within Cox model)	NA	NA	[redacted][b]	NA	[redacted][d]
HR CRd vs Rd (95% CI) IPW-adjusted (stepwise selection within logit model)	NA	NA	[redacted][c]	NA	[redacted][e]

[a] Stratification factors applied in ASPIRE were: β2-microglobulin level (<2.5 mg/L vs ≥2.5 mg/L), previous therapy with bortezomib (no vs yes), and previous therapy with lenalidomide (no vs yes).
[b] Variables adjusted for: [redacted].
[c] Variables adjusted for: [redacted].
[d] Variables adjusted for: [redacted].
[e] Variables adjusted for: [redacted].
Abbreviations: 2L, second line; AIC, Akaike 's Information Criterion; C, carfilzomib; CI, confidence interval; CS, company submission; d, dexamethasone; ERG, Evidence Review Group; HR, hazard ratio; IPW, inverse probability weighted; ITT, intention to treat; NA, not applicable; NR, not reported; PFS, progression-free survival; R, lenalidomide; SE, standard error.

Figure 2. Kaplan–Meier plot for unadjusted progression-free survival as determined by investigator for the ITT population from ASPIRE based on April 2017 cut off (reproduced from CS, Figure 6, page 40)

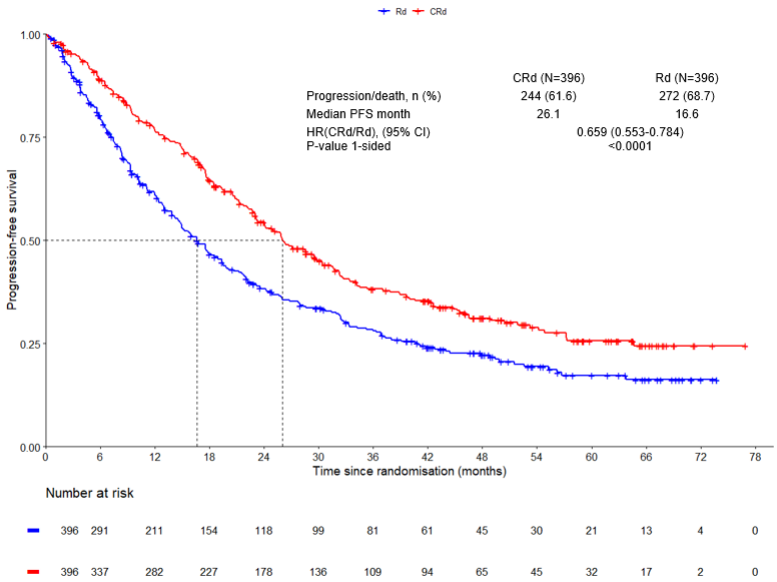


Figure 3. Kaplan–Meier plot for unadjusted progression-free survival for the subgroup receiving CRd at 2L after prior bortezomib and no lenalidomide (reproduced from Figure 1 of the company's response to clarification)

[redacted]

Abbreviations: 2L, second line; C, carfilzomib; d, dexamethasone; R, lenalidomide.

Figure 4. Kaplan–Meier plot for IPW-adjusted progression-free survival for the subgroup receiving CRd at 2L after prior bortezomib and no lenalidomide (reproduced from Figure 2 of the company's response to clarification; covariates selected using Cox proportional hazards regression)

Abbreviations: 2L, second line; C, carfilzomib; d, dexamethasone; R, lenalidomide.

Figure 5. Kaplan–Meier plot of IPW-adjusted PFS in the 1 prior therapy, prior bortezomib subgroup (ASPIRE, 5 December 2017 data cut; reproduced from the CS, Figure 9, page 50)

3.3.1.3 *Overall survival*

More mature data are now available to inform the analysis of OS, with a median follow-up of 67.1 months for the ITT population of ASPIRE and a total of 513 OS events (513/792 [64.8%]; Table 9): KM plot for unadjusted OS for ITT population of ASPIRE is presented in Figure 6.

In the subgroup of those receiving CRd or Rd at 2L after prior bortezomib and no lenalidomide, CRd was associated with an [redacted] in median survival of [redacted] months compared with Rd, and [redacted] in risk of death of [redacted], [redacted], the difference between treatments in OS [redacted] statistical significance (HR [redacted]; 95% CI: [redacted] to [redacted]; Table 9 and Figure 7). However, IPW analysis generates [redacted] for CRd compared with Rd, with [redacted] in risk of death from [redacted] (covariates selected by stepwise logistic regression), with the difference between groups [redacted] (HR [redacted]; 95% CI: [redacted] to [redacted]; Table 9 and Figure 8).

As is expected with OS, results are potentially confounded due to people moving on to non-randomised treatments due to progression of disease. Limited data are available in the CS on subsequent therapies received in the subgroup of interest to the STA. As part of the clarification process, the company provided details on treatments given at 3L for the ERG's preferred subgroup, and treatments given as subsequent treatments for the ITT population of ASPIRE (please see Table 32 in the company's response to clarification questions). The ERG noted that [redacted] at 3L, with [redacted] available in the NHS for the 3L setting (e.g. [redacted]). The proportion of people receiving individual therapies [redacted] across treatment groups, with the [redacted]. [redacted] people in the CRd group were given subsequent treatment with an investigational drug compared with Rd ([redacted] [[redacted]] with CRd vs [redacted] [[redacted]] with Rd). Conversely, a [redacted] ([redacted] [[redacted]] with CRd vs [redacted] [[redacted]] with Rd). The ERG's clinical experts commented that they would likely give bortezomib or ixazomib at third line to someone treated with a non-proteasome-inhibitor containing regimen (e.g., Rd) at second line.

Taking the ERG's reservations around the potential for confounding due to subsequent treatments, the ERG considers that the results for OS should be interpreted with a measure of caution.

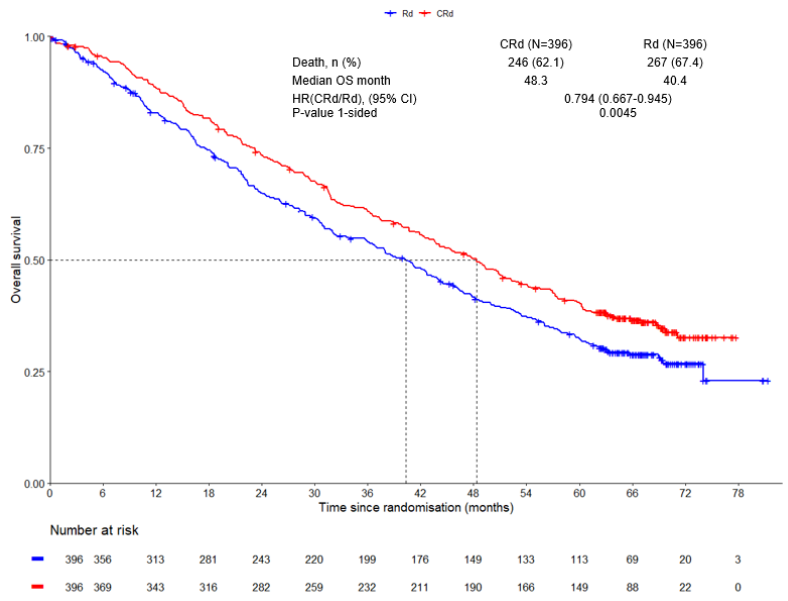
Table 9. Estimates of effect for overall survival for ITT population of ASPIRE and subgroup of those receiving CRd at 2L after prior bortezomib and no prior lenalidomide (ERG favoured subgroup) (adapted from Table 13 of the CS and Tables 8 and 9 of the company's response to clarification)

	ASPIRE ITT population April 2017 cut off		2L prior bortezomib/no prior lenalidomide December 2017 cut off				2L prior bortezomib December 2017 cut off			
			Unadjusted		IPW adjusted		Unadjusted		IPW adjusted	
	CRd (N = 396)	Rd (N = 396)	CRd (N = 74)	Rd (N = 66)	CRd (N = 69)	Rd (N = 68)	CRd (N = 93)	Rd (N = 73)	CRd (N = 82)	Rd (N = 81)
Total number of events, n (%)	246 (62.1%)	267 (67.4%)								
Median OS (95% CI), months	48.3 (42.4 to 52.8)	40.4 (33.6 to 44.4)								
Restricted mean OS time (95% CI) [SE]	NR	NR			NR	NR			NR	NR
Median follow-up (95% CI), months	67.1	67.1								
Mean follow-up (95% CI), months					NR	NR			NR	NR

HR CRd vs Rd (95% CI) Unadjusted	0.794 (0.667 to 0.945)	[redacted]	NA	[redacted]	NA
HR CRd vs Rd (95% CI) adjusted for stratification variables[a]	NR	[redacted]	NA	[redacted]	NA
HR CRd vs Rd (95% CI) IPW-adjusted (stepwise selection within Cox model)	NA	NA	[redacted][b]	NA	[redacted][d]
HR CRd vs Rd (95% CI) IPW-adjusted (stepwise selection within logit model)	NA	NA	[redacted][c]	NA	[redacted][e]

[a] Stratification factors applied in ASPIRE were: β2-microglobulin level (<2.5 mg/L vs ≥2.5 mg/L), previous therapy with bortezomib (no vs yes), and previous therapy with lenalidomide (no vs yes).
[b] Variables adjusted for: [redacted].
[c] Variables adjusted for: [redacted].
[d] Variables adjusted for: [redacted].
[e] Variables adjusted for: [redacted].
Abbreviations: 2L, second line; AIC, Akaike 's Information Criterion; C, carfilzomib; CI, confidence interval; CS, company submission; d, dexamethasone; ERG, Evidence Review Group; HR, hazard ratio; IPW, inverse probability weighted; ITT, intention to treat; NA, not applicable; NR, not reported; OS, overall survival; R, lenalidomide; SE, standard error.

Figure 6. Kaplan–Meier plot for unadjusted overall survival for the ITT population from ASPIRE based on April 2017 cut off (reproduced from CS, Figure 7, page 40)



Abbreviations: 2L, second line; C, carfilzomib; d, dexamethasone; R, lenalidomide.

Figure 7. Kaplan–Meier plot for unadjusted overall survival for the subgroup receiving CRd at 2L after prior bortezomib and no lenalidomide (reproduced from Figure 5 of the company's response to clarification)

Abbreviations: 2L, second line; C, carfilzomib; d, dexamethasone; R, lenalidomide.

Figure 8. Kaplan–Meier plot for IPW-adjusted overall survival for the subgroup receiving CRd at 2L after prior bortezomib and no lenalidomide (covariates selected using Cox proportional hazards regression model; reproduced from Figure 6 of the company's response to clarification)

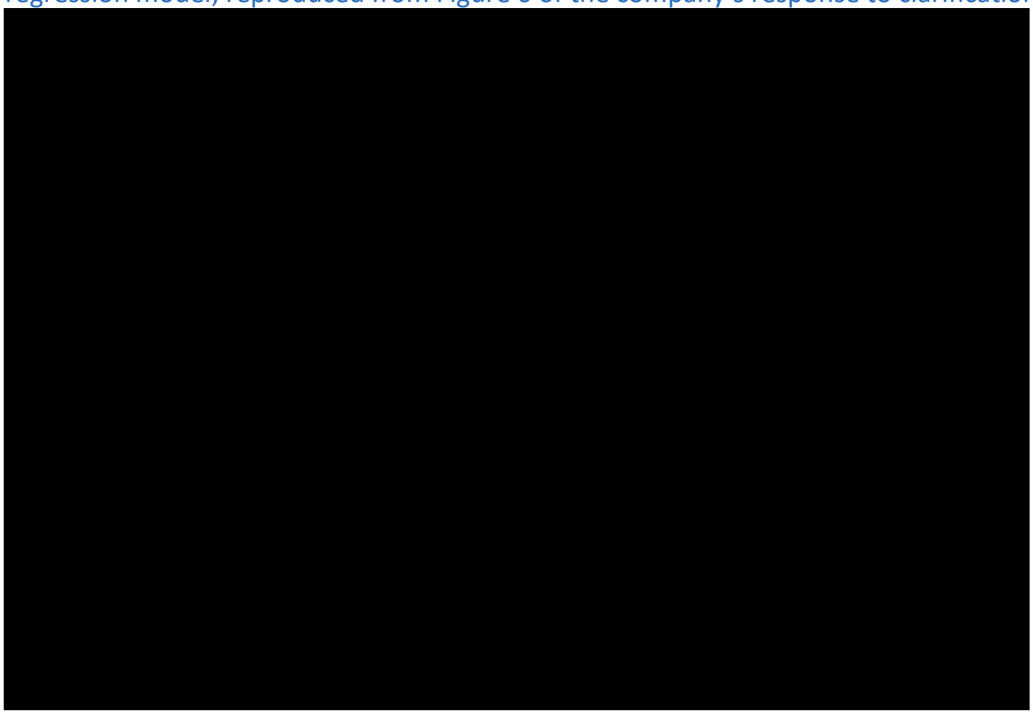


Figure 9. Kaplan–Meier plot of IPW-adjusted OS in the 1 prior therapy, prior bortezomib subgroup (ASPIRE, 5 December 2017 data cut; reproduced from the CS, Figure 9, page 50)

Abbreviations: 2L, second line; C, carfilzomib; d, dexamethasone; IPW, inverse probability weighted; R, lenalidomide.

3.3.2 Summary of other clinically relevant outcomes

A summary of results from ASPIRE for overall response rate and time to next treatment are available in Appendix 9.2.

3.3.2.1 Health-related quality of life

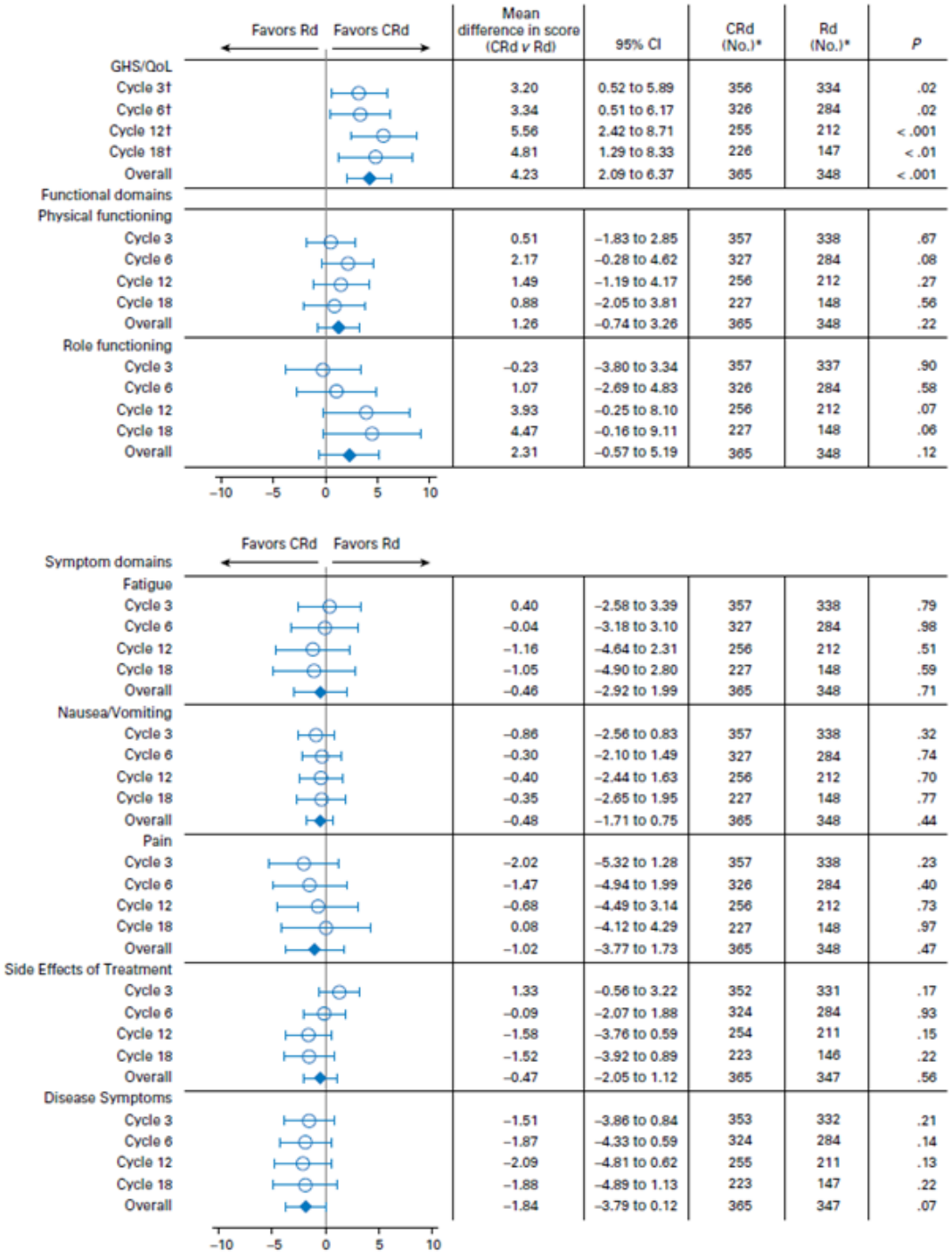
HRQoL was captured in ASPIRE using the EORTC QLQ-C30 questionnaire and the 20-item myeloma-specific EORTC QLQ-MY20 module. The company used the data on collected on EORTC QLQ-C30 score to predict an EQ-5D-3L utility score for each patient through application of a mapping algorithm.[22] As HRQoL data captured from ASPIRE have been used to inform the economic model, the results for HRQoL are presented here.

Analysis of HRQoL is based on results from the interim analysis (June 2014), as, like response, data on HRQoL was no longer collected on demonstration of a benefit of PFS.

Of the 792 people forming the ITT population, 713 (90%) completed at least 1 HRQoL assessment after baseline evaluation and were included in the analyses (CRd, n = 365; Rd, n = 348). Baseline QLQ-C30 and QLQ-MY20 subscale scores were similar between treatment groups.

Over 18 cycles of treatment, global health status scores as assessed using QLQ-C30 GHS/QoL were statistically significantly higher for the group receiving CRd compared with those treated with Rd (two-sided p <0.001; Figure 10). The minimal important difference (MID) for between-group differences on QLQ-C30 GHS/QoL is 5 points.[23-26] Based on the predefined threshold, the MID between CRd and Rd was met at cycle 12 (MID = 5.56) and was approached at cycle 18 (4.81). No statistically significant differences between CRd and Rd were recorded on other components of the HRQoL tools and no other MID was met, but a trend in favour of CRd was observed in differences across subscales (Figure 10).

Figure 10. Treatment difference in EORTC QLQ-C30 and myeloma-specific EORTC QLQ-MY20 module based on the interim analysis data cut-off for the ASPIRE population[a,b] (reproduced from the CS, Figure 8, page 43)



[a] Based on patients completing at least 1 post-baseline HRQoL assessment.

[b] Values shown are the adjusted least squares mean treatment difference in scores from a restricted maximum likelihood-based model for repeated measures under the assumption of missing at random. Scores are adjusted for baseline score, baseline score by visit interaction and the randomisation stratification factors (β2-microglobulin levels [<2.5 mg/L vs ≥2.5 mg/L], prior bortezomib (no vs. yes), and prior lenalidomide (no vs. yes).

Abbreviations: C, carfilzomib; CI, confidence interval; CS, company submission; d, dexamethasone; EORTC, European Organization for Research and Treatment of Cancer; HRQoL, health-related quality of life; R, lenalidomide.

3.3.2.2 *Adverse effects*

The CS (Section B.2.10, pages 62–67) gives a detailed overview of the adverse effects experienced by people enrolled in ASPIRE. Here, the ERG provides an overview of adverse effects, with a focus on those that have been included in the economic model, or those that have been omitted from the model and the ERG considers important to include.

Safety and tolerability data derived from ASPIRE and reported in the CS are based on the data cut-off of 28 April 2017, which includes approximately 3 additional years of follow-up compared with the interim analysis presented in TA457.[8] The company comments that the results based on longer follow-up are consistent with those from the interim analysis presented in TA457 and no new risks have been identified. The company highlights that there was no additional exposure to carfilzomib during longer term follow-up, as all people in the CRd group had completed carfilzomib treatment before the cut-off date for the interim analysis. As per the protocol for ASPIRE, people could continue treatment with Rd in both groups. At the time of the primary OS analysis data cut-off, [redacted] ([redacted]) and [redacted] ([redacted]) people in the CRd and Rd groups, respectively, remained on study treatment.

People continued allocated study treatment longer in the CRd arm than the Rd arm, with median treatment duration of [redacted] weeks and [redacted] weeks, respectively (April 2017 cut-off date).

Adverse effects accounted for in the economic model are:

- Neutropenia;
- Anaemia;
- Thrombocytopenia;
- Cataract;
- Hyperglycaemia;
- Lymphopenia;
- Hypertension;
- Fatigue;
- Hypokalaemia;
- Hypophosphataemia;
- Pneumonia.

Serious adverse reactions that could occur during treatment with carfilzomib that are not included in the economic model are cardiac disorders (e.g., congestive cardiac failure, pulmonary oedema,

decreased ejection fraction),[15] which the ERG's clinical experts advised was an important omission (discussed in greater detail in Section 4.2.6). The Summary of Product Characteristics (SmPC) for carfilzomib reports cardiac disorders as a special warning and precaution for use.[15] The SmPC details that cardiac toxicity typically occurs early in the course of treatment but also advises that people are assessed for cardiovascular risk factors before starting treatment with carfilzomib.[15]

Cardiac failure and ischaemic heart disease occurred [redacted] in the CRd group compared with the Rd group (Table 10). [redacted] were also more common with CRd than with Rd (Table 10). For the adverse effects considered in the economic model, with the exception of lymphopenia, a larger proportion of people receiving CRd experienced the event compared with people allocated to Rd (Table 10).

Table 10. Selected adverse events of interest from the ASPIRE safety population based on the April 2017 data cut-off (adapted from CS, Tables 25 [page 64] and 26 [page 66])

Adverse effect	CRd (N = 392) n (%)		Rd (N = 389) n (%)	
Preferred term	All Grades	Grade ≥3	All Grades	Grade ≥3
Cardiac failure	[redacted]	[redacted]	[redacted]	[redacted]
Ischaemic heart disease	[redacted]	[redacted]	[redacted]	[redacted]
Venous thromboembolic events	[redacted]	[redacted]	[redacted]	[redacted]
Peripheral neuropathy	[redacted]	[redacted]	[redacted]	[redacted]
Hypertension	[redacted]	[redacted]	[redacted]	[redacted]
Neutropenia	157 (40.1)	122 (31.1)	136 (35.0)	107 (27.5)
Anaemia	169 (43.1)	73 (18.6)	158 (40.6)	68 (17.5)
Thrombocytopenia	115 (29.3)	66 (16.8)	94 (24.2)	51 (13.1)
Cataract[a]	44 (11.2)	20 (5.1)	37 (9.5)	17 (4.4)
Hyperglycaemia[a]	50 (12.8)	21 (5.4)	39 (10.0)	18 (4.6)
Lymphopenia[a]	13 (3.3)	11 (2.8)	14 (3.6)	8 (2.1)
Fatigue	131 (33.4)	32 (8.2)	124 (31.9)	26 (6.7)
Hypokalaemia	116 (29.6)	41 (10.5)	58 (14.9)	23 (5.9)
Hypophosphataemia[a]	57 (14.5)	35 (8.9)	33 (8.5)	20 (5.1)
Pneumonia	91 (23.2)	63 (16.1)	66 (17.0)	47 (12.1)

[a] Taken from Clinical Study Report.[11]
Abbreviations: C, carfilzomib; CS, company submission; d, dexamethasone; R, lenalidomide.

3.4 Critique of trials identified and included in the indirect comparison and/or multiple treatment comparison

As highlighted in Section 2.3.3, the ERG does not consider daratumumab (D) in combination with bortezomib (V) and dexamethasone (d) at 2L after prior bortezomib to be a valid comparator for the STA reported here. The indirect comparison between CRd and DVd, presented by the company in the CS, is therefore not described or critiqued in this report.

3.5 Conclusions of the clinical effectiveness section

Evidence in support of the clinical effectiveness of CRd in the management of multiple myeloma at 2L is derived from ASPIRE, a randomised controlled trial enrolling adults with multiple myeloma who had received one or more previous lines of therapy. Thus, the population relevant to the decision problem is a subgroup of those taking part in ASPIRE, and, moreover, is not a pre-specified subgroup. As a *post hoc* subgroup, relevant estimates of comparative clinical effectiveness for CRd versus Rd in those receiving treatment at 2L after prior bortezomib are at a higher risk of bias than those reported for the full trial population.

Considering the *post hoc* subgroup that forms the basis of the CS and informs the company's base case in their economic evaluation, data presented by the company are derived from a subgroup in which a proportion of people have not received bortezomib, and others have undergone treatment with lenalidomide, as part of their last treatment regimen (2L prior bortezomib). No lenalidomide-based regimen is recommended by NICE as a treatment option in combination with bortezomib in the first-line setting. Thus, as part of the clarification process, the ERG requested a second *post hoc* subgroup comprising people who had received bortezomib-based regimen as their first treatment and excluding those who had received lenalidomide as part of their first-line regimen (2L prior bortezomib/no lenalidomide). The ERG notes that the two *post hoc* subgroups have similar baseline characteristics for the CRd and Rd treatment groups, and also that comparable hazard ratios (HRs) are derived for PFS and OS for CRd versus Rd.

As would be expected for non-randomised, *post hoc* subgroups, the ERG noted imbalances in some baseline characteristics between those given CRd and those administered Rd for both subpopulations. The direction of bias introduced by the differences in baseline characteristics, and the impact on estimates of relative treatment effect, is unclear. To account for imbalances in baseline characteristics, and to address the limitations associated with use of data derived from a *post hoc* subgroup, the company carried out an inverse probability weighted (IPW) analysis to generate estimates for PFS and OS for CRd versus Rd in the subgroups of people receiving CRd at 2L, for both the company's and ERG's preferred subgroup. In TA457, results from subgroup analyses adjusted to account for imbalances in baseline characteristics arising from non-randomised groups were accepted by the committee. The ERG agrees with the company's approach to mitigate against the issues arising from use of a *post hoc* subgroup. In the CS, the company highlights that for PFS, "*there is a consistent treatment effect across baseline covariate subgroups*". As HRs derived from an

ITT population of an RCT are, by their nature, more robust than those generated from a subgroup analysis, the ERG considers that the results from the ITT population are also relevant to the STA.

In their IPW analysis, the company implemented stepwise Cox regression analysis to select covariates for retention in the model that would subsequently be adjusted using logistic regression to generate IPW estimates. The ERG had reservations around the use of Cox regression to select covariates. On request, the company provided IPW analyses of PFS and OS in which covariates were selected using logistic regression. Similar HRs for PFS and OS are generated from IPW analyses adjusted for covariates selected using the Cox proportional hazards regression and the logistic regression.

Estimates of comparative treatment effectiveness for PFS and OS reported in the CS are based on an additional 3 years of follow up compared with the data presented in TA457. The ERG notes that PFS assessed by the IRC was the primary outcome in ASPIRE, and was met at the time of the interim analysis (data cut-off June 2014). Assessments of PFS after June 2014 are based on determinations by investigators and are not supported by determinations from an IRC. Given that ASPIRE was an open label study, the ERG notes that the determination of progression subsequent to June 2014 is at increased risk of bias.

For the 2L prior bortezomib/no prior lenalidomide subgroup, without adjustment for imbalances in baseline characteristics, CRd was associated with an absolute increase in median PFS of [redacted] months compared with Rd (median PFS [months]: [redacted] [95% CI: [redacted] to [redacted]] with CRd vs [redacted] [95% CI: [redacted] to [redacted]] with Rd). Treatment with CRd was associated with a [redacted] in risk of progression or death (PFS), with the difference between CRd and Rd [redacted] in PFS compared with Rd (HR [redacted]; 95% CI: [redacted] to [redacted]). After IPW adjustment for imbalances in key baseline characteristics (either method for stepwise selection of covariates), the [redacted] associated with treatment with CRd in the unadjusted analysis [redacted]. The difference between CRd and Rd in PFS [redacted], with [redacted] in risk of progression or death from [redacted] to [redacted] or to [redacted] for Cox proportional hazards and logistic regression analyses, respectively.

In the subgroup of those receiving CRd or Rd at 2L after prior bortezomib and no lenalidomide, CRd was associated with an [redacted] in median survival of [redacted] months compared with Rd, and [redacted] in risk of death of [redacted], [redacted], the difference

between treatments in OS [redacted] statistical significance (HR [redacted]; 95% CI: [redacted] to [redacted]). However, IPW analysis generates [redacted] for CRd compared with Rd, with [redacted] in risk of death from [redacted] (covariates selected by stepwise logistic regression), with the difference between groups [redacted] (HR [redacted]; 95% CI: [redacted] to [redacted]). As is expected with OS, results are potentially confounded due to people moving on to non-randomised treatments due to progression of disease. The ERG noted that [redacted] at 3L, with [redacted] available in the NHS for the 3L setting (e.g. [redacted]). The ERG notes that similar estimates of comparative treatment effectiveness for PFS and OS are generated from IPW analyses for the company's and ERG's preferred subgroups.

Although the ERG predominantly considers the company's approach to identification of relevant covariates appropriate, the ERG considers it important to highlight that the regression analyses [redacted] for some individual covariates. The ERG considers that the results could suggest that the characteristics are potential treatment effect modifiers. In particular, adjustment for prior SCT and for β2-microglobulin level suggest that, compared with Rd, treatment with CRd is associated with a [redacted] (Table 7): [redacted] (Table 7). Similar [redacted] was noted in the IPW analyses presented in the CS relating to the subgroup preferred by the company. During clarification, the ERG queried whether the reported effect estimates were for, [redacted] irrespective of treatment received. The company confirmed that the estimates presented were for CRd versus Rd after adjustment for the individual covariate. The company commented that [redacted]. The ERG agrees that [redacted] are likely to have different characteristics [redacted] but considers that there is no clear

clinical rationale

[redacted]. As data are derived from *post hoc* subgroup analyses, the ERG emphasises that any inferences from the results are hypothesis generating.

4 Cost effectiveness

4.1 ERG comment on the company's review of cost effectiveness evidence

The company performed a systematic literature review (SLR) to identify published studies of economic evaluation, health-related quality of life (HRQoL), resource-utilisation, and costs, relating to patients with relapsed or refractory multiple myeloma (R/RMM) who have received at least one prior therapy. The SLR was an update of the company's original appraisal (TA457)[27] and was conducted most recently on 16th March 2018 in anticipation of this part-review of the appraisal. A summary of the ERG's critique of the company's SLR is given in Table 11.

Table 11. Summary of ERG's critique of company's SLR

Systematic review step	Section of CS in which methods are reported			ERG assessment of robustness of methods
	Cost effectiveness evidence	HRQoL evidence	Resource use and costs evidence	
Search strategy	Appendix G	Appendix G	Appendix G	The search was performed around 2 years ago, so it may not include all the latest evidence.
Inclusion/exclusion criteria	Appendix G	Appendix G	Appendix G	Appropriate
Screening	Appendix G	Appendix G	Appendix G	Appropriate
Data extraction	Appendix G	Appendix G	Appendix G	Appropriate
Quality assessment of included studies	Appendix G	Appendix G	Appendix G	Appropriate

Abbreviations: CS, company submission; ERG, evidence review group; HRQoL, health related quality of life.

The company's SLR resulted in the inclusion of 43 economic evaluations, 15 cost/resource use studies and 22 HRQoL studies.

The company's SLR was generally sound but was performed around 2 years ago and therefore may have missed relevant studies published since then. Despite this, the ERG considers the sources used by the company throughout the analysis to be generally reasonable and unlikely to be limited by the restriction of the SLR date.

4.2 Summary and critique of company's submitted economic evaluation by the ERG

4.2.1 NICE reference case checklist

Table 12 summarises the ERG's appraisal of the company's economic evaluation against the requirements set out in the NICE reference case checklist for the base-case analysis, with reference to the NICE final scope outlined in Section 2.

Table 12. NICE reference case checklist

Element of health technology assessment	Reference case	ERG comment on company's submission
Perspective on outcomes	All direct health effects, whether for patients or, when relevant, carers	All relevant health effects for adult patients with multiple myeloma who have received only one prior therapy with bortezomib have been included.
Perspective on costs	NHS and PSS	All relevant costs have been included and are based on the NHS perspective.
Type of economic evaluation	Cost–utility analysis with fully incremental analysis	Cost-utility analysis with fully incremental analysis has been provided by the company.
Time horizon	Long enough to reflect all important differences in costs or outcomes between the technologies being compared	Lifetime horizon (40 years).
Synthesis of evidence on health effects	Based on systematic review	The company performed an appropriate systematic review.
Measuring and valuing health effects	Health effects should be expressed in QALYs. The EQ-5D is the preferred measure of health-related quality of life in adults.	QALYs using data from the EORTC QLQ-C30 and the myeloma-specific EORTC QLQ-MY20 taken from ASPIRE and mapped to the EQ-5D-3L.
Source of data for measurement of health-related quality of life	Reported directly by patients and/or carers	EORTC QLQ-C30 and the myeloma-specific EORTC QLQ-MY20 reported directly from the subgroup of interest in ASPIRE, mapped to obtained EQ-5D-3L utility values.
Source of preference data for valuation of changes in health-related quality of life	Representative sample of the UK population	The subgroup of interest from ASPIRE is representative of the UK population.
Equity considerations	An additional QALY has the same weight regardless of the other characteristics of the individuals receiving the health benefit	The economic evaluation matches the reference case.

Evidence on resource use and costs	Costs should relate to NHS and PSS resources and should be valued using the prices relevant to the NHS and PSS	Costs included in the analysis have been sourced using NHS reference costs[28], MIMS[29], eMIT[30] and published literature and are reported in pounds sterling for the price year 2018.
Discounting	The same annual rate for both costs and health effects (currently 3.5%)	Discount rate of 3.5% has been used for both costs and health effects.

Abbreviations: eMIT, Drug and pharmaceutical electronic market information tool; EORTC QLQ-C30, European Organization for the Research and Treatment of Cancer Quality of Life Questionnaire C30; EQ-5D-3L, EuroQoL five dimensions three levels; ERG, evidence review group; MIMS, monthly index of medical specialities; NHS, national health service; PSS, personal social services; QALY, quality adjusted life year

4.2.2 Population

The population considered by the company for this single technology appraisal (STA) is adult patients with multiple myeloma who have received only **one** prior therapy with bortezomib (hereafter referred to as the 2L prior bortezomib subgroup). The population under consideration is a restricted sub-population of the marketing authorisation (MA) for the triplet therapy, carfilzomib in combination with lenalidomide and dexamethasone (hereafter referred to as CRd), which does not restrict prior therapy to bortezomib.

The restricted population proposed by the company is a deviation from the NICE final scope, which proposes the relevant population to be adult patients with multiple myeloma who have received **at least** one prior therapy. The company justify the positioning of CRd for the 2L prior bortezomib subgroup as they state it reflects the need for triplet therapies earlier in the pathway, greater benefit of the treatment is demonstrated in this subgroup and thus offers the greatest economic value, and lastly, it aligns with the NICE recommendation for Rd, which is deemed the most relevant comparator. The ERG's clinical experts agreed that there is an unmet need in the second-line setting. Furthermore, the ERG considers that not exploring subgroups where cost-effectiveness cannot be demonstrated is appropriate and pragmatic.

However, as discussed in Section 2.3.1, the ERG noted that the company's subgroup included a proportion of people who had not received bortezomib as part of their last round of therapy ([redacted]), as well as people who had undergone treatment with lenalidomide in their last regimen ([redacted]). Given that the company is positioning CRd at the 2L setting and after prior bortezomib, as part of the clarification process, the ERG requested that the company generate a new subgroup comprising people who had undergone only one round of therapy that was bortezomib-based and who had not received prior lenalidomide, and to provide revised estimates of

PFS and OS for the new subgroup (hereafter referred to as 2L prior bortezomib/no prior lenalidomide) and incorporate these into the cost-effectiveness analysis.

The company provided revised cost-effectiveness results for the 2L prior bortezomib/no prior lenalidomide subgroup as a scenario only and maintained that the 2L prior bortezomib subgroup analysis is more appropriate because patients could receive lenalidomide as first-line treatment in combination with bortezomib. Thus, the company did not change its base-case assumptions. However, in England, bortezomib plus lenalidomide is not an approved therapy at first-line. As such, the ERG considers the 2L prior bortezomib/no prior lenalidomide subgroup to be more reflective of patients who would be eligible for CRd and Rd in England and has implemented this subgroup for the ERG base-case analysis (Section 6.4).

4.2.3 Interventions and comparators

The intervention under consideration for the economic analysis is CRd, which is a triplet therapy consisting of carfilzomib in combination with lenalidomide and dexamethasone.

The comparators considered by the company are Rd, which is doublet combination therapy of lenalidomide and dexamethasone, and daratumumab plus bortezomib and dexamethasone (DVd). However, DVd is only available through the Cancer Drugs Fund and is not approved for use in routine commissioning. Therefore, NICE has advised that DVd is not a relevant comparator for this analysis and is not discussed further in this ERG report.

The company's choice to limit the comparator to Rd, based on the restriction of the population to the 2L prior bortezomib subgroup, only partially reflects the NICE final scope. The NICE final scope splits the population by line of therapy and outlines comparators for each subgroup. For the 2L subgroup, the comparators of interest in the NICE final scope are Rd, carfilzomib in combination with dexamethasone (Cd) and bortezomib monotherapy. However, the company's restriction to one prior therapy with bortezomib removes Cd and bortezomib monotherapy as comparators and ignores the third- and fourth-line subgroups. As mentioned in Section 4.2.2, the ERG considers the company's justification to restrict to the 2L prior bortezomib subgroup reasonable and as such considers that Rd is the most relevant comparator at second-line of therapy.

The dosing regimen for the individual components of CRd and Rd (carfilzomib, lenalidomide and dexamethasone) is presented in Table 13.

Table 13. Treatment dosing regimen

Treatment	Dose	Dose regimen	Treatment duration
Carfilzomib	Starting dose of 20mg/m^2 on days 1 and 2 of cycle 1 (maximum dose of 44mg). Target dose of 27mg/ m^2 thereafter (maximum dose of 60mg).	**Cycles 1-12**: 10-minute IV infusion on days 1,2,8,9,15 and 16 of a 28-day treatment cycle. **Cycles 13-18**: 10-minute IV infusion on days 1,2,15 and 16 of a 28-day cycle.	Up to 18 cycles
Lenalidomide	25mg per dose	One tablet, taken orally on days 1-21 of a 28-day treatment cycle.	Treatment until progression of disease or unacceptable toxicity.
Dexamethasone	40mg per dose	20 tablets, taken orally on days 1, 8, 15 and 22 of a 28-day treatment cycle.	Treatment until progression of disease or unacceptable toxicity.

Abbreviations: IV, intravenous; m^2, metre-squared; mg, milligram.

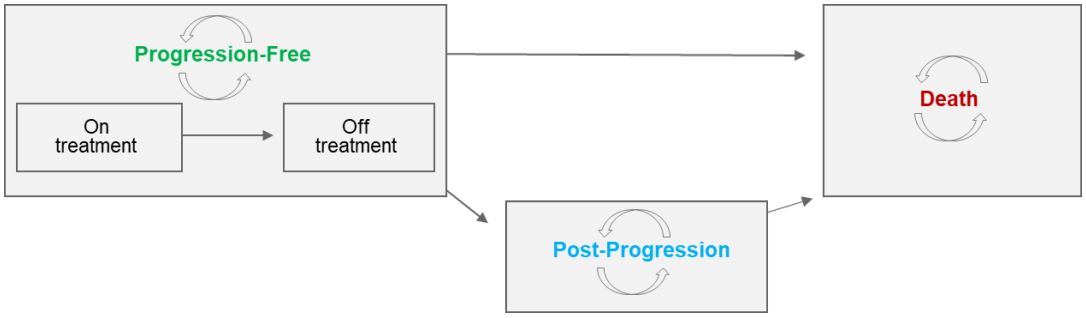
Time to treatment discontinuation (TTD) for carfilzomib, lenalidomide and dexamethasone in the CRd and Rd arms is based on data from the ASPIRE trial, extrapolated over a lifetime horizon using standard parametric survival distributions (described further in Section 4.2.5). It should be noted that in ASPIRE, treatment for carfilzomib was capped to 18 cycles and this is reflected in the economic analysis. For lenalidomide and dexamethasone, discontinuation of treatment was primarily due to disease progression or because of unacceptable toxicity.

4.2.4 Modelling approach and model structure

A single *de novo* economic model was developed in Microsoft© Excel to assess the cost-effectiveness of CRd compared with Rd for the treatment of adult patients with multiple myeloma who have received at least one prior therapy with bortezomib (2L prior bortezomib subgroup).

The model structure is based on a partitioned survival analysis structure, with three health states: progression-free, progressed and dead. The progression-free health state is further sub-divided into progression-free and on-treatment and progression-free off-treatment. Figure 11 presents the company's model schematic. The company state that the chosen model structure is in line with previous HTA oncology models, specifically in the area of multiple myeloma.[4, 7, 31-33]

Figure 11. Model structure (adapted from the schematic presented in the company's economic model)



All patients enter the model in the progression-free health state and are assumed to start treatment on CRd or Rd. During each model cycle, patients in the progression-free health state can be either on-treatment or off-treatment if they are experiencing unacceptable toxicity. Furthermore, from the progression-free health state, patients can transition to either the progressed health state when they experience disease progression or die (thus transitioning to the dead health state). When patients transition to the progressed health state, they remain there until death.

Extrapolations of clinical outcomes data, including progression-free survival (PFS), overall survival (OS) and TTD, using standard parametric curves are implemented in the model to estimate the proportion of patients occupying a health state in any given model cycle. PFS is used to estimate the proportion of patients occupying the progression-free health state, OS is used to model the death state and TTD is used to estimate the proportion of patients who are progression-free and on-treatment. The proportion of patients occupying the progressed health state for any given cycle is calculated as the difference between OS and PFS per cycle. A detailed description of how the survival curves were estimated and implemented in the model is provided in Section 4.2.5.

A model cycle length of 28-days with half-cycle correction applied was implemented in the model and is reflective of a treatment cycle length for carfilzomib. The model time horizon was set to 40 years, considered by the company to be sufficiently long enough to capture a lifetime as the median age in ASPIRE at baseline was 64 years. The perspective of the analysis was based on the UK national health service (NHS), with costs and benefits discounted using a rate of 3.5%, as per the NICE reference case.[34]

4.2.4.1 ERG Critique

The ERG considers the structure of the company's model is appropriate, capturing all relevant health states and clinically plausible transitions between health states that are largely similar to other appraised oncology models. The 28-day cycle length used in the model is suitable to capture important changes in the health state of patients, allowing for robust estimates of costs and benefits to be calculated for each treatment. Half-cycle correction has been appropriately applied in the model to prevent over or under-estimation of costs and quality-adjusted life-years (QALYs).

4.2.5 Treatment effectiveness

Overview of the company's approach to survival analysis

Treatment effectiveness estimates in the economic model for CRd and Rd are calculated using extrapolations of ASPIRE inverse probability weighted (IPW) Kaplan Meier (KM) PFS and OS data for the 2L prior bortezomib subgroup (company base-case). At the request of the ERG, the company also provided a scenario where alternative treatment effectiveness estimates for CRd and Rd are based on extrapolations of IPW KM PFS and OS data for the 2L prior bortezomib/no prior lenalidomide subgroup, which is discussed further in Section 4.2.5.1. The data cut-off point for all analyses was 5 December 2017.

For the company's base-case analysis, OS estimates used in the economic model incorporate extrapolated real-world data from the MyelomaToul registry. Time-on-treatment estimates in the model for carfilzomib, lenalidomide and dexamethasone for each treatment arm are based on extrapolations of TTD KM data for the 2L prior bortezomib subgroup from ASPIRE.

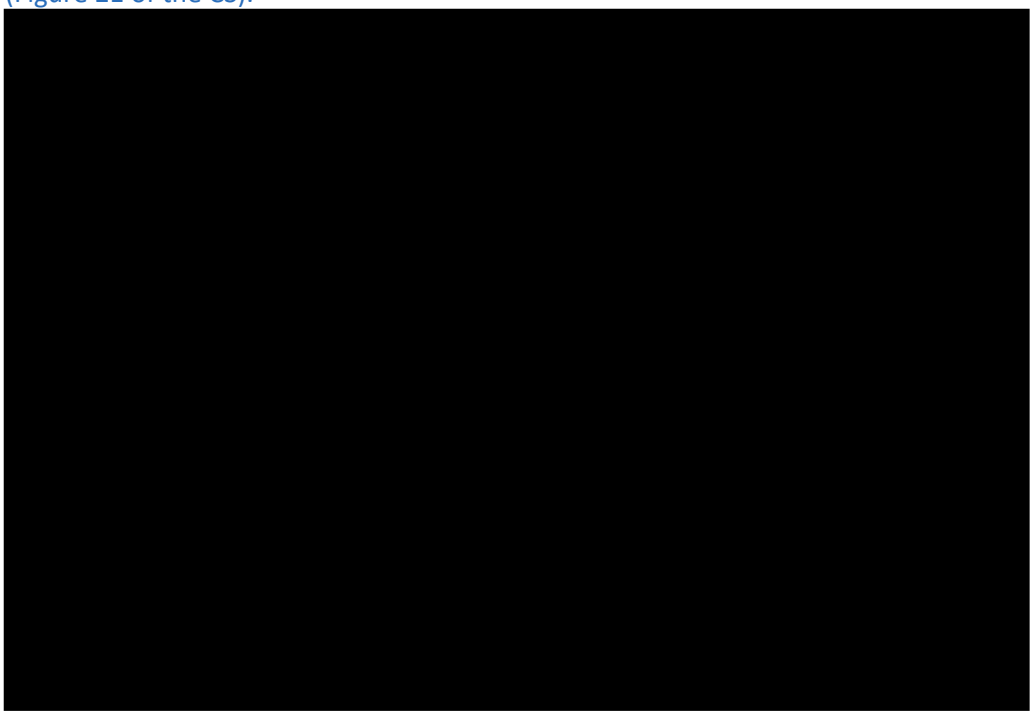
The company first assessed whether the assumption of proportional hazards (PH) held for PFS and OS outcomes from ASPIRE for both the intention-to-treat (ITT) population, as well as the 2L prior bortezomib subgroup using log-cumulative hazard plots. The company used the outcomes of the PH assessment to decide to either jointly or separately fit survival distributions. Extrapolations of the KM data were then performed using standard parametric survival distributions (exponential, Weibull, Gompertz, log-normal, log-logistic and generalised gamma). For the extrapolation of the MyelomaToul registry data, the company also assessed piecewise exponential models with different time-point cut-offs.

The process of curve selection recommended in the NICE decision support unit technical support document (DSU TSD) 14 was implemented by the company to select an appropriate distribution for the extrapolation of each outcome.[20] The company assessed the fit of each modelled curve against the KM data using statistical goodness of fit statistics, including Akaike information criterion (AIC) and Bayesian information criterion (BIC) statistics, visual inspection of the curves and clinical plausibility of the extrapolation over the time horizon of the model.

Progression-free survival

Based on AIC/BIC statistics and visual fit, the company selected the generalised gamma distribution for the PFS extrapolation, presented in Figure 12. Plots of all the assessed distributions compared with the KM data and AIC/BIC statistics can be found in Appendix M.5 of the CS.

Figure 12. PFS curves used for the company's base-case analysis; 2L prior bortezomib subgroup (Figure 21 of the CS).



Abbreviations: CS, company submission; CRd, Carfilzomib plus lenalidomide and dexamethasone; KM, Kaplan Meier; PFS, progression-free survival; Rd, lenalidomide plus dexamethasone.

A comparison of the IPW KM PFS curves in Figure 12 shows that the curves cross at around 66 months after randomisation (please refer to Figure 32, Appendix M of the CS for a more detailed presentation of the KM curves). The company state this is due to the small numbers at risk towards the end of the data cut-off (5 December 2017), such that one event causes a substantial change in the KM curve. The crossing of the curves was deemed clinically implausible by the company and its

clinical experts as the treatment effect for OS remained consistent throughout follow-up. Furthermore, the KM PFS curve for the ITT population (Figure 26, Appendix M of the CS) demonstrated a consistent separation of the curves for CRd and Rd and the log-cumulative hazard plots demonstrated that the PH assumption holds (Figure 40, Appendix M of the CS).

Thus, the company chose to model PFS jointly instead of separately as they state that the ITT data for PFS are more informative to decide on the approach to modelling the 2L prior bortezomib subgroup data as they are based on more patients and demonstrate consistency of the treatment effect for CRd.

Overall survival

The company explored whether the assumption of PH held for the IPW KM OS data for the 2L prior bortezomib subgroup to determine the choice of jointly or separately modelling the parametric survival curves. Based on the log-cumulative hazards plots, presented in Figure 19 of Appendix M 3.2 of the CS, the company concluded that the PH assumption held and jointly modelled the OS curves for CRd and Rd.

The company explored the statistical and visual fit of standard parametric distributions to the IPW KM OS data, as well as the clinical plausibility of the extrapolations. The company selected the Weibull distribution as the best fit but found that the estimates of survival produced for the Rd curve towards the end of the extrapolation (0% at 20 years) were conservative when compared with survival estimates of 11% at 25 years, presented for the technology assessment of Rd (TA586), though the ERG for TA586 found the estimates implausible.[7]

As such, the company used real-world data from a French registry of multiple myeloma patients, MyelomaToul, to inform the extrapolations of OS for both the Rd and the CRd arms of the model.[35] The company digitised published data from the registry of patients treated with second-line lenalidomide (n=1,890) and explored three piecewise exponential models with cut off points of 48, 60 and 72 months to extrapolate the data. The company stated that all three models visually fit the data well and chose the piecewise exponential model with the cut-off point of 72 months as it had the best statistical fit (lowest AIC value). The exponential model pieces can be defined as period one, which is months 0 to 72 and period two, which is 72 months onwards.

In their clarification response, the company confirmed that standard parametric curves (exponential, Weibull, Gompertz, log-logistic, lognormal and generalised gamma) were also explored but found that the piecewise exponential model (no cut-off stated) had the best statistical fit.

Based on the piecewise exponential model for the MyelomaToul data, a survival probability for one cycle was estimated separately for period one and period two. The company then calculated an adjusted MyelomaToul survival curve (referred to as the matched MyelomaToul curve in the CS) to account for the difference in the mortality rate between the registry data and the IPW subgroup data from ASPIRE. The adjusted curve was calculated by applying time-dependent hazard ratios (HR) to the survival probability for one cycle for period one and then period two.

The company calculated the time-dependent HRs by fitting a Cox model to the MyelomaToul KM data and the IPW KM data for the 2L prior bortezomib subgroup. The company selected time-dependent HRs to adjust the MyelomaToul registry data because up to month 10 the two data sets overlapped and thereafter separated out (Figure 13). As such, the HR for the period 0-10 months was 1.01 and for 11 months onwards, the HR was 2.04. The company explored the use of a constant HR in a scenario, presented in Section 5.2.

Figure 13. Comparison of Kaplan-Meier overall survival data for MyelomaToul and ASPIRE (2L prior bortezomib subgroup) (Figure 17 of the CS)

Abbreviations: 2L, second-line; CS, company submission; CRd; pB, prior bortezomib; Rd, lenalidomide plus dexamethasone.

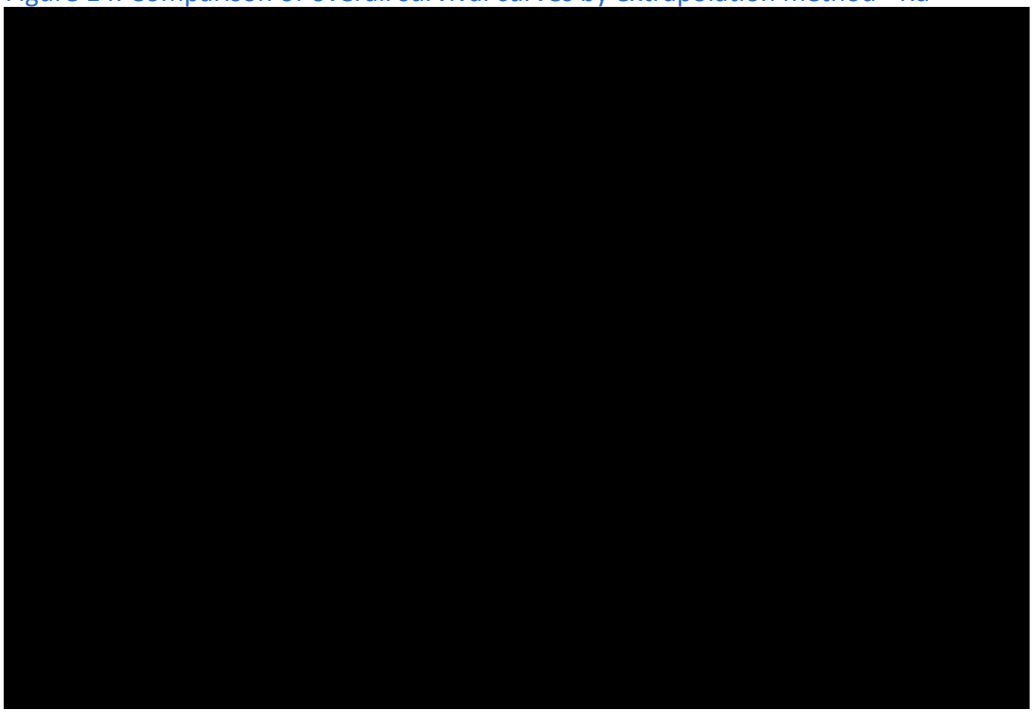
The company's base-case OS curve for Rd is a hybrid of the Weibull survival curve, based on IPW KM subgroup data, truncated to month 72 and then from month 72 onwards, the hazards from the adjusted MyelomaToul piecewise exponential survival curve are applied to the survival proportion estimated in the previous cycle. To estimate the CRd OS curve, first the company applied the IPW OS HR for the subgroup derived from ASPIRE ([redacted]) to the adjusted MyelomaToul Rd OS curve to calculate the per-cycle hazards. Then, for the first 72 months, the ASPIRE Weibull OS curve for CRd was used and thereafter the hazards from the first step were applied to survival proportion estimated in the previous cycle to construct the remaining portion of the OS curve. Table 14, Figure 14 and Figure 15 present comparisons of OS predictions for the Weibull extrapolation of ASPIRE IPW subgroup data and the hybrid method using MyelomaToul data.

Table 14. Comparison of overall survival predictions by extrapolation method

OS assumptions	10 years		20 years	
	CRd	Rd	CRd	Rd
ASPIRE Weibull distribution	16%	5%	2%	0%
Adjusted MyelomaToul model + HR	21%	9%	9%	1%

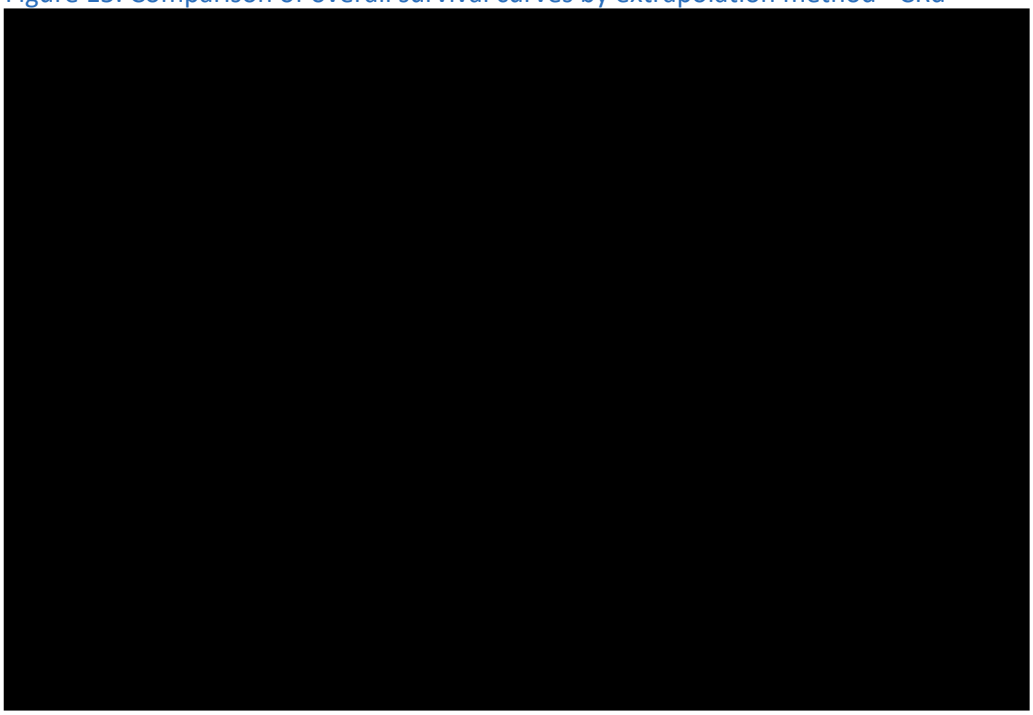
Abbreviations: CRd, carfilzomib plus lenalidomide and dexamethasone; HR, hazard ratio; ITT, intention-to-treat; OS, overall survival; Rd, lenalidomide plus dexamethasone

Figure 14. Comparison of overall survival curves by extrapolation method - Rd



Abbreviations: KM, Kaplan Meier; Rd, lenalidomide plus dexamethasone.

Figure 15. Comparison of overall survival curves by extrapolation method - CRd



Abbreviations: CRd, carfilzomib plus lenalidomide and dexamethasone; KM, Kaplan Meier.

Time-to-treatment discontinuation

Time-to-treatment discontinuation for each treatment in the CRd and Rd arms was modelled separately as the company states that patients may discontinue different components of combination therapy at different times. As such, IPW TTD KM subgroup data from ASPIRE for carfilzomib, lenalidomide and dexamethasone for CRd and Rd were used to inform the survival extrapolations. It should be noted that in ASPIRE, carfilzomib treatment was capped at 18 treatment cycles. Thus, the company truncated the carfilzomib survival extrapolation to 18 model cycles. In addition, the company states that as PFS is longer for patients on CRd compared with Rd, and treatment duration with lenalidomide and dexamethasone is also longer, lenalidomide and dexamethasone have been modelled separately for CRd and Rd. In the model, TTD is capped to PFS to ensure patients are not accruing treatment costs if they have disease progression.

Table 15 presents the parametric survival distributions selected by the company for use in the base case analysis, based on AIC/BIC statistics and visual fit of the curve to the KM data. Plots of all the assessed distributions compared with the KM data and AIC/BIC statistics can be found in Appendix M.6 of the CS.

Table 15. Selected TTD survival distributions for components of CRd and Rd (Table 30 of the CS)

Treatment component	CRd	Rd
Carfilzomib	Gompertz	-
Lenalidomide	Exponential	Log-logistic
Dexamethasone	Exponential	Log-logistic

Abbreviations: CRd, Carfilzomib plus lenalidomide and dexamethasone; CS, company submission; Rd, lenalidomide plus dexamethasone; TTD, time to treatment discontinuation.

4.2.5.1 ERG critique

The company's base-case cost-effectiveness analysis is based on the IPW 2L prior bortezomib subgroup from ASPIRE. As mentioned in Sections 2.3 and 4.2.2, this subgroup includes patients who have received prior lenalidomide, which the ERG considers does not reflect UK clinical practice. Thus, the ERG's critique of treatment effectiveness is based on the IPW 2L prior bortezomib/no prior lenalidomide subgroup data from ASPIRE and the analysis provided by the company in their response to ERG clarification questions. It should be noted that the methods of analysing the data remain the same as in the company base-case and it is only the underlying data sources and extrapolations that have been updated (presented in Figures 8, 9 and 10 of the company's response to ERG clarification questions). Table 16 presents a comparison of the treatment effectiveness parameters used for the company's base-case and the company's scenario for the 2L prior bortezomib/no lenalidomide subgroup.

Table 16. Comparison of treatment effectiveness parameters for company base-case vs company scenario for the 2L prior bortezomib/no lenalidomide subgroup

Model parameter	Company base case (2L prior bortezomib)		Company scenario (2L prior bortezomib/no prior lenalidomide)	
	CRd	Rd	CRd	Rd
PFS	Joint fitted generalised gamma		Joint fitted generalised gamma	
OS	Joint fitted Weibull (first 72months), then MyelomaToul Rd with ASPIRE IPW OS HR applied for the remainder of the model	Joint fitted Weibull (first 72 months) + matched MyelomaToul (piecewise exponential, cut-off at 72 months)	Joint fitted Weibull (first 72 months), then MyelomaToul Rd with ASPIRE IPW OS HR applied for the remainder of the model	Joint fitted Weibull (first 72 months) + matched MyelomaToul (piecewise exponential, cut-off at 72 months)
TTD - carfilzomib	Gompertz	-	Exponential	-
TTD – lenalidomide	Exponential	Log-logistic	Exponential	Log-logistic
TTD - dexamethasone	Exponential	Log-logistic	Exponential	Log-logistic

IPW OS HR (CRd vs Rd)	[redacted]	-	[redacted]	-
Time dependent HRs for MyelomaToul adjustment	-	1.02 before 10 months, 2.04 thereafter	-	[redacted]

Abbreviations: 2L, second-line; CRd, Carfilzomib plus lenalidomide and dexamethasone; HR, hazard ratio; IPW, inverse probability weighted; OS, overall survival; PFS, progression-free survival; Rd, lenalidomide plus dexamethasone; TTD, time to treatment discontinuation.

As mentioned in Section 3.4, the ERG considers the company's IPW analysis to adjust the ASPIRE ITT data for the 2L prior bortezomib/no prior lenalidomide subgroup is reasonable. The ERG's main concern with the modelling of treatment effectiveness is how the company has estimated OS, using real-world data to adjust mature trial-based data. The company extrapolated ASPIRE trial data but found the estimates for key time points (10 and 20 years) for the Rd arm did not pass clinical validity based on a comparison of estimates produced by previous TAs (TA586, TA573, TA457)[7, 27, 36]. Namely, the company deemed the estimates produced by the best-fitting Weibull curve for Rd to be pessimistic, predicting survival at 10 and 20 years to be 5% and 0%, respectively. Using data from the MyelomaToul registry for the 2L lenalidomide population, the company estimated what they believed to be more clinically plausible survival estimates of 9% and 1% for 10 and 20 years (Table 14).

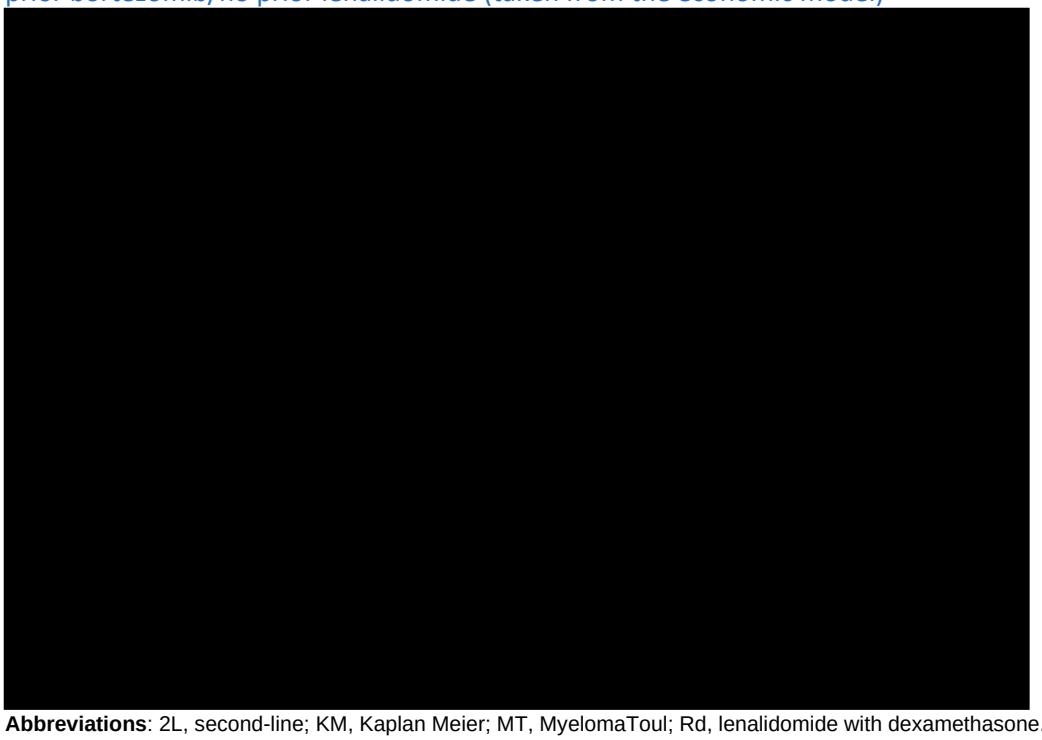
The company also provided results from a multistate model based on ASPIRE ITT data (Appendix N of the CS) which estimated survival at 20 years to be between 1.9% and 3%. Even though the estimates are not based on the subgroup of interest, the company state the multi-state model results are generalisable to the subgroup of interest. The ERG highlights that the model was not submitted to the ERG and estimates of survival for CRd were not provided for comparison. However, the ERG considers that it was not necessary to investigate the model further, as mature trial data from ASPIRE for the subgroup of interest are available and as mentioned previously, the three-state model is appropriate.

In the statistical analysis report for MyelomaToul, produced specifically for Amgen, the ERG found that the subgroup data used in the modelling is not adjusted for "only one prior therapy that was a bortezomib-based regimen", which would match the company base-case population.[35] In the report, of patients not undergoing SCT (referred to as a graft; second plot of Figure 15), approximately 50% of patients received a bortezomib-based regimen as their first-line treatment and nearly 30% of patients received therapy classed as "other". The ERG assumes that the analyses were bespoke for Amgen and thus considers that the company could have requested a subgroup analysis of OS for patients who only received a bortezomib-based treatment as their first-line therapy and then went

on to 2L lenalidomide. The mixture of first-line treatments may be an influential factor on OS for 2L lenalidomide patients from MyelomaToul as OS in this group is longer than that of 2L Rd patients who had one prior treatment with bortezomib and no lenalidomide from ASPIRE (Figure 16).

The ERG acknowledges that the company has adjusted the data to account for the mortality difference between the two datasets, but because the shape of the tail of the adjusted MyelomaToul curve is different to the ASPIRE KM curve (Figure 16), the adjustment influences the extrapolation of the data and results in an increase in survival in the tail compared with Weibull extrapolation of the ASPIRE data (see Figure 14). Furthermore, as can be seen in Figure 16, the company's adjustment of the MyelomaToul data shows that only 1% are alive from year 12 onwards, rather than the 20 years reported by the company.

Figure 16. Comparison of OS KM for 2L lenalidomide treated patients in MyelomaToul and ASPIRE 2L prior bortezomib/no prior lenalidomide (taken from the economic model)



Abbreviations: 2L, second-line; KM, Kaplan Meier; MT, MyelomaToul; Rd, lenalidomide with dexamethasone.

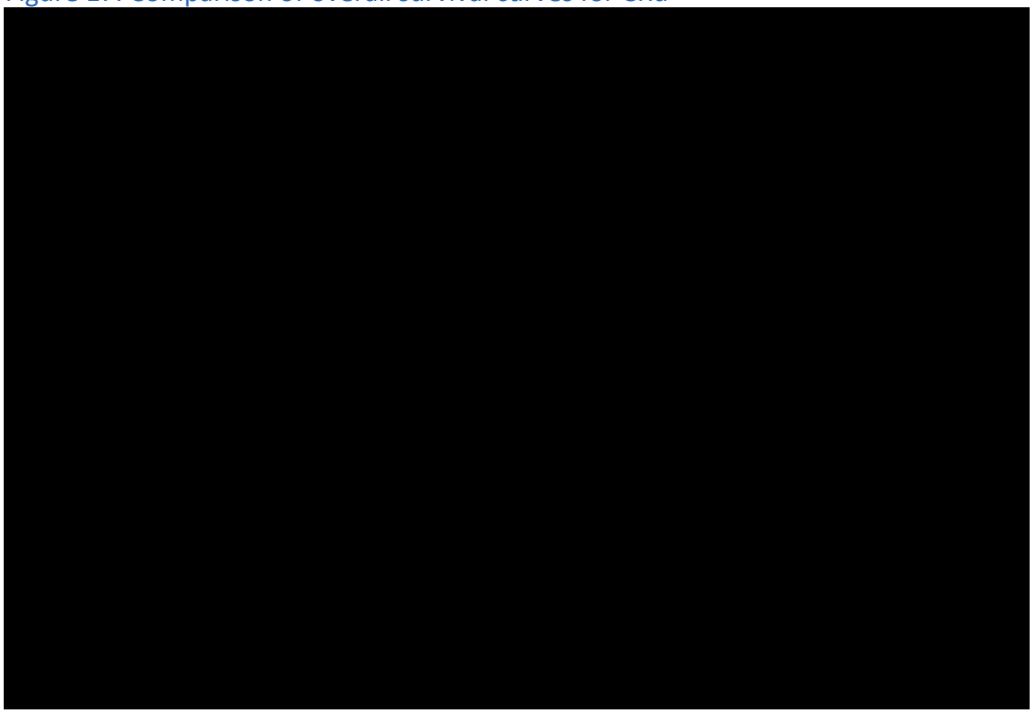
The ERG consulted its clinical experts who confirmed that longer-term survival estimates for Rd patients based on ASPIRE are conservative. However, the consequence of the company's adjustment when using real-world evidence is that survival is inflated for CRd compared with the estimates based on IPW ASPIRE data presented in Figure 15, improving its cost-effectiveness.

Where trial data are available DSU TSD 14 recommends selecting a different extrapolation based on trial data that produces more clinically valid estimates of survival.[20] In their response to the ERG's

clarification question B7, the company states that when using the MyelomaToul extrapolation to validate the ASPIRE extrapolation, the exponential distribution provides the most plausible long-term predictions of survival, with results comparable between the models (Table 17). However, the company considered their base-case approach more appropriate as the model had a better statistical fit to the observed data. Though in their main submission, the company states that all models for the ASPIRE OS data performed similarly well in terms of statistical fit. Moreover, for the company base-case subgroup (2L prior bortezomib) the exponential model was the second-best fitting distribution. For the ERG preferred subgroup (2L prior bortezomib/ no prior lenalidomide), the exponential model was statistically the best fit to the KM data.

Figure 17 presents a comparison of the different modelling approaches for CRd OS and it can be seen that the exponential distribution is less pessimistic than the Weibull distribution but is also less optimistic than the company's base-case approach using MyelomaToul data.

Figure 17. Comparison of overall survival curves for CRd



Abbreviations: CRd, carfilzomib plus lenalidomide and dexamethasone, KM, Kaplan Meier.

The ERG considers that IPW OS data from ASPIRE should be used for the base-case analysis as it is now mature, which was a considerable limitation in TA457[27] and thus a clinically plausible extrapolation of OS for CRd can be estimated entirely from trial data. Furthermore, data from ASPIRE are based on the subgroup of interest, the patient characteristics have been adjusted for to limit bias and it maintains the observed treatment effect between the two trial arms, increasing the

robustness of the cost-effectiveness analysis. Therefore, the ERG deems it appropriate to revert to the company's survival modelling of the ASPIRE subgroup data wholly for CRd and Rd and considers the exponential model to be appropriate to model OS and explores this in a scenario presented in Table 17 and Section 6.3.

The ERG highlights an additional issue regarding the treatment effect for subgroups. In the CS the company states that "*based on stepwise Cox regression modelling, there was a lack of evidence of treatment-covariate interactions for PFS suggesting an overall consistent treatment effect across the baseline covariate subgroups"*. This statement infers that while the absolute benefit may be different based on the particular subgroup the relative benefit between the two treatment groups is consistent irrespective of subgroup. As such, the ERG considers that the HRs for PFS (HR 0.66) and OS (HR 0.794) derived from the ITT population are relevant for consideration and requested the company to provide a scenario applying the ITT HRs to the baseline Rd PFS and OS extrapolations to construct alternative CRd PFS and OS curves.

The company did not supply the requested scenario with their clarification response and instead provided what the ERG considers to be a circular argument for why it is inappropriate to use the ITT HRs. The company state that, "*ASPIRE was not primarily designed to detect significant treatment effects within subgroups, and consequently lacked power to detect significant treatment-covariates interactions. In addition, we also noted that there may be important differences in baseline characteristics across study arms in subgroups (particularly if the subgroup is constructed by multiple covariates such as for the current assessment) that confound the subgroup-specific treatment effect estimates"*.

From the company's statement, the ERG understands that the HRs for the base-case subgroup (2L prior bortezomib) and the ERG preferred subgroup (2L prior bortezomib/no prior lenalidomide) are likely to be confounded whereas the results from the randomised ITT population are not. Therefore, the ERG considers it is still relevant to explore the impact of applying the ITT HRs to construct alternative CRd PFS and OS curves and conducted the following two scenario analyses for the ERG preferred subgroup (2L prior bortezomib/no prior lenalidomide):

1. Applying the ITT PFS and OS HRs to the company's preferred PFS and OS survival curves for Rd.

2. Applying the ITT PFS HR to the company's preferred Rd PFS curve and the ITT OS HR to the ERG's preferred modelling of Rd OS using the exponential distribution to extrapolate ASPIRE IPW data.

Results of these scenarios can be found in Section 6.3 and a comparison of the OS predictions by extrapolation method are presented in Table 17.

Table 17. Comparison of overall survival predictions by extrapolation method for the 2L prior bortezomib/ no prior lenalidomide subgroup

OS assumptions	10 years		20 years	
	CRd	Rd	CRd	Rd
ASPIRE Weibull distribution	16%	5%	2%	0%
ASPIRE exponential distribution (ERG preferred)	**19%**	**8%**	**4%**	**1%**
Adjusted MyelomaToul model + HR (company base case)	**21%**	**9%**	**6%**	**1%**
ITT PFS and OS HRs applied to company scenario for PFS and OS for Rd	15%	9%	3%	1%
ITT PFS HR applied to company scenario Rd PFS curve and ITT OS HR applied to ERG alternative OS modelling for Rd using the exponential distribution for ASPIRE data only.	13%	8%	2%	1%

Abbreviations: CRd, carfilzomib plus lenalidomide and dexamethasone; ERG, evidence review group; HR, hazard ratio; ITT, intention-to-treat; OS, overall survival; PFS, progression-free survival; Rd, lenalidomide plus dexamethasone

Aside from the issues with OS, the ERG considers the modelling of PFS using the jointly fitted generalised gamma distribution is appropriate. Furthermore, the use of the log-logistic distribution to extrapolate TTD for lenalidomide and dexamethasone for the Rd arm of the model is reasonable.

The ERG had concerns with the modelling of TTD for CRd treatment components. Specifically, the Weibull distribution for the TTD modelling of carfilzomib, lenalidomide and dexamethasone for the CRd arm provided a better fit to the observed KM data then the company's base-case choice of the exponential distribution. However, a scenario using the ERG's preferred survival curves for the extrapolation of CRd TTD had minimal impact on the ICER. Results of the scenario can be found in Section 6.2.

4.2.6 Adverse events

For the base case analysis, the company included grade 3 or higher treatment-emergent adverse events (TEAEs) that were reported by at least 5% of patients in the safety population in either treatment arm of ASPIRE, presented in Table 18.

Table 18. Grade 3 or higher AEs implemented in the model (Table 47 of the CS)

Adverse events	CRd (%)	Rd (%)
Neutropenia	31.12	27.51
Anaemia	18.62	17.48
Thrombocytopenia	16.84	13.11
Cataract	5.10	4.37
Hyperglycaemia	5.36	4.63
Lymphopenia	2.81	2.06
Hypokalaemia	10.46	5.91
Fatigue	8.16	6.68
Hypertension	5.36	2.31
Hypophosphataemia	8.93	5.14
Pneumonia	16.07	12.08

Abbreviations: CRd, Carfilzomib plus lenalidomide and dexamethasone; Rd, lenalidomide plus dexamethasone

The company then estimated per-cycle probabilities of experiencing each adverse event using the following formula:

Per-cycle probability of AE = 1-EXP(((LN(1-incidence of AE))/mean number of treatment cycles in ASPIRE))

Table 19 presents the per-cycle AEs for each treatment arm included in the model.

Table 19. Probability of AEs per cycle implemented in the model (Table 47 of the CS)

Adverse events	CRd (%)	Rd (%)
Neutropenia	1.24%	1.39%
Anaemia	0.69%	0.83%
Thrombocytopenia	0.61%	0.61%
Cataract	0.17%	0.19%
Hyperglycaemia	0.18%	0.21%
Lymphopenia	0.09%	0.09%
Hypokalaemia	0.37%	0.26%
Fatigue	0.28%	0.30%
Hypertension	0.18%	0.10%
Hypophosphataemia	0.31%	0.23%
Pneumonia	0.58%	0.56%

Abbreviations: CRd, Carfilzomib plus lenalidomide and dexamethasone; Rd, lenalidomide plus dexamethasone

The impact of AEs on patients' quality of life is considered in the model and is described further in Section 4.2.7, while the costs of managing AEs are discussed in Section 4.2.8.

4.2.6.1 ERG critique

After consultation with the ERG's clinical experts, cardiac failure was found to be an omission from the model. In ASPIRE, [redacted] of CRd patients and [redacted] of Rd patients experienced grade 3 or higher cardiac failure. Furthermore, the company presented grade 3 or higher adverse events of interest (which included cardiac failure) that were also not included in the analysis (Table 26 of the CS). The ERG requested the company to provide a scenario where grade 3 or higher adverse events of interest are included in the model in addition to the TEAEs.

The company advised the ERG that it was not possible to provide the requested scenario within the timeframe to respond to ERG clarification questions and instead took a pragmatic approach to provide a scenario where costs of cardiac failure are included in the model and a second scenario where AE costs are increased by 50%. Both scenarios were found to have minimal impact on the ICER. Details of the scenarios can be found in Table 25 to Table 28 of the company's response to ERG clarification questions.

Overall, the ERG considers that AEs are not a primary driver of cost-effectiveness in the model and that any amendments to how these are incorporated in the economic model are unlikely to have a substantial impact on the ICER.

4.2.7 Health-related quality of life

4.2.7.1 Health-State Utility Values

The ASPIRE study did not collect utility data directly but did collect HRQoL data using two disease-specific measures; the cancer-specific European Organization for the Research and Treatment of Cancer Quality of Life Questionnaire C30 (EORTC QLQ-C30) and the myeloma-specific EORTC QLQ-MY20. Using these data, the company applied a published mapping algorithm by Proskorovsky *et al.* 2014[22], to predict an EQ-5D-3L utility score for each patient based on their EORTC QLQ-C30 score.

This mapping study was used in the original TA457 appraisal based on an SLR of mapping studies. The company did not perform an update to this SLR but instead searched the University of Oxford Health Economics Research Centre mapping database in April 2016 to identify any more recently published studies that may be relevant. Only one study was identified but was not considered further as it was based on newly diagnosed MM patients (MYELOMA-IX). The Proskorovsky *et al.*

2014 algorithm was, therefore, used again as per the original TA457, as well as for the NICE appraisal of panobinostat (TA380).[22, 27, 31]

The predicted EQ-5D-3L utility values were then analysed using a repeated-measures mixed-effects linear regression model. The company stated that the regression model included subject-level random intercepts to account for repeated measures, and fixed effects included treatment group, baseline characteristics, and a time-dependent progression covariate. The outcome of the model was defined as change in utility from baseline.

The regression was performed in two steps. The first step assessed the significance of the effect of each potential covariate in a univariate model to determine if it was associated with the outcome based on a p-value threshold of 0.2. The next step was to include the covariates that were associated with the outcome in a multivariate regression model with a backwards stepwise variable selection procedure performed to remove variables that became non-significant at each step based on a threshold p-value of 0.1. For categorical variables, the company included the variable if at least one of the categories had a p-value < 0.2 and excluded if none of the categories had a p-value < 0.1. The resulting significant variables that were associated with affecting the outcome were carfilzomib treatment, baseline utility, ECOG performance, progression, age, neutrophil count, measurable disease category, and number of prior therapies. The final model results are given in Table 20.

Table 20. Final utility regression model results (Table 34 of the CS)

Covariate	Value	SE	p-value
(Intercept)	0.467	0.042	0.000
CRd (vs Rd)	0.016	0.009	0.075
Progression	-0.047	0.008	0.000
Baseline utility	-0.403	0.025	0.000
Age	-0.001	0.001	0.010
ECOG PS 1	-0.032	0.010	0.001
ECOG PS 2	-0.044	0.019	0.020
Absolute neutrophil count ≥ 1.5 × 109/L	-0.033	0.016	0.036
Measurable disease category: SPEP only	-0.025	0.013	0.050
Measurable disease category: UPEP only	0.009	0.020	0.637
Number of prior therapies: ≥ 2	-0.031	0.009	0.001

Abbreviations: CRd, carfilzomib/lenalidomide/dexamethasone; ECOG PS, Eastern Cooperative Oncology Group Performance Status; Rd, lenalidomide/dexamethasone; SE, standard error; SPEP, serum protein electrophoresis; UPEP, urine protein electrophoresis.

The mean predicted change from baseline for the full population was estimated to be 0.0145 and for the 2L prior bortezomib subgroup, the estimated mean change from baseline was 0.047.

For the economic model, the company used the mapped EORTC QLQ-C30 baseline utility value from the ASPIRE study based on patients with one prior therapy with bortezomib (0.714) for cycles 1 and 2. For patients in the later cycles of the pre-progression health-state, the company added the mean change from baseline estimate of 0.047 for the CRd treatment group (0.761), and from the resulting value, the company took off the treatment effect of 0.016 for the Rd treatment group (0.745). From this value, the company removed the effect estimated for progression, 0.047, and used this as the post-progression utility value for patients for both the CRd and Rd treatment groups (0.698).

Table 21. Base case health-state utility values (adapted from Table 36 of the CS)

Health state	CRd	Rd
Pre-progression (cycles 1 and 2)	0.714	0.714
Pre-progression (later cycles)	0.761	0.745
Post-progression	0.698	0.698

Abbreviations: CRd, carfilzomib/lenalidomide/dexamethasone; Rd, lenalidomide/dexamethasone.

4.2.7.2 Adverse Event Disutility values

The company modelled the impact of AEs based on the event rates observed in the ASPIRE study for treatment-related AEs that occurred in at least 5% of patients in either treatment group. The event rates are discussed in more detail in Section 4.2.6. The company used disutility values sourced from various publications. The disutility values applied as well as the sources are detailed in Table 22.

Table 22. Disutility values for AEs (adapted from Table 35 of the CS)

Adverse event	Disutility	Duration (Days)	Duration-adjusted utility decrement (per event)	Source
Neutropenia	0.145	13.20	0.005	NICE TA573[36]
Anaemia	0.310	10.70	0.009	
Thrombocytopenia	0.310	14.10	0.012	
Cataract	0.140	182.63	0.070	NICE TA297[37]
Hyperglycaemia	0.060	4.02	0.001	Disutility stated to be from Wehler *et al.* (2018): hard copy of the paper and complete reference details were not provided; Duration estimated as weighted average length of stay from NHS reference costs 2017/18; Non-elective inpatients long stay: Fluid or Electrolyte Disorders, with Interventions, KC05G to KC05N.
Lymphopenia	0.065	15.50	0.003	NICE TA573[36]
Hypertension	0.000	0.00	0.000	
Fatigue	0.115	14.60	0.005	

Hypokalaemia	0.200	0.02	0.000	Consistent with assumption made in NICE TA510[38]
Hypophosphataemia	0.000	0.00	0.000	Assumption.
Pneumonia	0.190	12.00	0.006	NICE TA573[36]

4.2.7.3 ERG critique

As mentioned previously, the ERG considers that the relevant population for this appraisal is the 2L prior bortezomib/no prior lenalidomide subgroup. In response to ERG clarification questions, the company provided the equivalent utility values for the ERG's preferred subgroup. This population showed a lower baseline utility value of [redacted] and the resulting change from baseline over time was greater at [redacted]. The resulting utility values are given alongside those for the company's base case population in Table 23. The methodology for estimating the utility values remains unchanged from the company base-case.

Table 23. Health-state utility values used in the economic model

Health state	Company base case (2L prior bortezomib)		Company scenario (2L prior bortezomib/no prior lenalidomide)	
	CRd	Rd	CRd	Rd
Pre-progression (cycles 1 and 2)	0.714	0.714	[redacted]	[redacted]
Pre-progression (later cycles)	0.761	0.745	[redacted]	[redacted]
Post-progression	0.698	0.698	[redacted]	[redacted]

Abbreviations: 2L, second-line; CRd, carfilzomib/lenalidomide/dexamethasone; Rd, lenalidomide/dexamethasone.

The ERG has two primary concerns with regard to the company's estimation of health state utility values (HSUVs). One issue relates to the company's use of the estimated mean change in utility over time to increase the HSUVs for model cycles 3 onwards, in addition to changes that relate to progression or treatment effects for instance. Change from baseline was the outcome of the utility model so the mean change from baseline is estimated from the individual effects of each covariate that is adjusted for. However mean change in utility over time was [redacted] for CRd than the Rd, even though all patients have progression-free disease.

In addition, clinical expert advice sought by the ERG suggested that there was no clinical reason for there to be a treatment-specific utility benefit in addition to the benefit provided by any gains in progression-free survival. They considered that there may be a quicker response to treatment in patients receiving carfilzomib (in CRd) compared to Rd but there would be no additional benefit beyond being progression-free.

Therefore, the ERG considers the company's application of both mean change in utility over time for the pre-progression health state and treatment-specific values may be unreliable and may overestimate the overall quality-adjusted life-years (QALYs) in favour of carfilzomib. The ERG recommends applying treatment utilities based on progression status alone without a treatment effect applied or an increase in utility from baseline. The ERG has included these assumptions in the ERG's preferred base case, presented in Section 6.4.

A secondary issue that the ERG was concerned about the lack of information on the company's variable selection procedure for the adjustment of utility values and the inclusion of urine protein electrophoresis (UPEP), which was not statistically significant. However, as a result of their response to ERG clarification questions, the company provided more details on their methods and highlighted UPEP was part of a categorical variable called "Measurable disease category". This variable also included serum protein electrophoresis (SPEP) as a category and this had a significant p-value of 0.05, and therefore, the whole categorical variable was included.

The ERG considers that even though the effect estimate for "UPEP only" was relatively small at 0.009, it may have been more appropriate to specify the model differently where categorical variables produced levels with non-significant effect estimates. Particularly for variables like UPEP and SPEP, which could have been included as separate independent variables rather than a single categorical variable. This may have made the variable selection procedure more robust and provided potentially more reliable results but unlikely to have a large impact on the results.

Lastly, the ERG was concerned that the company's SLR was performed nearly two years ago but considers the sources of evidence used to be reasonable and that the date of the SLR is unlikely to have missed any evidence that could have impacted on utility estimates.

4.2.8 Resource use and costs

In the economic analyses, the company included the costs of drug acquisition, administration of drugs, concomitant medications, routine monitoring, treatment of AEs and the costs of palliative care. Each of these is described in the following subsections.

4.2.8.1 Drug acquisition

The company sourced the unit costs for each of the branded drugs in the model using the Monthly Index of Medical Specialties (MIMS)[29], while for generic drugs the company used the electronic market information (eMIT) tool[30].

Carfilzomib and lenalidomide both have a patient access scheme (PAS), and both provide a simple discount on the list price to the health care provider. For carfilzomib, the simple discount is [redacted]. A confidential appendix accompanies this document to provide the results of the cost-effectiveness analysis with the comparator PAS for lenalidomide as well as the carfilzomib PAS applied. A summary of the drug acquisition costs relating to the actual doses received in ASPIRE for the 2L prior bortezomib subgroup, and assuming no wastage, is given in Table 24. Where applicable, the prices with the PAS discounts applied are given in brackets.

Further details of the regimens for the intervention and comparator are given in Section 4.2.3.

Table 24. Drug acquisition costs per 28-day cycle

Health state	CRd List prices (PAS prices)	Rd List prices
Carfilzomib	Cycle 1: £4,230 ([redacted]) Cycle 2-12: £4,630 ([redacted]) Cycles 13-18: £3,087 ([redacted])	NA
Lenalidomide	£4,050	£4,058
Dexamethasone	£16	£16
Total	**Cycle 1: £8,295 ([redacted])** **Cycle 2-12: £8,695 ([redacted])** **Cycles 13-18: £7,152 ([redacted])**	**£4,075**

Abbreviations: CRd, carfilzomib with lenalidomide and dexamethasone; PAS, patient access scheme; Rd, lenalidomide with dexamethasone.

The company also accounted for the relative dose intensity (RDI) of each of the regimens to factor in doses that were not received and therefore did not incur costs. The RDIs were calculated as the percentage of planned doses that were actually received and these were multiplied by the drug acquisition costs per cycle. The RDIs for each regimen in each treatment group are given in Table 25.

Table 25. Relative dose intensity

Regimen	CRd	Rd
Carfilzomib	90.72%	NA
Lenalidomide	80.27%	79.46%
Dexamethasone	79.93%	82.90%

Abbreviations: CRd, carfilzomib with lenalidomide and dexamethasone; Rd, lenalidomide with dexamethasone.

4.2.8.2 Administration costs

Administration costs were included for carfilzomib based on the simple parenteral chemotherapy at first attendance cost code (SB12Z) from NHS reference costs 2018[28]. Specifically, it was based on an outpatient setting cost, which was estimated to be £174.40 per administration. The overall administration costs per cycle were estimated as £1,010 for cycles 1-12, and £674 for cycles 13 onwards. The difference is a result of the reduced frequency of doses after cycle 12.

For lenalidomide and dexamethasone, no administration costs were assumed, as these are oral drugs that do not require any resource use for administration.

Further details on the regimens for the intervention and comparator are given in Section 4.2.3.

4.2.8.3 Concomitant Medication Costs

The costs of concomitant medications were applied in the model based on those received in the ASPIRE trial. The medications received were valacyclovir, lansoprazole and aspirin. The proportions of patients receiving these medications in each group of the ASPIRE trial were used to estimate a weighted per-cycle cost to be applied in the model. The estimated costs per-cycle were £5.88 for the CRd group and £4.27 for the Rd group. Further details can be found in Table 43 of the CS.

4.2.8.4 Routine Monitoring Costs

The company included costs of routine monitoring in addition to the costs incurred from administration of drugs. The expected resource use was estimated by the company based on a non-interventional, observational chart review study using retrospective data collected from medical records of patients with symptomatic multiple myeloma.[39]

To collect data for the chart review, 56 oncologists and haematologists in the UK were asked to complete electronic forms to provide retrospective data on patient characteristics, treatments, response, costs and resource use.[39] Costs included outpatient consultations, lab tests, scans and other relevant procedures, and were separated by on-treatment and off-treatment for the pre-progression phase but did not split by treatment regimen as the average cost of resource use was considered similar across treatment regimens. Post-progression costs were also considered separately and were not treatment specific but did consider the subsequent treatment phase separately from the best supportive care phase.

A summary of the costs used in the economic model, inflated to 2018 prices using the PSSRU hospital and community health services pay and prices index[40], are summarised in Table 26.

Table 26. Monitoring costs per 28-day cycle

Health state	Costs per 28-day cycle (inflated to 2018 prices)
Progression-free (on treatment)	£94.51
Progression-free (off treatment)	£64.32
Post-progression (on subsequent treatment)	£94.51
Post-progression (BSC)	£194.78

Abbreviations: BSC, best supportive care; CRd, carfilzomib with lenalidomide and dexamethasone; PAS, patient access scheme; Rd, lenalidomide with dexamethasone.

4.2.8.5 Adverse Event Costs

The company included the costs associated with treating AEs for both the CRd and Rd groups based on the safety population of the ASPIRE trial. The company restricted the AEs included in the model to those that were grade 3 and above and occurred in at least 5% of patients in at least one group of the ASPIRE trial. Further detail of AEs included in the economic model is given in Section 4.2.6.

The unit costs applied to the proportion of AEs estimated for each model cycle were based on either inpatient, outpatient, day case or general practice treatment settings depending on the AE. Details of the specific AEs can be found in Table 48 on page 112 of the CS.

Unit costs for each AE in each setting were based on NHS reference costs 2018. Further details can be found in Table 49 on page 113 of the CS. The total cost of AE treatment per model cycle for the CRd treatment group was estimated to be £54.40, while for the Rd treatment group it was estimated to be £56.15.

4.2.8.6 Subsequent Treatment Costs

The company included the costs of subsequent treatments that would be expected to be received by patients in each of the treatment groups of the ASPIRE trial. Following either CRd or Rd as the primary treatment, the company assumed that patients would subsequently receive panobinostat in combination with bortezomib and dexamethasone (PVd) followed by pomalidomide in combination with dexamethasone (Pd), based on the current treatment pathway in England and Wales and the proposed positioning of CRd. A treatment-free interval of three cycles was included in the model, based on data from ASPIRE which estimated time between progression and start of subsequent treatment. The treatment-free interval was assumed to be the same, irrespective of prior lines of therapy.

Unit costs of the regimens were sourced from MIMS[29] or eMIT[30] as per the primary intervention and comparator regimens. The company stated in the CS that they assumed that 80% of patients would receive 3L therapy and that 20% received no treatment but did not specify their assumptions regarding 4L treatment. The company's model appears to assume that PVd should be received by 100% of patients and Pd by 66% of patients, based on a Kantar Health chart review.[41] The total cost of PVd estimated per-cycle was £8,432 and the estimated cost of Pd per-cycle was £8,900. The resulting cost per-cycle applied for both the CRd and Rd treatment groups in the model was £7,295. The company assumed a duration of five months for PVd based on median duration from the PANORMA-1 trial, and a duration of four months for Pd based on the pomalidomide NICE appraisal[32], resulting in a total of nine cycles for subsequent therapy.

The company included an administration cost of £89 per cycle for bortezomib based on the specialist nursing, cancer related, Adult, Face to face cost code (N10AF) from NHS reference costs 2018[28]. The company did not include administration costs for panobinostat, lenalidomide, pomalidomide and dexamethasone as these are oral treatments.

4.2.8.7 Palliative Care Costs

All patients are assumed to incur costs of palliative care covering resources used in the 90 days prior to death. The company used estimates from Georghiou and Bardsley 2014[42], which were inflated to 2018 prices using the PSSRU hospital and community health services pay and prices index[40]. A summary of these costs is presented in Table 27.

Table 27. Palliative care costs per 28-day cycle

Health state	Costs per 28-day cycle (inflated to 2018 prices)
District nurse	£308
Nursing and residential care	£1,141
Hospice care (in-patient)	£609
Hospice care (final 3 months)	£4,985
Marie Curie nursing service	£609
Total	**£7,653**

Abbreviations: BSC, best supportive care; CRd, carfilzomib with lenalidomide and dexamethasone; PAS, patient access scheme; Rd, lenalidomide with dexamethasone.

4.2.8.8 ERG critique

As mentioned previously, the ERG considers that the relevant population for this appraisal is the 2L prior bortezomib/no prior lenalidomide subgroup. In response to ERG clarification questions, the

company provided updated RDI for each of the regimens (Table 28) and weighted average cost per dose for lenalidomide for CRd and Rd (Table 29).

Table 28. Relative dose intensity by subgroup

Regimen	Company base case (2L prior bortezomib)		Company scenario (2L prior bortezomib/no prior lenalidomide)	
	CRd	Rd	CRd	Rd
Carfilzomib	90.72%	NA	[redacted]	[redacted]
Lenalidomide	80.27%	79.46%	[redacted]	[redacted]
Dexamethasone	79.93%	82.90%	[redacted]	[redacted]

Abbreviations: 2L, second-line; CRd, carfilzomib with lenalidomide and dexamethasone; NA, not applicable; Rd, lenalidomide and dexamethasone.

Table 29. Weighted average lenalidomide cost per dose by subgroup

Treatment arm	Company base case (2L prior bortezomib)	Company scenario (2L prior bortezomib/no prior lenalidomide)
CRd	£192.84	[redacted]
Rd	£193.24	[redacted]

Abbreviations: 2L, second-line; CRd, carfilzomib with lenalidomide and dexamethasone; Rd, lenalidomide with dexamethasone.

The ERG considers the company's methods regarding the estimation of unit costs and resource use to be generally reasonable. However, the ERG highlights two issues regarding monitoring costs and subsequent treatment, which warrant further investigation.

An issue that could have an important impact on the cost-effectiveness results is the subsequent treatment costs that are applied for both CRd and Rd in the economic model. The company's application of costs relating to the anticipated treatment pathway in England may seem plausible; however, it may not necessarily reflect the treatments received in the ASPIRE trial from which the treatment effectiveness estimates were acquired. This potentially causes bias in the economic analysis and the ERG considers that it is more appropriate to apply treatment costs based on the treatments received by patients in the 2L prior bortezomib subgroup of the ASPIRE trial. This then aligns the treatment effectiveness data with the costs of the treatments that have impacted on those data. The potential drawback with this approach is that some patients may have received treatments that are not recommended by NICE and, therefore, may have prices that do not reflect a cost-effective use of resources.

In response to the ERG's clarification questions, the company provided details of the subsequent treatments received by patients in the ASPIRE trial (Table 30). The company also supplied the simplified analysis in which subsequent treatment costs in the model were estimated based on some

of the key treatments received in the ASPIRE trial, which included Vd, Pd and Rd as subsequent treatments. This resulted in total per-cycle costs of £2,497 and £2,032 for CRd and Rd, respectively.

Bortezomib appears to be the key treatment that has a relatively large difference in usage across the treatment groups. The ERG's clinical experts advised that it is reasonable that more patients on Rd would be given bortezomib as a third-line treatment compared with CRd patients.

However, there is also a notable difference in investigational drugs, which appears to be largely monoclonal antibodies including daratumumab, based on the footnotes in the company's table (Table 32 of the company's response to clarification document). Daratumumab is an expensive and effective drug, and therefore, the company's omission of this from their estimation of subsequent treatment costs is likely to underestimate the total costs. The benefits, however, are likely to have overestimated the overall survival observed in the ASPIRE trial and therefore it is important that these costs are included to align with the overall survival. The ERG has provided a scenario analysis to include investigational drugs in the subsequent treatment costs with the assumption that costs are based on daratumumab costs. The results of this scenario are given in Section 6.2.

The ERG considers that the company may have misinterpreted the evidence it has used in its approach to weighting costs for subsequent treatments. In the economic model, 80% of progressed patients go on to receive subsequent therapy. However, the company has assumed that of those 80% of patients, 66% of patients will receive fourth- and fifth-line treatment, based on data from a conference poster.[41] However, the ERG considers that the fourth-line cohort in the study is a percentage of the total cohort and not a sub-population of the third-line cohort, as has been assumed in the economic-model. The ERG has conducted a scenario, where the weighting of subsequent treatment costs assumes 80% of costs for third-line treatment and 66% of costs for fourth- and fifth-line treatments. Results of the scenario are presented in Section 6.2.

A secondary issue, raised by the ERG's clinical experts, was that monitoring costs seemed quite low and were likely to be an underestimate of the true monitoring costs for 2L multiple myeloma patients. However, the ERG reviewed relevant submissions from previous appraisals and noted that for the daratumumab appraisal (TA573) the routine monitoring costs were actually lower and these were accepted by the ERG and subsequently the committee.[36]As such, the ERG considers the company's estimates to be conservative and acceptable. Nonetheless, the ERG tested the impact of increasing the routine monitoring costs in the PFS health state by 50% and found that this had a minimal impact on the ICER. The full results of this scenario are given in Section 6.3

The ERG had some concerns regarding the resource use assumed for the treatment of AEs based on clinical expert opinion but found that changes in the cost assumptions had minimal impact on the ICER.

Table 30. Subsequent antimyeloma therapies reported for ≥2% of patients in any treatment arm of the intent-to-treat population (Adapted from Table 32 of the company's clarification response)

	2L / prior bortezomib			2L / prior bortezomib / no prior lenalidomide		
	CRd (N=93) n (%)	Rd (N=73) n (%)	Mean DOT	CRd (N=74) n (%)	Rd (N=66) n (%)	Mean DOT
Nr of patients experienced progression	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Nr. of patients treated with ≥1 antimyeloma therapy	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Antineoplastic agents						
Bortezomib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Cyclophosphamide	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Doxorubicin	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Melphalan	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Pomalidomide	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Bendamustine	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Carfilzomib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Etoposide	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Cisplatin	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Immunosuppressants						
Lenalidomide	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Thalidomide	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Corticosteroids						
Dexamethasone	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Prednisone	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
All other therapeutic products						
Investigational drug‡	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Blood substitutes and perfusion solutions						
Blood and related products	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: 2L, second-line; CRd, carfilzomib with lenalidomide and dexamethasone; DOT, duration of treatment; Rd, lenalidomide with dexamethasone.

5 Cost-effectiveness results

5.1 Company base-case results

The results of the company's base-case analysis are given in Table 31, showing an incremental cost-effectiveness ratio of £43,952 per QALY gained for CRd versus Rd. These results include the company's agreed PAS for carfilzomib, which provides a discount of [redacted] on the list price.

Table 31. Company's base case results

Interventions	Total Costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£/QALY)
Rd	[redacted]	4.08	2.58	-	-	-	-
CRd	[redacted]	6.62	3.96	60,467	2.54	1.38	43,952

Abbreviations: CRd, carfilzomib with lenalidomide and dexamethasone; ERG, evidence review group; ICER, incremental cost-effectiveness ratio; QALY, quality-adjusted life-year; Rd, lenalidomide and dexamethasone.

5.2 Company's sensitivity analyses

5.2.1 Probabilistic sensitivity analysis

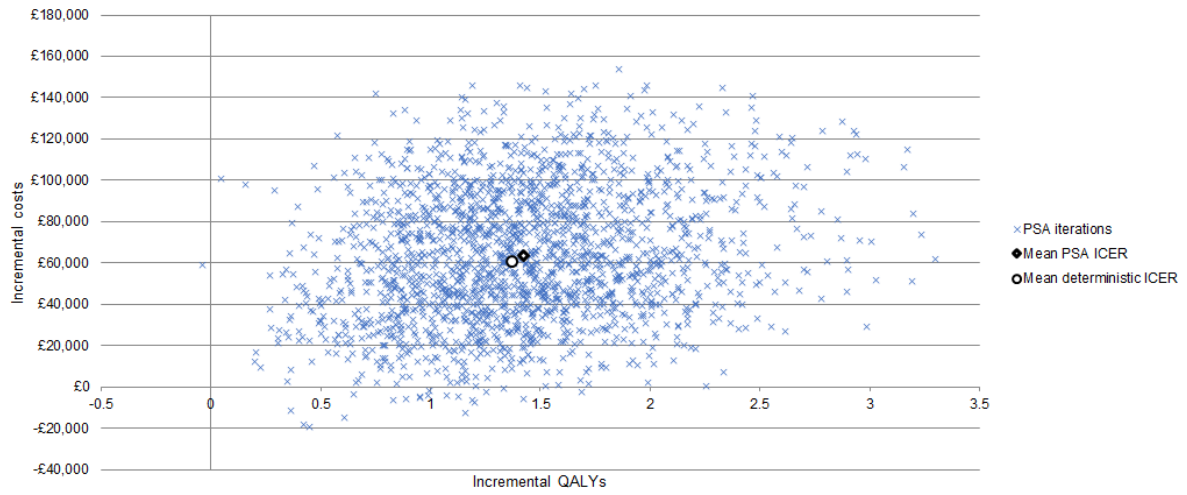
The company conducted a probabilistic sensitivity analysis (PSA) based on 2,000 samples. In response to ERG clarification questions, the company provided a corrected analysis after the ERG identified an unusual clustering of points on the cost-effectiveness plane produced when the ERG ran 10,000 samples in the company's economic model. The corrected PSA results are presented in Table 32, and a scatterplot of the 2,000 sampled costs and QALYs on the cost-effectiveness plane are presented in Figure 18.

Table 32. Company's PSA results

Interventions	Total Costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£/QALY)
Rd	[redacted]	4.08	2.58	-	-	-	-
CRd	[redacted]	6.78	4.00	63,873	2.70	1.42	44,902

Abbreviations: CRd, carfilzomib with lenalidomide and dexamethasone; ERG, evidence review group; ICER, incremental cost-effectiveness ratio; QALY, quality-adjusted life-year; Rd, lenalidomide and dexamethasone.

Figure 18. Scatterplot of PSA samples on cost-effectiveness plane (pairwise – CRd vs Rd)

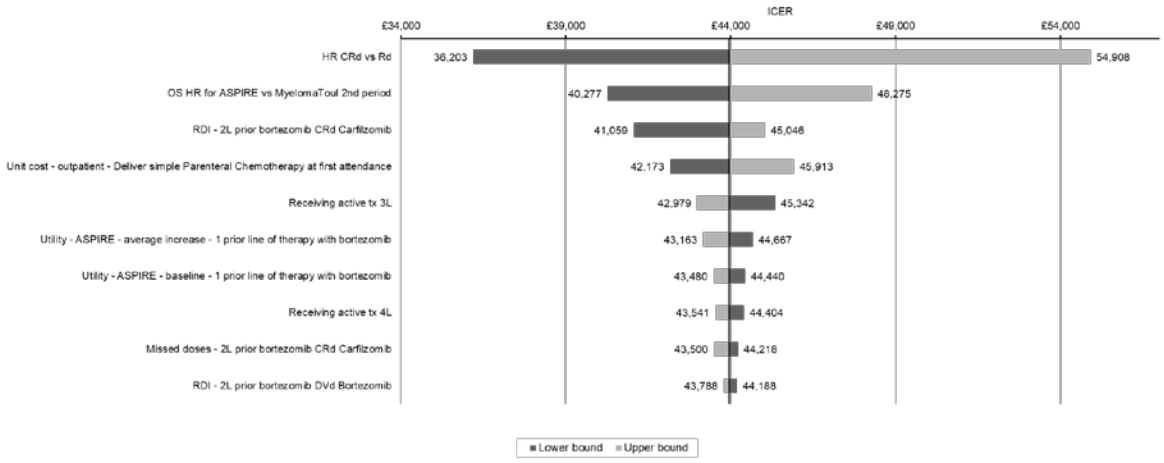


Abbreviations: CRd, Carfilzomib plus lenalidomide and dexamethasone; PSA, probabilistic sensitivity analysis; Rd, lenalidomide plus dexamethasone.

5.2.2 One-way sensitivity analyses

The company conducted a range of one-way sensitivity analyses (OWSAs) to test the impact that plausible changes on parameters have on the overall results. The tornado plot in Figure 19 shows the parameters that had the greatest impact, with the OS HR for CRd versus Rd having the greatest impact, resulting in ICERs ranging from £36,203 to £54,908 per QALY.

Figure 19. Tornado plot of OWSA results

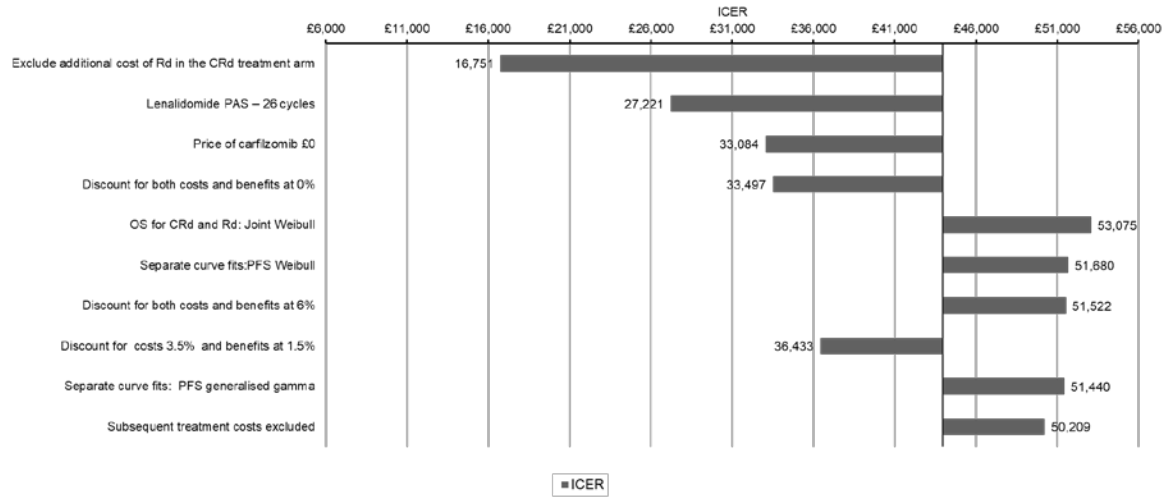


Abbreviations: 2L, second-line; 3L, third-line; 4L, fourth-line; CRd, Carfilzomib plus lenalidomide and dexamethasone; HR, hazard ratio; OS, overall survival; PFS, progression-free survival; Rd, lenalidomide plus dexamethasone; RDI, relative dose intensity.

5.2.3 Scenario analyses

The company provided a range of scenario analyses around their base case, which are detailed in full in Table 56 on page 128 of the CS. The results of the scenarios that had the greatest impact are shown in the tornado plot in Figure 20.

Figure 20. Tornado plot of scenario analysis results



Abbreviations: CRd, Carfilzomib plus lenalidomide and dexamethasone; ICER, incremental cost-effectiveness ratio; OS, overall survival; PFS, progression-free survival; Rd, lenalidomide plus dexamethasone; RDI, relative dose intensity.

5.3 Model validation and face validity check

Quality assurance was performed by the external company who developed the model. A health economist not involved with model development reviewed the model for coding errors, inconsistencies and validity of model parameters.

6 Additional economic analysis undertaken by the ERG

6.1 Model corrections

The ERG did not identify any model errors.

6.2 Exploratory and sensitivity analyses undertaken by the ERG

In Section 4 of this report, the ERG has described several scenarios that warrant further exploration in addition to the company's own sensitivity and scenario analyses to ascertain the impact of these changes on the incremental cost-effectiveness ratio (ICER). The deterministic scenarios the ERG has produced are applied to the company's alternative cost-effectiveness scenario for the 2L prior bortezomib/no prior lenalidomide subgroup provided by the company in their response to ERG clarification questions and are as follows:

1. Implementation of the company's jointly fitted exponential distribution for ASPIRE inverse probability weighted (IPW) overall survival (OS) subgroup data - Section 4.2.5.1
2. Construction of progression-free survival (PFS) and OS curves for the carfilzomib with lenalidomide and dexamethasone treatment arm (hereafter to referred to as CRd) using intention-to-treat (ITT) PFS and OS hazard ratios (HRs) applied to the company scenario PFS and OS curves for the lenalidomide with dexamethasone treatment arm (hereafter referred to as Rd) - Section 4.2.5.1
3. Alternative construction of PFS and OS curves for CRd using the ITT PFS HR applied to the company scenario Rd PFS curve and ITT OS HR applied to ERG alternative OS modelling for Rd using the exponential distribution for ASPIRE data only - Section 4.2.5.1
4. Weibull distribution for CRd time-to-treatment discontinuation (TTD) – Section 4.2.5.1
5. No treatment effect applied for pre-progression health state utility value – Section 4.2.7.3
6. No average increase in baseline utility from cycle three onwards – Section 4.2.7.3
7. Combination of scenarios five and six.
8. Inclusion of investigational drugs in the company's subsequent treatment scenario using the ASPIRE trial data and assumed costs of daratumumab – Section 4.2.8.8
9. Assuming a 50% increase in costs for routine monitoring in the PFS health state – Section 4.2.8.8
10. Alternative approach to weighting costs of subsequent treatment – Section 4.2.8.8

6.3 ERG scenario analysis

Table 35 presents the results of the ERG exploratory analyses described in Section 6.2. Results reported include the company's proposed patient access scheme (PAS) of [redacted].

Table 35. Results of the ERG's scenario analyses

	Results per patient	Intervention - CRd	Comparator - Rd	Incremental value
0a	**Company base case**			
	Total costs (£)	[redacted]	[redacted]	£60,467
	QALYs	3.96	2.58	1.38
	ICER (£/QALY)			**43,952**
0b	**Company scenario for the 2L prior bortezomib/no prior lenalidomide subgroup**			
	Total costs (£)	[redacted]	[redacted]	54,626
	QALYs	3.94	2.58	1.35
	ICER (£/QALY)			**40,335**
1	**Jointly fitted exponential distribution for OS – ASPIRE only**			
	Total costs (£)	[redacted]	[redacted]	£3,017
	QALYs	3.68	2.52	1.15
	ICER (£/QALY)			**45,919**
2	**PFS and OS CRd curves using ITT PFS and OS HR applied to company scenario PFS and OS**			
	Total costs (£)	[redacted]	[redacted]	52,235
	QALYs	3.26	2.58	0.68
	ICER (£/QALY)			**76,716**
3	**PFS and OS CRd curves using ITT PFS HR applied to company scenario Rd PFS curve and ITT OS HR applied to ERG preferred Rd OS curve**			
	Total costs (£)	[redacted]	[redacted]	52,261
	QALYs	3.16	2.52	0.64
	ICER (£/QALY)			**81,593**
4	**Weibull distribution for CRd TTD**			
	Total costs (£)	[redacted]	[redacted]	54,918
	QALYs	3.94	2.58	1.35
	ICER (£/QALY)			**40,552**
5	**No treatment effect applied for pre-progression health state utility value**			
	Total costs (£)	[redacted]	[redacted]	54,626
	QALYs	3.96	2.64	1.32
	ICER (£/QALY)			**41,303**
6	**No average increase in baseline utility from cycle three onwards**			
	Total costs (£)	[redacted]	[redacted]	54,626
	QALYs	3.68	2.43	1.25
	ICER (£/QALY)			**43,583**
7	**Scenarios 5 and 6**			
	Total costs (£)	[redacted]	[redacted]	54,626
	QALYs	3.68	2.46	1.23
	ICER (£/QALY)			**44,438**
8	**Inclusion of investigational drugs cost for subsequent therapy based on ASPIRE**			

	Total costs (£)	[redacted]	[redacted]	57,768
	QALYs	3.94	2.58	1.35
	ICER (£/QALY)			**42,657**
9	**50% increase in costs for routine monitoring in the PFS health state**			
	Total costs (£)	[redacted]	[redacted]	55,396
	QALYs	3.94	2.58	1.35
	ICER (£/QALY)			**40,903**
10	**Alternative weighting of subsequent treatment costs**			
	Total costs (£)	[redacted]	[redacted]	54,512
	QALYs	3.94	2.58	1.35
	ICER (£/QALY)			**40,253**

Abbreviations: 2L, second-line; CRd, carfilzomib with lenalidomide and dexamethasone; HR, hazard ratio; ICER, incremental cost-effectiveness ratio; ITT, intention-to-treat; OS, overall survival; PFS, progression-free survival; QALY, quality-adjusted life-year; Rd, lenalidomide plus dexamethasone; TTD, time-to-treatment discontinuation

6.4 ERG preferred assumptions

In this section, the ERG presents its base-case ICER for the 2L prior bortezomib/no prior lenalidomide subgroup. Deterministic results are presented in Table 36 and incorporate the company's patient access scheme (PAS) simple discount of [redacted]. The PSA ICER for the ERG preferred base-case is £55,530.

Table 36. ERG's preferred model assumptions

Preferred assumption	Section in ERG report	Cumulative ICER £/QALY
Company base case	*5.1*	**43,952**
Corrected company scenario for the 2L prior bortezomib/no prior lenalidomide subgroup	*4.2.5.1 & 6.2*	**40,335**
Jointly fitted exponential distribution for OS – ASPIRE only	*4.2.5.1*	**45,919**
Removal of treatment effect and average increase in utility for cycle three onwards for pre-progression health state utility value	*4.2.7.3*	**50,960**

Abbreviations: 2L, second-line; ERG, evidence review group; ICER, incremental cost-effectiveness ratio; OS, overall survival; QALY, quality-adjusted life-year

6.5 Conclusions of the cost-effectiveness sections

The final scope provide by the National Institute of Health and Care Excellence (NICE) listed that the relevant population to assess the cost-effectiveness of carfilzomib in combination with lenalidomide and dexamethasone (CRd) is adult patients with multiple myeloma who have received at least one prior therapy. The company deviated from the NICE final scope by restricting the proposed population to only one prior therapy with bortezomib (2L prior bortezomib subgroup). Based on advice from clinical experts, the ERG accepts the company's justifications for the positioning of CRd for the 2L prior bortezomib subgroup, which they state reflects the need for triplet therapies earlier in the pathway, greater benefit of the treatment is demonstrated in this subgroup and thus offers the greatest economic value, and lastly, it aligns with the NICE recommendation for Rd, which is deemed the most relevant comparator. Furthermore, the ERG considers that not exploring subgroups where cost-effectiveness cannot be demonstrated is appropriate and pragmatic.

However, as mentioned in Section 3.5, the company's subgroup analysis included a proportion of people who had not received bortezomib as part of their last round of therapy as well as people who had undergone treatment with lenalidomide in their last regimen. In response to ERG clarification

questions, the company provided scenario analyses for the subgroup excluding patients who had received prior lenalidomide (2L prior bortezomib/no prior lenalidomide), which the ERG considers more accurately represents the company's population of interest and as such is the subgroup considered for the ERG preferred analyses.

Overall, the ERG considers the company's approach to estimating PFS and TTD is appropriate and unbiased. With regards to the modelling of adverse events (AEs), the ERG had some concerns but on balance found that any changes to the modelling assumptions for AEs had minimal impact on the ICER and thus were not considered a primary driver of cost-effectiveness. Moreover, the ERG investigated the impact on the ICER of alternative assumptions for estimating monitoring costs and the inclusion of investigational drugs costs in the subsequent therapy pathway based on data from ASPIRE but found these did not produced a meaningful difference.

One of the primary issues with the cost-effectiveness analysis is the company's approach to estimating OS, using real-world data to adjust mature trial-based data. For the base-case analysis of OS for Rd, the company constructed a hybrid survival curve based on extrapolated ASPIRE IPW OS data and real-world evidence from a French registry of multiple myeloma patients, MyelomaToul. For the CRd arm, OS is also based on extrapolated ASPIRE IPW OS data and MyelomaToul data adjusted using the IPW OS HR from ASPIRE. The company chose this approach as they deemed the survival estimates based solely on ASPIRE using the Weibull distribution, which they deemed the best-fitting distribution to the observed data, produced pessimistic results for the Rd arm.

The ERG consulted its clinical experts who confirmed that longer-term survival estimates for Rd patients based on ASPIRE are conservative. However, the consequence of the company's adjustment when using real-world evidence is that survival is inflated for CRd compared with the extrapolated estimates based on IPW OS ASPIRE data. As such, the ERG considers that the company could have chosen a more clinically plausible extrapolation of the ASPIRE data to use for the base-case. The company confirmed that if they used MyelomaToul to validate their extrapolations, the exponential distribution would have been appropriate to estimate OS. The ERG considers that the exponential distribution produced similar survival estimates for Rd compared with company's base-case estimates. Furthermore, the CRd OS survival estimates are based entirely on mature ASPIRE OS data, which the ERG deems is appropriate and reduces the uncertainty in the analysis.

It should be noted that in the company submission (CS), the company highlight that for PFS, "*there is a consistent treatment effect across baseline covariate subgroups*". HRs derived from an ITT

population of an RCT are, by their nature, more robust than those generated from a subgroup analysis, which is based on *post-hoc* data where randomisation has been broken and the sample size reduced. However, this appraisal is a part-review of TA457 where results from subgroups adjusted to account for imbalances in baseline characteristics arising from non-randomised groups was accepted by the committee. As such, the ERG considers that the company's IPW analysis to adjust subgroup data for imbalances can be considered for appropriate for decision-making. However, as an illustrative scenario the ERG tested the impact of utilising ITT hazard ratios (HRs) for PFS and OS and found that it increased the ICER by almost £40,000 when combined with the other ERG preferences for modelling OS.

Separately from OS, the ERG had a concern with the assumptions made by the company for the estimation of utility values for the progression-free health state. Specifically, pre-progression utility values in the model capture both mean increase in utility from baseline for both treatment arms as well as treatment-specific increase in utility if a patient is on CRd. Change from baseline was the outcome of the utility model so the mean change from baseline is estimated from the individual effects of each covariate that is adjusted for. However mean change in utility over time was [redacted] for CRd than the Rd, even though all patients have progression-free disease. Furthermore, clinical expert advice sought by the ERG suggested that there was no clinical reason for there to be a treatment-specific utility benefit in addition to the benefit provided by any gains in progression-free survival. Thus, the ERG considers that it is more appropriate for pre-progression utility values for both treatment arms to be equal and that difference in pre-progression quality-adjusted life-years (QALYs) should be determined by length of time spent in the progression-free health state.

In conclusion, the ERG considers that the original uncertainty in TA457 has been resolved by more mature OS and PFS from ASPIRE.[27] As such, the ERG considers the ICER for the ERG preferred analysis to be robust.

7 End-of-Life

NICE end-of-life status should be applied when the following criteria are satisfied:

(i) the treatment provides an extension to life of more than an average of three months compared to current NHS treatment, and;

(ii) the treatment is indicated for patients with a short life expectancy, normally a mean life expectancy of less than 24 months.

The company state that second-line (2L) carfilzomib plus lenalidomide and dexamethasone (CRd) meets the first criterion of extension to life but does not meet the second criterion of short life expectancy. The ERG agrees with the company's evaluation and the case has not been made for end-of-life.

However, the company highlighted that for the appraisal of pertuzumab for HER2 positive metastatic cancer (TA509)[43], committees can use the following criteria to apply discretion and agree to end-of-life status for treatments for metastatic cancer when:

- OS without new drug exceeds 24 months;
- The new drug provides significant extension to life beyond three months, and;
- The new drug is combined with existing treatment, and;
- Both the existing treatment and the new drug are used until disease progression.

The company stated that 2L CRd meets these additional, discretionary criteria. However, the final appraisal document for TA509 stated that pertuzumab, "*has been available on the cancer drugs fund for several years and the committee recognised this as an exceptional circumstance. In this context, committee considered it reasonable to apply flexibility in its interpretation of the criteria for special consideration as a life-extending treatment for people with a short life expectancy, but that the weight applied to the quality adjusted life years gained would not be at the maximum allocated in other, more regular, circumstances where the end of life criteria have been applied*".[43]

Thus, the ERG does not consider the company's request for flexibility is warranted as CRd is not in the Cancer Drugs Fund for the subgroup under consideration in this appraisal.

8 References

1. National Institute for Health and Care Excellence. Carfilzomib for previously treated multiple myeloma (part review of TA457) (ID1493), 2019. Available from: https://www.nice.org.uk/guidance/gid-ta10510/documents/final-scope. Date accessed: 28 Jan 2020.
2. National Institute for Health and Care Excellence. Bortezomib for induction therapy in multiple myeloma before high-dose chemotherapy and autologous stem cell transplantation, 2014. Available from: https://www.nice.org.uk/guidance/TA311. Date accessed: 28 Jan 2020.
3. Kurteva E, Rajani S, Chadokufa S, Sider S, Sawney T, Buckingham R, et al. A case of de novo Inflammatory Bowel Disease (IBD) post thymic transplant in a patient with di George syndrome. *Journal of Pediatric Gastroenterology and Nutrition* 2018; **66**: 548.
4. Excellence NIfHaC. Bortezomib and thalidomide for the first-line treatment of multiple myeloma, 2011. Available from: https://www.nice.org.uk/guidance/TA228. Date accessed: 28 Jan 2020.
5. National Institute for Health and Care Excellence. Lenalidomide plus dexamethasone for previously untreated multiple myeloma, 2019. Available from: https://www.nice.org.uk/guidance/TA587. Date accessed: 28 Jan 2020.
6. National Institute for Health and Care Excellence. Daratumumab with bortezomib and dexamethasone for previously treated multiple myeloma, 2019. Available from: https://www.nice.org.uk/guidance/TA573. Date accessed: 28 Jan 2020.
7. National Institute for Health and Care Excellence. Lenalidomide plus dexamethasone for multiple myeloma after 1 treatment with bortezomib, 2019. Available from: https://www.nice.org.uk/guidance/TA586. Date accessed: 28 Jan 2020.
8. National Institute for Health and Care Excellence. Carfilzomib for previously treated multiple myeloma, 2017. Available from: https://www.nice.org.uk/guidance/TA457. Date accessed: 28 Jan 2020.
9. Stewart AK, Rajkumar SV, Dimopoulos MA, Masszi T, Spicka I, Oriol A, et al. Carfilzomib, lenalidomide, and dexamethasone for relapsed multiple myeloma. *N Engl J Med* 2015; **372**: 142-52.
10. European Medicines Agency. Kyprolis (carfilzomib). An overview of Kyprolis and why it is authorised in the EU, 2018. Available from: https://www.ema.europa.eu/en/documents/overview/kyprolis-epar-medicine-overview_en.pdf. Date accessed: 04 Feb 2020.
11. Amgen. Abbreviated Clinical Study Report: Onyx. A randomized, multicenter, Phase 3 study comparing carfilzomib, lenalidomide, and dexamethasone (CRd) vs lenalidomide and dexamethasone (Rd) in subjects with relapsed multiple myeloma. 2017.
12. National Institute for Health and Care Excellence. Carfilzomib for previously treated multiple myeloma. Final Appraisal Determination, 2017. Available from: https://www.nice.org.uk/guidance/ta457/documents/final-appraisal-determination-document. Date accessed: 04 Feb 2020.
13. Dimopoulos M, Spencer A, Attal M, Prince HM, Harousseau JL, Dmoszynska A, et al. Lenalidomide plus dexamethasone for relapsed or refractory multiple myeloma. *N Engl J Med* 2007; **357**: 2123-32.
14. Weber DM, Chen C, Niesvizky R, Wang M, Belch A, Stadtmauer EA, et al. Lenalidomide plus dexamethasone for relapsed multiple myeloma in North America. *N Engl J Med* 2007; **357**: 2133-42.
15. electronic Medicines Compendium (eMC). Kyprolis powder for solution for infusion, 2019. Available from: https://www.medicines.org.uk/emc/product/5061. Date accessed: 03 Feb 2020.
16. National Institute for Health and Care Excellence. Managing myeloma, 2020. Available from: https://pathways.nice.org.uk/pathways/myeloma#path=view%3A/pathways/myeloma/managing-myeloma.xml&content=view-index. Date accessed: 28 Jan 2020.

17. Centre for Reviews and Dissemination (CRD). CRD's guidance for undertaking reviews in healthcare, 2011. Available from: https://www.york.ac.uk/media/crd/Systematic_Reviews.pdf. Date accessed: 17 Feb 2020.
18. Higgins JPT GS. *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0*. Collaboration TC, editor2011.
19. Sperandei S. Understanding logistic regression analysis. *Biochem Med (Zagreb)* 2014; **24**: 12-8.
20. Latimer N. NICE DSU Technical Support Document 14: Survival analysis for Economic Evaluations Alongside Clinical Trials - Extrapolation with Patient-Level Data, 2013. Available from: http://www.nicedsu.org.uk/NICE%20DSU%20TSD%20Survival%20analysis.updated%20March%202013.v2.pdf. Date accessed: 01/03/2017.
21. National Institute for Health and Care Excellence: Decision Support Unit (NICE DSU). 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. , 2015. Available from: http://nicedsu.org.uk/wp-content/uploads/2016/03/TSD17-DSU-Observational-data-FINAL.pdf. Date accessed: 03 Mar 2020.
22. Proskorovsky I, Lewis P, Williams CD, Jordan K, Kyriakou C, Ishak J, et al. Mapping EORTC QLQ-C30 and QLQ-MY20 to EQ-5D in patients with multiple myeloma. *Health Qual Life Outcomes* 2014; **12**: 35.
23. Cocks K, King MT, Velikova G, Martyn St-James M, Fayers PM, Brown JM. Evidence-based guidelines for determination of sample size and interpretation of the European Organisation for the Research and Treatment of Cancer Quality of Life Questionnaire Core 30. *J Clin Oncol* 2011; **29**: 89-96.
24. Delforge M, Dhawan R, Robinson D, Jr., Meunier J, Regnault A, Esseltine DL, et al. Health-related quality of life in elderly, newly diagnosed multiple myeloma patients treated with VMP vs. MP: results from the VISTA trial. *Eur J Haematol* 2012; **89**: 16-27.
25. Kvam AK, Fayers P, Wisloff F. What changes in health-related quality of life matter to multiple myeloma patients? A prospective study. *Eur J Haematol* 2010; **84**: 345-53.
26. Kvam AK, Fayers PM, Wisloff F. Responsiveness and minimal important score differences in quality-of-life questionnaires: a comparison of the EORTC QLQ-C30 cancer-specific questionnaire to the generic utility questionnaires EQ-5D and 15D in patients with multiple myeloma. *Eur J Haematol* 2011; **87**: 330-7.
27. National Institute for Health and Care Excellence. Carfilzomib for previously treated multiple myeloma, 2017. Available from: https://www.nice.org.uk/guidance/TA457. Date accessed: 28 Jan 2020.
28. NHS improvement. National Schedule of Reference Costs 2017/18. Available from: https://improvement.nhs.uk/resources/reference-costs/. Date accessed: Oct 2019.
29. MIMS. Monthly Index of Medical Specialties (MIMS). Available at: https://www.mims.co.uk/.
30. Department of Health. Drug and pharmaceutical electronic market information tool (eMit) .
31. National Institute for Health and Care Excellence. Panobinostat for treating multiple myeloma after at least 2 previous treatments, 2016. Available from: https://www.nice.org.uk/guidance/TA380. Date accessed: 28 Jan 2020.
32. National Institute for Health and Care Excellence. Pomalidomide for multiple myeloma previously treated with lenalidomide and bortezomib, 2017. Available from: https://www.nice.org.uk/guidance/TA427. Date accessed: 28 Jan 2020.
33. National Institute for Health and Care Excellence. Lenalidomide for the treatment of multiple myeloma in people who have received at least 2 prior therapies, 2019. Available from: https://www.nice.org.uk/guidance/TA171. Date accessed: 28 Jan 2020.

34. National Institute for Health and Care Excellence (NICE). Guide to the methods of technology appraisal. 2013. Available from: https://www.nice.org.uk/process/pmg9/resources/guide-to-the-methods-of-technology-appraisal-2013-pdf-2007975843781. Date accessed: Feb 2019.
35. Expertise Ph. MyelomaToul Statistical Analysis Report. 2020.
36. National Institute for Health and Care Excellence. Daratumumab with bortezomib and dexamethasone for previously treated multiple myeloma, 2019. Available from: https://www.nice.org.uk/guidance/TA573. Date accessed: 28 Jan 2020.
37. National Institute for Health and Care Excellence. Ocriplasmin for treating vitreomacular traction, 2013. Available from: https://www.nice.org.uk/guidance/ta297. Date accessed: March 2020.
38. National Institute for Health and Care Excellence. Daratumumab monotherapy for treating relapsed and refractory multiple myeloma, 2018. Available from: https://www.nice.org.uk/guidance/TA510. Date accessed: 28 Jan 2020.
39. Amgen. Haematological Malignancy Research Network: Clinical Management and Outcomes in Multiple Myeloma. 2016.
40. Curtis L and Burns A. Unit Costs of Health and Social Care 2018. 2018. Available from: https://www.pssru.ac.uk/project-pages/unit-costs/unit-costs-2018/. Date accessed: February 2020.
41. Gonzalez-McQuire S YK, Flinois A, et al. Retrospective Chart Review Study to Evaluate the Cost of Care of Patients with Symptomatic Multiple Myeloma in the UK. *Poster presented at the ISPOR 21st Annual Meeting; May 21-25, 2016; Washington DC, United States* 2016.
42. Georghiou T, Bardsley M. Exploring the cost of care at the end of life. 2014.
43. National Institute for Health and Care Excellence. Pertuzumab with trastuzumab and docetaxel for treating HER2-positive breast cancer, 2019. Available from: https://www.nice.org.uk/guidance/ta509. Date accessed: March 2020.
44. International Myeloma Working Group. International Myeloma Working Group (IMWG) Uniform Response Criteria for Multiple Myeloma, 2020. Available from: https://www.myeloma.org/resource-library/international-myeloma-working-group-imwg-uniform-response-criteria-multiple. Date accessed: 06 Feb 2020.

9 Appendices

9.1 Baseline characteristics

Table 37. Baseline characteristics for the subgroup of people from ASPIRE who received carfilzomib 2L after bortezomib-based regimen and no lenalidomide (adapted from Table 3 provided as part of the company's response to clarification questions)

Characteristic	CRd	Rd
	(N = 74)	(N = 66)
Age group, n (%)		
• <65	[redacted]	[redacted]
• 65–74	[redacted]	[redacted]
• ≥75	[redacted]	[redacted]
ECOG performance status, n (%)		
• 0	[redacted]	[redacted]
• 1	[redacted]	[redacted]
• 2	[redacted]	[redacted]
Baseline creatinine clearance, n (%)		
• 30–<50 mL/min	[redacted]	[redacted]
• 50–<80 mL/min	[redacted]	[redacted]
• ≥80 mL/min	[redacted]	[redacted]
Time (months) since initial diagnosis		
• Mean (SD)	[redacted]	[redacted]
Time (months) since last relapse		
• Mean (SD)	[redacted]	[redacted]
Baseline ISS Stage, n (%)		
• Stage I	[redacted]	[redacted]
• Stage II	[redacted]	[redacted]
• Stage III	[redacted]	[redacted]
Baseline β2 microglobulin, n (%)		
• <3.5 mg/L	[redacted]	[redacted]
• ≥3.5 mg/L	[redacted]	[redacted]
Prior SCT, n (%)		
• Yes	[redacted]	[redacted]
• No	[redacted]	[redacted]
Prior therapy, n (%)		
• Bortezomib	[redacted]	[redacted]
• Lenalidomide	[redacted]	[redacted]
Refractory in any prior regimen, n (%)		
• Bortezomib	[redacted]	[redacted]

Abbreviations: C, carfilzomib; CS, company submission; d, dexamethasone; ECOG PS, Eastern Cooperative Oncology Group Performance Status; IMiD, immunomodulatory drug; ISS, International Staging System; ITT, intention-to-treat; max, maximum; min, minimum; NR, not reported; R, lenalidomide; SCT, stem cell transplantation; SD, standard deviation.

Table 38. Overview of baseline characteristics in ASPIRE (ITT population, adapted from CS, Table 7, page 27)

Characteristic	CRd (N = 396)	Rd (N = 396)	Total (N = 792)
Age, years, median (min, max)	64.0 (38.0, 87.0)	65.0 (31.0, 91.0)	64.0 (31.0, 91.0)
Female, n (%)	181 (45.7)	164 (41.4)	345 (43.6)
Race, n (%)			
• White	377 (95.2)	377 (95.2)	754 (95.2)
• Black	12 (3.0)	11 (2.8)	23 (2.9)
• Asian	1 (0.3)	3 (0.8)	4 (0.5)
• Native Hawaiian/Pacific Islander	[redacted]	[redacted]	[redacted]
• NR/other	6 (1.5)	4 (1.0)	10 (1.3)
Time since diagnosis, years, median (min, max)	3.0 (0.4, 19.7)[a]	3.2 (0.5, 27.3)	3.1 (0.4, 27.3)
Body surface area (m^2), mean (SD)	[redacted]	[redacted]	[redacted]
ECOG PS, n (%)			
• 0	165 (41.7)	175 (44.2)	340 (42.9)
• 1	191 (48.2)	186 (47.0)	377 (47.6)
• 2	40 (10.1)	35 (8.8)	75 (9.5)
ISS stage at diagnosis, n (%)			
• I	64 (16.2)	74 (18.7)	138 (17.4)
• II	99 (25.0)	94 (23.7)	193 (24.4)
• III	185 (46.7)	161 (40.7)	3 (43.7)
• Unknown	48 (12.1)	67 (16.9)	115 (14.5)
Calculated ISS stage at baseline, n (%)[b]			
• I	[redacted]	[redacted]	[redacted]
• II	[redacted]	[redacted]	[redacted]
• III	[redacted]	[redacted]	[redacted]
• Unknown	[redacted]	[redacted]	[redacted]
Cytogenetic risk (%)[c]			
• High	48 (12.1)	52 (13.1)	100 (12.6)
• Standard	147 (37.1)	170 (42.9)	317 (40.0)
• Unknown	201 (50.8)	174 (43.9)	375 (47.3)
Number of prior regimens			
• Median (min, max)	2.0 (1, 4)	2.0 (1, 4)	2.0 (1, 4)
• 1, n (%)	184 (46.5)	157 (39.6)	341 (43.1)
• 2, n (%)	[redacted]	[redacted]	[redacted]
• 3, n (%)	[redacted]	[redacted]	[redacted]
• 4, n (%)	[redacted]	[redacted]	[redacted]
Prior therapy received, n (%)			
• SCT	217 (54.8)	229 (57.8)	446 (56.3)
• Bortezomib	261 (65.9)	260 (65.7)	521 (65.8)
• Lenalidomide	79 (19.9)	78 (19.7)	157 (19.8)

• Thalidomide	[redacted]	[redacted]	[redacted]
• Pomalidomide	[redacted]	[redacted]	[redacted]
• Any IMiD[d]	233 (58.8)	229 (57.8)	462 (58.3)
• Bortezomib and IMiD	146 (36.9)	139 (35.1)	285 (36.0)
• Corticosteroids	[redacted]	[redacted]	[redacted]
• Anthracycline	[redacted]	[redacted]	[redacted]
• Alkylators	[redacted]	[redacted]	[redacted]
Received in last regimen, n (%)			
• Bortezomib	[redacted]	[redacted]	[redacted]
• Lenalidomide	[redacted]	[redacted]	[redacted]
Refractory to last regimen, n (%)	110 (27.8)	119 (30.1)	229 (28.9)

[a] N = 395 for this analysis.

[b] ISS sponsor-derived using central laboratory data for β2-microglobulin and local laboratory data for serum albumin.

[c] The high-risk group consisted of patients with the genetic subtypes t(4; 14), t(14;16), or deletion 17p in ≥ 60% of plasma cells. The standard-risk group consisted of patients without t(4; 14), t(14;16), and < 60% of plasma cells with deletion 17p. The unknown risk group included patients with FISH results that could not be analysed or from whom samples were not collected.

[d] Lenalidomide, thalidomide, or pomalidomide.

Abbreviations: C, carfilzomib; CS, company submission; d, dexamethasone; ECOG PS, Eastern Cooperative Oncology Group Performance Status; FISH, fluorescence in situ hybridisation; IMiD, immunomodulatory drug; ISS, International Staging System; ITT, intent-to-treat; max, maximum; min, minimum; NR, not reported; R, lenalidomide; SCT, stem cell transplantation; SD, standard deviation.

Table 39. Baseline characteristics for people receiving treatment at second line after one prior therapy with bortezomib (ASPIRE; adapted from Table 19 of Appendix E)

	CRd (N = 93)	Rd (N = 73)	Total (N = 166)
Age, years			
• Median (min, max)	[redacted]	[redacted]	[redacted]
• Mean (SD)	[redacted]	[redacted]	[redacted]
Age group N(%)			
• <65	[redacted]	[redacted]	[redacted]
• 65–74	[redacted]	[redacted]	[redacted]
• >=75	[redacted]	[redacted]	[redacted]
Female, n (%)	[redacted]	[redacted]	[redacted]
Race, n (%)			
• White	[redacted]	[redacted]	[redacted]
• Black	[redacted]	[redacted]	[redacted]
• Asian	[redacted]	[redacted]	[redacted]
• Native Hawaiian/Pacific Islander	[redacted]	[redacted]	[redacted]
• NR/other	[redacted]	[redacted]	[redacted]
Ethnicity, N(%)			
• Hispanic or Latino	[redacted]	[redacted]	[redacted]
• Not Hispanic or Latino	[redacted]	[redacted]	[redacted]
BMI			
• N(%) Missing	[redacted]	[redacted]	[redacted]
• Mean (SD)	[redacted]	[redacted]	[redacted]
• Median (Min, Max)	[redacted]	[redacted]	[redacted]

Time since diagnosis, years	[redacted]	[redacted]	[redacted]
• Median (min, max)	[redacted]	[redacted]	[redacted]
Time from last relapse, months			
• N	[redacted]	[redacted]	[redacted]
• Median (min, max)	[redacted]	[redacted]	[redacted]
Time from last regimen, months			
• N	[redacted]	[redacted]	[redacted]
• Median (min, max)	[redacted]	[redacted]	[redacted]
Body surface area (m^2)			
• N	[redacted]	[redacted]	[redacted]
• Mean (SD)	[redacted]	[redacted]	[redacted]
Body surface area (m^2)			
• N(%) Missing	[redacted]	[redacted]	[redacted]
• <=2.2	[redacted]	[redacted]	[redacted]
• >2.2	[redacted]	[redacted]	[redacted]
Region, N (%)			
• Missing	[redacted]	[redacted]	[redacted]
• Europe	[redacted]	[redacted]	[redacted]
• North America	[redacted]	[redacted]	[redacted]
• ROW	[redacted]	[redacted]	[redacted]
ECOG PS, n (%)			
• 0	[redacted]	[redacted]	[redacted]
• 1	[redacted]	[redacted]	[redacted]
• 2	[redacted]	[redacted]	[redacted]
ECOG PS, n (%)			
• 0	[redacted]	[redacted]	[redacted]
• 1–2	[redacted]	[redacted]	[redacted]
Baseline Hemoglobin N(%)			
• Median (min-max)	[redacted]	[redacted]	[redacted]
• <105 g/L	[redacted]	[redacted]	[redacted]
• >=105 g/L	[redacted]	[redacted]	[redacted]
Absolute Neutrophil count N(%)			
• Median (min-max)	[redacted]	[redacted]	[redacted]
• <1.5 g/L	[redacted]	[redacted]	[redacted]
• >=1.5 g/L	[redacted]	[redacted]	[redacted]
Platelet count (10^9/L), N(%)			
• Median (min-max)	[redacted]	[redacted]	[redacted]
• <150	[redacted]	[redacted]	[redacted]
• >=150	[redacted]	[redacted]	[redacted]
Corrected Calcium (mg/dl), N(%)			
• N	[redacted]	[redacted]	[redacted]
• Median (min-max)	[redacted]	[redacted]	[redacted]
• <=11.5	[redacted]	[redacted]	[redacted]
• >11.5	[redacted]	[redacted]	[redacted]

• Missing	[redacted]	[redacted]	[redacted]
Serum Creatinine (umol/L)			
• Mean (SD)	[redacted]	[redacted]	[redacted]
• Median (Min, Max)	[redacted]	[redacted]	[redacted]
Creatinine Clearance Sponsor Calculated (mL/min)			
• N	[redacted]	[redacted]	[redacted]
• Median (Min, Max)	[redacted]	[redacted]	[redacted]
• <30	[redacted]	[redacted]	[redacted]
• 30–<50	[redacted]	[redacted]	[redacted]
• 50–<80	[redacted]	[redacted]	[redacted]
• ≥80	[redacted]	[redacted]	[redacted]
• Missing	[redacted]	[redacted]	[redacted]
Creatinine Clearance Reported (mL/min)			
• N	[redacted]	[redacted]	[redacted]
• Median (Min, Max)	[redacted]	[redacted]	[redacted]
• <30	[redacted]	[redacted]	[redacted]
• 30–<50	[redacted]	[redacted]	[redacted]
• 50–<80	[redacted]	[redacted]	[redacted]
• ≥80	[redacted]	[redacted]	[redacted]
ISS stage at diagnosis, n (%)			
• I	[redacted]	[redacted]	[redacted]
• II	[redacted]	[redacted]	[redacted]
• III	[redacted]	[redacted]	[redacted]
• Unknown	[redacted]	[redacted]	[redacted]
Calculated ISS stage at baseline, n (%)[a]			
• I	[redacted]	[redacted]	[redacted]
• II	[redacted]	[redacted]	[redacted]
• III	[redacted]	[redacted]	[redacted]
• Unknown	[redacted]	[redacted]	[redacted]
Measurable disease category at baseline N(%)			
• SPEP Only	[redacted]	[redacted]	[redacted]
• SPEP and UPEP	[redacted]	[redacted]	[redacted]
• UPEP Only	[redacted]	[redacted]	[redacted]
M-protein heavy chain isotype N(%)			
• IGA	[redacted]	[redacted]	[redacted]
• IGG	[redacted]	[redacted]	[redacted]
• IGD	[redacted]	[redacted]	[redacted]
• NOT DETECTED	[redacted]	[redacted]	[redacted]
M-protein light chain isotype N(%)			
• KAPPA	[redacted]	[redacted]	[redacted]
• LAMBDA	[redacted]	[redacted]	[redacted]

Baseline Beta 2 Microglobulin Level N(%)			
• <3.5	[redacted]	[redacted]	[redacted]
• >=3.5	[redacted]	[redacted]	[redacted]
• Missing	[redacted]	[redacted]	[redacted]
Baseline Beta 2 Microglobulin Level per Covance N(%)			
• <2.5	[redacted]	[redacted]	[redacted]
• >=2.5	[redacted]	[redacted]	[redacted]
• Missing	[redacted]	[redacted]	[redacted]
Presence of plasmacytoma N(%)			
• N(%) Missing	[redacted]	[redacted]	[redacted]
• Yes	[redacted]	[redacted]	[redacted]
• No	[redacted]	[redacted]	[redacted]
Presence of bone lesion N(%)			
• N(%) Missing	[redacted]	[redacted]	[redacted]
• Yes	[redacted]	[redacted]	[redacted]
• No	[redacted]	[redacted]	[redacted]
Cytogenetic risk (%)[b]			
• High	[redacted]	[redacted]	[redacted]
• Standard	[redacted]	[redacted]	[redacted]
• Unknown	[redacted]	[redacted]	[redacted]
Baseline Albumin (g/L)			
• N(%) Missing	[redacted]	[redacted]	[redacted]
• Mean (SD)	[redacted]	[redacted]	[redacted]
• Median (Min, Max)	[redacted]	[redacted]	[redacted]
Prior surgery for multiple myeloma N(%)			
• Yes	[redacted]	[redacted]	[redacted]
• No	[redacted]	[redacted]	[redacted]
Prior radiotherapy for multiple myeloma N(%)			
• Yes	[redacted]	[redacted]	[redacted]
• No	[redacted]	[redacted]	[redacted]
Prior hematopoietic cell transplant N(%)			
• Yes	[redacted]	[redacted]	[redacted]
• No	[redacted]	[redacted]	[redacted]
Prior therapy received, n (%)			
• SCT	[redacted]	[redacted]	[redacted]
• Bortezomib	[redacted]	[redacted]	[redacted]
• Lenalidomide	[redacted]	[redacted]	[redacted]
• Thalidomide	[redacted]	[redacted]	[redacted]
• Pomalidomide	[redacted]	[redacted]	[redacted]
• Any IMiD[c]	[redacted]	[redacted]	[redacted]
• Bortezomib and IMiD	[redacted]	[redacted]	[redacted]

Received in last regimen, n (%)			
• Bortezomib	[redacted]	[redacted]	[redacted]
• Lenalidomide	[redacted]	[redacted]	[redacted]
Refractory in Any Prior Regimen N(%)			
• Bortezomib	[redacted]	[redacted]	[redacted]
• Lenalidomide	[redacted]	[redacted]	[redacted]
• Bortezomib and IMiD	[redacted]	[redacted]	[redacted]
• Thalidomide	[redacted]	[redacted]	[redacted]
• Refractory to last regimen, n (%)	[redacted]	[redacted]	[redacted]
History of neuropathy N(%)			
• N(%) Missing			
• Yes	[redacted]	[redacted]	[redacted]
• No	[redacted]	[redacted]	[redacted]
Best response to last prior line regimen N(%)			
• Unknown	[redacted]	[redacted]	[redacted]
• Complete Response	[redacted]	[redacted]	[redacted]
• Partial Response	[redacted]	[redacted]	[redacted]
• Minimal Response	[redacted]	[redacted]	[redacted]
• Stable Disease	[redacted]	[redacted]	[redacted]
• Progressive Disease	[redacted]	[redacted]	[redacted]

[a] ISS sponsor-derived using central laboratory data for β2-microglobulin and local laboratory data for serum albumin

[b] The high-risk group consisted of patients with the genetic subtypes t(4; 14), t(14;16), or deletion 17p in ≥ 60% of plasma cells. The standard-risk group consisted of patients without t(4; 14), t(14;16), and < 60% of plasma cells with deletion 17p. The unknown risk group included patients with FISH results that could not be analysed or from whom samples were not collected.

[c] Lenalidomide, thalidomide, or pomalidomide.

Abbreviations: CRd, carfilzomib/lenalidomide/dexamethasone; ECOG PS, Eastern Cooperative Oncology Group Performance Status; IMiD, immunomodulatory drug; ISS, International Staging System; ITT, intent-to-treat; max, maximum; min, minimum; NR, not reported; Rd, lenalidomide/dexamethasone; SCT, stem cell transplantation

9.2 Overview of overall response rate and time to next treatment

9.2.1 Overall response rate

Primary analysis of overall response rate (ORR) was defined as achieving a partial response (PR) or better, and was based on classifications of response to treatment as evaluated by the IRC. Level of response was categorised as per criteria set out by the IMWG-URC (International Myeloma Working Group-Uniform Response Criteria),[44] with the exception of minimal response, which was based on European Group for Blood and Marrow Transplant (EBMT) criteria.

For the ITT population of ASPIRE, CRd was associated with a statistically significant higher ORR compared with Rd, with [redacted] ([redacted]/396) of people in the CRd group achieving a best response of at least a PR versus [redacted] ([redacted]/396) with Rd (odds ratio [redacted]; 95% confidence interval [CI]: [redacted]; p <0.0001; Table 40). Median time to response was 1 month for both CRd and Rd,

but a difference in mean time to response was noted (1.6 months with CRd vs 2.3 months for Rd; Table 40).

Table 40. Overall response rate as determined by the IRC for the ASPIRE ITT population (adapted from CS, Table 14, page 41)

Response	CRd (N = 396)	Rd (N = 396)
Best response[a]		
≥CR	126 (31.8)	37 (9.3)
• Stringent CR	56 (14.1)	17 (4.3)
• CR	70 (17.7)	20 (5.1)
≥VGPR	277 (69.9)	160 (40.4)
• VGPR	[redacted]	[redacted]
PR	[redacted]	[redacted]
Minimal response	[redacted]	[redacted]
Stable disease	[redacted]	[redacted]
Progressive disease	[redacted]	[redacted]
Not evaluable	[redacted]	[redacted]
ORR, n (%)[b] (95% CI of ORR)	345 (87.1) (83.4 to 90.3)	264 (66.7) (61.8 to 71.3)
p-value (one-sided)[c]	<0.0001[d,e]	
OR (95% CI)	[redacted]	
Time to response		
• Mean, months (SD)	1.6 (1.39)	2.3 (2.42)
• Median, months	1	1

[a] Best response was defined as a patient's best response during the study.
[b] Defined as patients who had a best response of sCR, CR, VGPR, or PR.
[c] Unadjusted p-value from Cochran–Mantel–Haenszel chi-square test with β2-microglobulin levels (<2.5 mg/L vs ≥2.5 mg/L), prior bortezomib (no vs yes), and prior lenalidomide (no vs yes) as stratification factors.
[d] p-value is statistically significant (per hierarchical testing strategy described in Siegel *et al.* 2018.
[e] Reported as a two-sided p-value (p <0.0001) in Stewart *et al.* 2015.
Abbreviations: C, carfilzomib; CI, confidence interval; CR, complete response; CS, company submission; d, dexamethasone; ITT, intention to treat; OR, odds ratio; ORR, overall response rate; PR, partial response; R, lenalidomide; SD, standard deviation; VGPR, very good partial response.

9.2.2 Time to next treatment

TTNT was defined as the median time from randomisation to commencement of a new anti-myeloma treatment. At the time of the interim analysis (June 2014) presented in TA457,[8] CRd was associated with a statistically significantly longer TTNT than Rd, with median TTNT of 17.3 months and 12.1 months, respectively (hazard ratio [HR] 0.63; 95% CI: 0.50 to 0.78; p <0.0001; Table 41). The benefit in TTNT reported for CRd was maintained at a later data cut-off (April 2017), with TTNT of [redacted] months reported for CRd compared with [redacted] months for RD (HR 0.65; 95% CI: 0.53 to 0.79; p <0.0001).

Table 41. Time to next treatment for the ASPIRE ITT population (adapted from ERG report for TA457[8] and CS, Table 15, page 42)

	Interim analysis (data cut-off 16 June 2014)		Primary OS analysis (data cut off 28 April 2017)	
	CRd (N = 396)	**Rd (N = 396)**	**CRd (N = 396)**	**Rd (N = 396)**
Participants who started next treatment, n (%)	151 (38.1)	184 (46.5)	182 (46.0)	211 (53.3)
Time to next treatment, median months (min, max)	17.3 (0.46 to 37.6)	12.1 (0.26 to 33.5)	[redacted]	[redacted]
K–M estimate of time to next treatment, median months (95% CI)	37.6 (31.8 to NE)	24.5 (20.8 to 32.8)	39.0 (31.8 to 55.1)	24.4 (20.8 to 28.4)
Hazard ratio CRd:Rd (95% CI)	0.63 (0.50 to 0.78)		0.65 (0.53 to 0.79)	
Descriptive p-value (1-sided)[a]	<0.0001		<0.0001	
Median follow-up for time to next treatment, months (95% CI)	31.5 (30.7 to 32.0)	30.0 (29.3 to 31.2)	[redacted]	[redacted]

[a] Unadjusted p-value is from a stratified log-rank test with β2-microglobulin levels (<2.5 mg/L vs ≥2.5 mg/L), prior bortezomib (no vs yes), and prior lenalidomide (no vs yes) as stratification factors. P-value is for descriptive purposes only.

Abbreviations: C, carfilzomib; CI, confidence interval; d, dexamethasone; ITT, intention to treat; K–M, Kaplan–Meier; NE, not estimable; OS, overall survival; R, lenalidomide.