



Journal of Health Economics and Outcomes Research

Oncology

Use of Minimal Residual Disease Status to Reduce Uncertainty in Estimating Long-term Survival Outcomes for Newly Diagnosed Multiple Myeloma Patients

Naomi van Hest^{1*}, Peter Morten¹, Keith Stubbs², Nicola Trevor²

¹Costello Medical Consulting, Cambridge, United Kingdom

²Janssen-Cilag Ltd, High Wycombe, United Kingdom

ARTICLE INFORMATION

Article history:

Received August 23, 2022

Received in revised form November 21, 2022

Accepted November 23, 2022

Keywords: multiple myeloma, minimal residual disease, PSM, landmark analysis, CASSIOPEIA, daratumumab, BDt, uncertainty

*Corresponding author:

Email address:

naomi.vanhest@costellomedical.com

➤ [Supplementary Material](#)

ABSTRACT

Background: Demonstrating the cost-effectiveness of new treatments for multiple myeloma (MM) often relies on the extrapolation of overall survival (OS) trial data. This method can introduce uncertainty in long-term survival estimates if OS data are immature, as is often the case in newly diagnosed MM (NDMM). We explore the use of the relationship between minimal residual disease (MRD) status and OS to reduce uncertainty of long-term survival outcomes.

Objectives: To evaluate if uncertainty in long-term modeled outcomes in NDMM is reduced using a response-based partitioned survival model (PSM), whereby patients were categorized as MRD-positive or -negative, relative to a standard PSM, when OS data are immature.

Methods: Standard and response-based PSMs, estimating patient life-years (LYs) over a lifetime horizon, were developed for NDMM patients treated with bortezomib, thalidomide, and dexamethasone (BDt) with or without daratumumab as induction and consolidation therapy. In the standard PSM, LYs were determined by extrapolations from individual patient data from CASSIOPEIA. In the response-based PSM, survival was dependent on MRD status at the time of the response assessment via a landmark analysis. Cox-proportional hazard ratios from external sources and CASSIOPEIA informed the relationship for OS between MRD-positive and MRD-negative, and between patients receiving BDt and daratumumab plus BDt, respectively. Uncertainty was assessed by comparing LYs and OS extrapolations from deterministic and probabilistic analyses.

Results: This response-based PSM demonstrated reduced uncertainty in long-term survival outcomes compared with the standard PSM (range across extrapolations of 3.4 and 7.7 LYs for daratumumab plus BDt and BDt, respectively, vs 14.8 and 11.8 LYs for the standard PSM). It also estimated a narrower interquartile range of LYs in the probabilistic analyses for the majority of parametric extrapolations.

Discussion: Alternative methods to estimate long-term survival outcomes, such as a response-based PSM, can reduce uncertainty in modeling predictions around cost-effectiveness estimates for HTA bodies and payers, thereby supporting faster market access for novel therapies with immature survival data.

Conclusions: Use of MRD status in a response-based PSM reduces uncertainty in modeling long-term survival in patients with NDMM and provides a greater number of clinically plausible extrapolations compared with a standard PSM.

BACKGROUND

Multiple myeloma (MM) is a rare and incurable blood cancer with orphan disease designation in both the United States and Europe.^{1,2} The disease follows a relapsing-remitting course with almost all surviving MM patients eventually relapsing or becoming refractory to ex-

isting treatment options.³⁻⁶ With each relapse, it becomes increasingly difficult to induce deep and durable responses, leading to increased attrition rates.^{7,8} Therefore, it is important to use the most effective treatments as early as possible, as patients may not survive or be fit enough to receive treatment at later lines of therapy.



This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CCBY-4.0). View this license's legal deed at <http://creativecommons.org/licenses/by/4.0> and legal code at <http://creativecommons.org/licenses/by/4.0/legalcode> for more information.

For patients with newly diagnosed MM (NDMM), induction therapy followed by autologous stem cell transplant (ASCT) with or without consolidation therapy represents the standard of care for patients fit enough to receive these interventions.^{9,10} Transplant-eligible NDMM patients have the best chance of achieving long-term sustained remission compared with other patients with MM.^{11,12} Despite this, survival outcomes in this population remain substantially lower than in the general population.¹⁷ This reflects, in part, the considerable heterogeneity in this patient population, with some patients relapsing early or not achieving a response to treatment, resulting in comparatively poorer outcomes.^{13,14} Therefore, there is still a substantial unmet need for more effective, well-tolerated treatment options in the front-line setting.

Overall survival (OS) and progression-free survival (PFS) are the standard, established endpoints used by regulatory agencies in their assessment of novel therapies for NDMM. However, for emerging therapies in earlier lines of treatment, the time frame needed to demonstrate statistically significant or clinically meaningful improvements in PFS alone in clinical trials is over 5 years.¹⁵

Long-term estimates of OS are needed as part of cost-effectiveness analyses, which are typically assessed over a lifelong time horizon. Therefore, cost-effectiveness models to inform these analyses, such as PSMs, rely on the extrapolation of survival data from clinical trials. Methods to extrapolate survival data are available, such as those described in Technical Support Document 14, published by the National Institute of Health and Care Excellence (NICE) Decision Support Unit (DSU).¹⁶ However, considerable uncertainty in long-term survival estimates is introduced if survival data are immature, which can hamper health technology assessment (HTA) decision-making. Recognizing the limitations of standard methods, the NICE DSU has published Technical Support Document 21, which details more flexible methods (eg, landmark analyses, among others) to model lifetime effects when hazard functions are complex.¹⁷ To effectively utilize this landmark method, a reliable marker of survival in NDMM to reduce uncertainty of long-term survival outcomes is required.

Value of Minimal Residual Disease Status in Multiple Myeloma

Relapse in MM is due to cancerous cells that resist treatment and undergo clonal expansion and evolution, resulting in tumor repopulation. These remaining cells that contribute to relapse are known as minimal residual disease (MRD).^{18,19} The state of MRD negativity is one where no remaining clonal or subclonal cancerous cells can be detected using currently available measurement techniques. Patients who achieve this state have been shown to be less likely to relapse, with long-term disease control achieved for some patients. In recent years, the International Myeloma Working Group has recognized MRD as the most sensitive measure of response assessment for MM.^{18,19}

The significance of MRD status and its relationship with long-term survival outcomes in patients with NDMM have been established in the literature through systematic literature reviews (SLRs) and meta-analyses.²⁰⁻²² The results of these studies have established the clinical utility of MRD assessment as an important marker of survival and could therefore be considered when estimating long-term outcomes when survival data are immature. Compared with other outcomes, MRD has also been shown to be superior and the most relevant predictor of clinical outcomes in MM via multivariate analyses.^{19,21,23-26} Furthermore, MRD status was recently used as the primary endpoint in the ongoing CEPHEUS trial for NDMM.²⁷ Recently, the European Medicines Agency considered there to be uncertainty in the quantitative, predictive association between MRD negativity and clinical endpoints like OS and PFS. However, it should be noted that assessment was made considering the use of MRD status as a primary endpoint

in clinical trials to be used for regulatory decision-making. The correlation between MRD status and survival has been well established in the literature and the European Medicines Agency acknowledged this correlation at the patient level.²⁸

A prior study by Yamamoto *et al* incorporated MRD status as a predictive marker in a cost-effectiveness analysis of daratumumab in NDMM used a Markov model structure.²⁹ The current study differs from the approach of Yamamoto *et al* as it incorporates external data, whereas the previous study relied only on trial data. Secondly, this study provides a comparison between a standard partitioned survival model (PSM) and a PSM incorporating MRD status to assess the impact on uncertainty. The previous study presented cost-effectiveness results alone and did not provide this comparison.

Introduction to Daratumumab

Daratumumab is a first-in-class, fully human IgG1κ monoclonal antibody that binds to CD38, a cell-surface glycoprotein found on many immune cells, including white blood cells.^{30,31} Among others, daratumumab is indicated in combination with bortezomib, thalidomide, and dexamethasone (DBTd) for the treatment of adult patients with NDMM who are eligible for ASCT.²²

Evidence of the clinical effectiveness of DBTd was demonstrated in CASSIOPEIA, a pivotal randomized, open-label, active-controlled, parallel-group, multicenter phase 3 study.³² The trial included 1085 patients, with a median follow-up of 29.2 months. The treatment phase was split into 2 parts; only data from part 1 were used in this study, which reflects the current approved use of daratumumab in this indication. In part 1, patients were randomized to receive four 28-day cycles of induction therapy with DBTd or bortezomib, thalidomide, and dexamethasone (BTd) prior to ASCT, followed by two 28-day cycles of consolidation therapy with DBTd or BTd. The primary endpoint in part 1 was the proportion of patients achieving stringent complete response. Key secondary endpoints included PFS, OS, and rate of MRD negativity. Minimal residual disease negativity was measured in CASSIOPEIA at 100 days post-ASCT, to assess efficacy following induction, ASCT and consolidation therapy.³² As measured at postconsolidation, a significantly higher proportion of patients were MRD-negative (MRD-) measured using multiparametric flow cytometry when treated with DBTd compared with BTd alone at a threshold of 1 tumor cell per 10⁻⁵ white cells (64% vs 44%; $P < .0001$).³² Survival data from CASSIOPEIA are, however, considered immature as only a small proportion of patients experienced an OS event in the BTd and DBTd arms at a median follow-up of 29.2 months (8.9% and 4.8%, respectively).³³

Study Objectives

The objective of this study was to assess the uncertainty in modeled outcomes associated with the use of a response-based PSM, relative to a standard PSM, when survival data are immature. Using this approach, the study also aimed to evaluate the suitability of MRD status as a marker for modeling long-term survival outcomes in NDMM by assessing the clinical plausibility of results from each modeling approach.

METHODS

Setting

A de novo PSM, which estimated patient life-years (LYs), was developed in Microsoft Excel for patients with NDMM who are eligible for ASCT. To reflect the CASSIOPEIA trial, patients were modeled to receive induction therapy, with either DBTd or BTd, followed by ASCT and consolidation therapy (as per their induction therapy).

Model Structure

In line with guidance from NICE DSU 19, both a PSM and state transition were developed initially. The PSM structure was deemed more appropriate as it allows intuitive incorporation of PFS and OS data from the CASSIOPEIA trial, and preliminary analyses indicated results would be comparable. Therefore, a PSM structure was pursued further.³⁴

A standard PSM and response-based PSM were independently explored. In the standard PSM, the occupancy of health states over time was derived from survival curves directly extrapolated from individual patient data (IPD). In the response-based PSM, long-term outcomes were modeled to be dependent on whether patients had achieved a response to treatment, with “responders” defined as patients who were MRD- at a particular time point. Compared with the standard PSM, this approach explicitly acknowledges, and can better account for, underlying patient heterogeneity between responders and nonresponders.

To mitigate against the effect of immortal time bias (ie, the need for patients to live long enough to experience an event), a “landmark” approach was taken in the response-based PSM in line with NICE DSU guidance.¹⁷ The selected landmark time point was 100 days post-ASCT, to align with the timing of the MRD response assessment in the CASSIOPEIA trial.³² This time point is frequently chosen in MM trials because it is considered to be sufficiently long to capture the benefit of consolidation while being short enough to ensure few events were censored.^{21,35} Survival after this time point was modeled depending on MRD status at the time of the response assessment, with separate survival inputs used for patients with MRD- and MRD-positive (MRD+) status. Prior to the landmark time point, patients were modeled based on the observed survival data from CASSIOPEIA.³²

The following approach was taken for modeling post-landmark outcomes in the response-based PSM (**Figure 1**). For the BTd MRD+ curve, IPD from CASSIOPEIA for this cohort of patients were extrapolated. Of all cohorts of interest, this was the most mature data set from CASSIOPEIA and was used as a reference curve. For

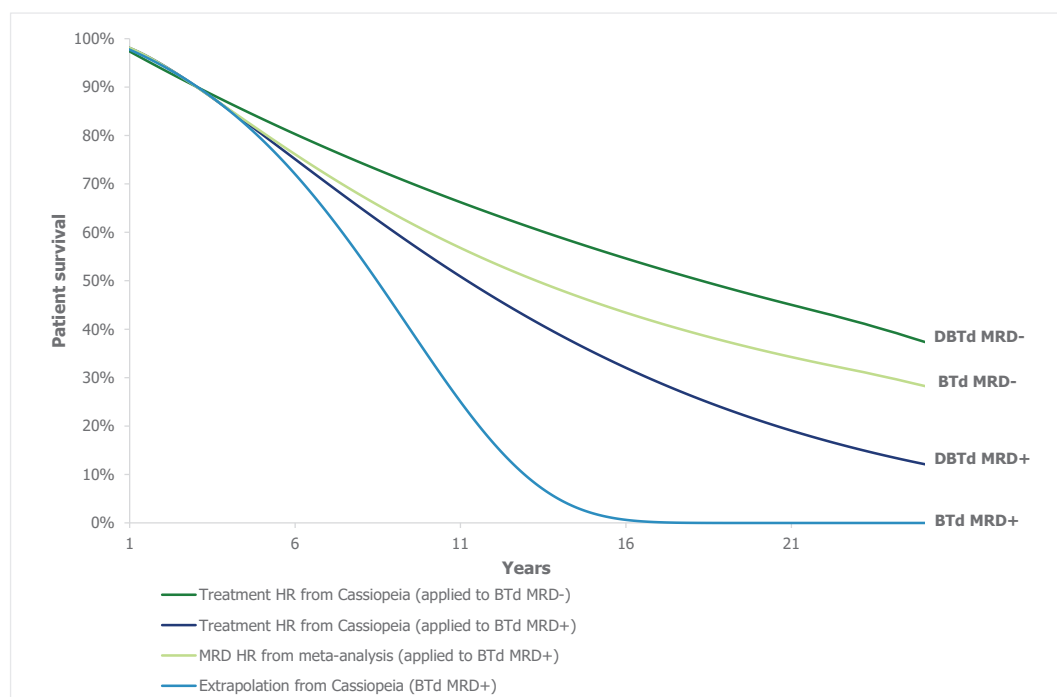
the BTd MRD- curve, a Cox-proportional hazard ratio (HR) for MRD- relative to MRD+ patients was applied to the BTd MRD+ survival curve, utilizing long-term data for NDMM transplant-eligible patients from the published SLRs and meta-analyses that assessed the relationship between MRD status and survival outcomes.²⁰⁻²² For the DBTd MRD+ and DBTd MRD- curves, MRD-positive- and MRD-negative-specific Cox-proportional HRs for DBTd vs BTd were applied to the corresponding BTd survival curves, utilizing information on the DBTd treatment effect derived from the landmark analysis of the CASSIOPEIA trial. The resultant extrapolations for each treatment (**Figure 2**) were validated against the observed Kaplan-Meier data and other published data (**Table 1**). The sources of these data and inputs are described below.

Clinical Inputs

Survival curves for DBTd and BTd were based on IPD from the intention-to-treat population of the CASSIOPEIA trial.³² Where appropriate, survival extrapolations were conducted in line with NICE DSU guidance,¹⁶ considering the full range of recommended parametric distributions (exponential, Weibull, log logistic, log normal, Gompertz, and generalized gamma). The appropriateness of modeling survival via the use of HRs was dependent on the assumption of proportional hazards. This was assessed and confirmed for all cases via inspection of log-cumulative hazard Schoenfeld residual plots, as per NICE DSU guidance (**Supplementary Figure S1**).¹⁶

The HR between MRD- and MRD+ applied to the reference curve was based on an SLR/meta-analysis conducted by Munshi *et al*, which identified 143 publications reporting survival outcomes in MM by MRD status.²⁰⁻²² The results of the SLR were subsequently adapted to limit the analysis to patients in whom transplant was performed and MRD was assessed at 100 days post-SCT, and who were not treated with DBTd. Therefore, the DBTd arm of the CASSIOPEIA trial was excluded. This ensured that the included HRs aligned to the patient

Figure 1. Approach to Modeling Survival by Response and Treatment Arm for the Response-based PSM



Survival presented is illustrative only and not expected to reflect the absolute or relative survival of these treatment and response curves.

Abbreviations: BTd, bortezomib, thalidomide and dexamethasone; DBTd, daratumumab, bortezomib, thalidomide and dexamethasone; HR, hazard ratio; MRD+/-, minimal residual disease positive/negative; PSM, partitioned survival model.

population in the model. As a result, 7 OS HRs informed the value implemented in the model. The clinical inputs used in the response-based PSM are shown in **Table 1**.

The risk of mortality for transplant-eligible patients with NDMM is expected to be higher than those of the general population when matched for age and gender. Therefore, to ensure that the extrapolated OS did not exceed that of the general population, age- and gender-matched general population mortality (GPM) in England was used in any cycle where the predicted probability of death was lower than GPM.

Model Outcomes and Validation

The model estimated LYs for the total patient population over a lifetime horizon (43 years, based on a mean age of 57 years from CASSIOPEIA and assuming all patients were dead at 100 years). Long-term OS extrapolations were assessed using a combination of statistical goodness-of-fit criteria (see **Supplementary Table S1**), visual inspection, and clinical expert opinion.¹⁶

For the base case analysis, the choice of extrapolation was informed by feedback from UK clinicians on the long-term survival outcomes expected in clinical practice. In the standard PSM, a Weibull and log logistic curve was selected for BTd and DBTd, respectively. For the response-based PSM, an exponential curve was selected for BTd MRD+.

The survival outcomes predicted by the response-based PSM, which incorporated both MRD+ and MRD- patients, and the standard PSM were then compared against: (1) observed data from the CASSIOPEIA trial (DBTd and BTd)³²; (2) observed data from the GIMEMA trial,³⁶ a phase 3 randomized controlled trial

(n = 474; median follow-up, 10.3 years), which compared BTd with thalidomide and dexamethasone as induction and consolidation therapy after double ASCT for NDMM; and (3) data from National Health Service (NHS) Digital datasets on real-world survival outcomes for 1049 patients in the United Kingdom who received first-line BTd.³⁷

Evaluation of Model Uncertainty

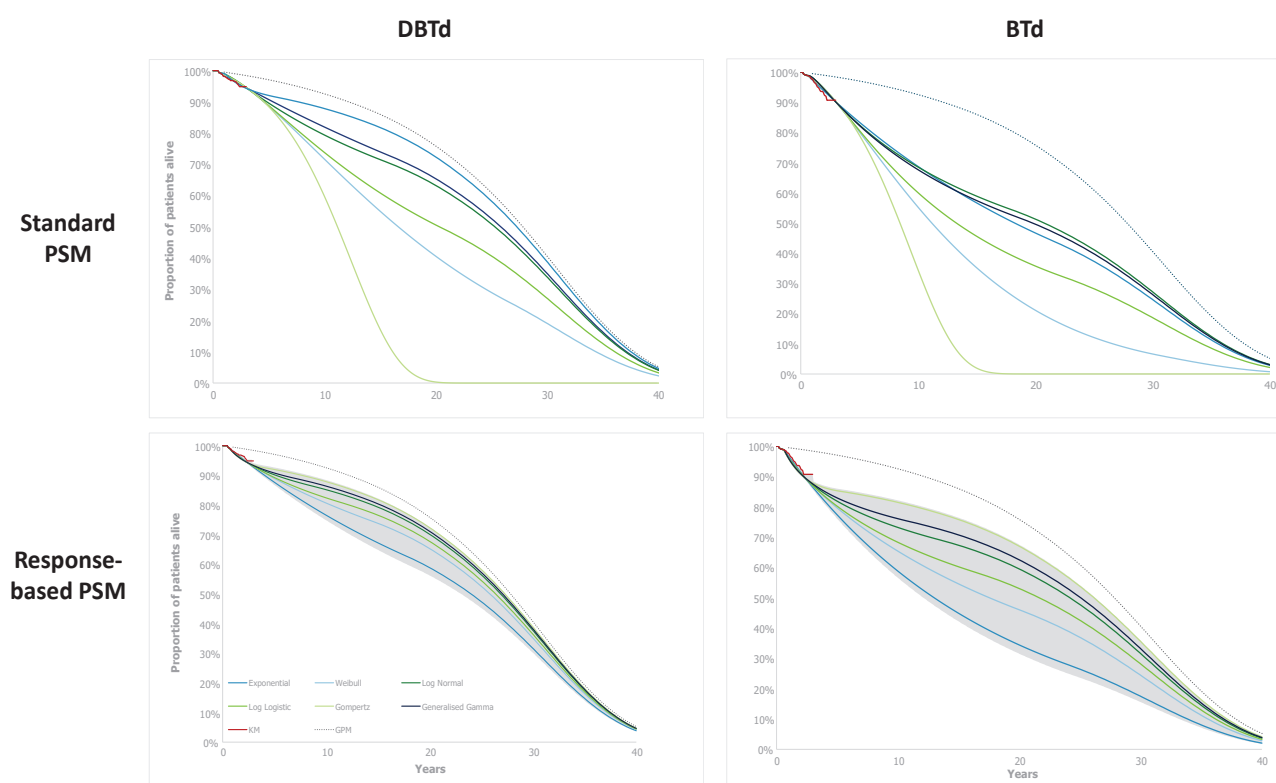
To evaluate uncertainty in model estimates, mean and median LYs for each of the 6 parametric extrapolations were evaluated. To explore uncertainty within each parametric extrapolation, probabilistic sensitivity analyses (PSAs) were conducted. A total of 2000 simulations were conducted, based on the number of simulations required for stabilization of the model results, per parametric extrapolation and model structure, in which the individual parameters informing each extrapolation were varied according to a normal distribution (Cholesky decomposition). In the response-based PSM, the HRs (log normal distribution), and the MRD- rate (beta distribution) were also sampled.

RESULTS

External Validation

With the selected base case extrapolations, patients were estimated to live an average of 14 and 21 years for BTd and DBTd, respectively, in the standard PSM and 16 and 22 years, respectively, in the response-based PSM (**Table 2**). Comparison of the base case LYs estimated by both models to the observed data demonstrated that both generate clinically plausible results: the plausibility of model estimates was supported both in the short- and longer-term by comparing with

Figure 2. Long-term Modeled Extrapolations by Treatment and Model Structure



For the response-based PSM, curve choices refer to the extrapolation of MRD+ BTd patients; HRs are then applied to this extrapolation to generate survival estimates for the full BTd cohort and the DBTd cohort. Areas shaded grey show where the curves sit when the HRs are varied across the 95% CIs.

Abbreviations: BTd, bortezomib, thalidomide and dexamethasone; DBTd, daratumumab, bortezomib, thalidomide and dexamethasone; GPM, general population mortality; HR, hazard ratio; PSM, partitioned survival model.

Table 1. Clinical Inputs for the Response-Based PSM Model

Model Input	Parameter Value	Source	Model Application
MRD-negative patients, % (95% CI)^a			
DBTd	63.7 (59.5–67.8) ^b	CASSIOPEIA (ITT population)	To determine proportion of patients achieving MRD negativity
BTd	43.5 (39.3–47.8) ^b	CASSIOPEIA (ITT population)	
HR applied to OS parametric extrapolation, HR (95% CI)			
MRD- vs MRD+	0.60 (0.50–0.72) ^c	Analysis conducted on Munshi et al ²⁰	To estimate survival outcomes for MRD- BTd patients (applied to BTd MRD+)
DBTd MRD+ vs BTd MRD+	0.73 (0.40–1.35)	CASSIOPEIA (landmark analysis)	To estimate survival outcomes for MRD+ DBTd patients (applied to BTd MRD+)
DBTd MRD- vs BTd MRD-	0.41 (0.16–1.04)	CASSIOPEIA (landmark analysis)	To estimate survival outcomes for MRD- DBTd patients (applied to BTd MRD-)

Abbreviations: BTd, bortezomib, thalidomide and dexamethasone; CI, confidence interval; DBTd, daratumumab, bortezomib, thalidomide and dexamethasone; HR, hazard ratio; ITT, intention to treat; MRD+/-, minimal residual disease positive/negative; OS, overall survival.

^aThe proportion of patients achieving postconsolidation MRD negativity in the model was based on that in CASSIOPEIA, with a sensitivity threshold of 10⁻⁵ for MRD.

^bDBTd vs BTd: *P* < .0001.

^cThe HR used in the model differs slightly from the HR reported in Munshi et al²⁰ due to the exclusion of the DBTd arm of the CASSIOPEIA trial and studies measuring MRD at 100-days post-autologous stem cell transplant from the analysis.

Table 2. Estimated Mean Life-Years by Treatment and Model Structure

Extrapolation Curve	Standard PSM (LYs)				Response-Based PSM (LYs) ^a			
	DBTd		BTd		DBTd		BTd	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median
Exponential	24.18	26.52	19.60	18.77	22.49	24.68	16.36	13.31
Weibull	18.78	17.02	13.58	11.46	23.98	26.52	19.01	18.08
Log normal	23.58	25.99	20.31	21.38	25.21	27.60	22.00	24.92
Log logistic	20.97	20.85	16.94	13.62	24.56	27.06	20.43	22.31
Gompertz	11.09	11.42	8.48	8.69	25.87	27.98	24.04	26.92
Generalized gamma	25.90	27.98	19.95	20.38	25.50	27.75	22.78	25.85
Range^b	14.81	16.56	11.84	12.69	3.37	3.30	7.68	13.62

Abbreviations: BTd, bortezomib, thalidomide, and dexamethasone; DBTd, daratumumab, bortezomib, thalidomide and dexamethasone; LY, life-year; MRD, minimal residual disease; NHS, National Health Service; PSM, partitioned survival model.

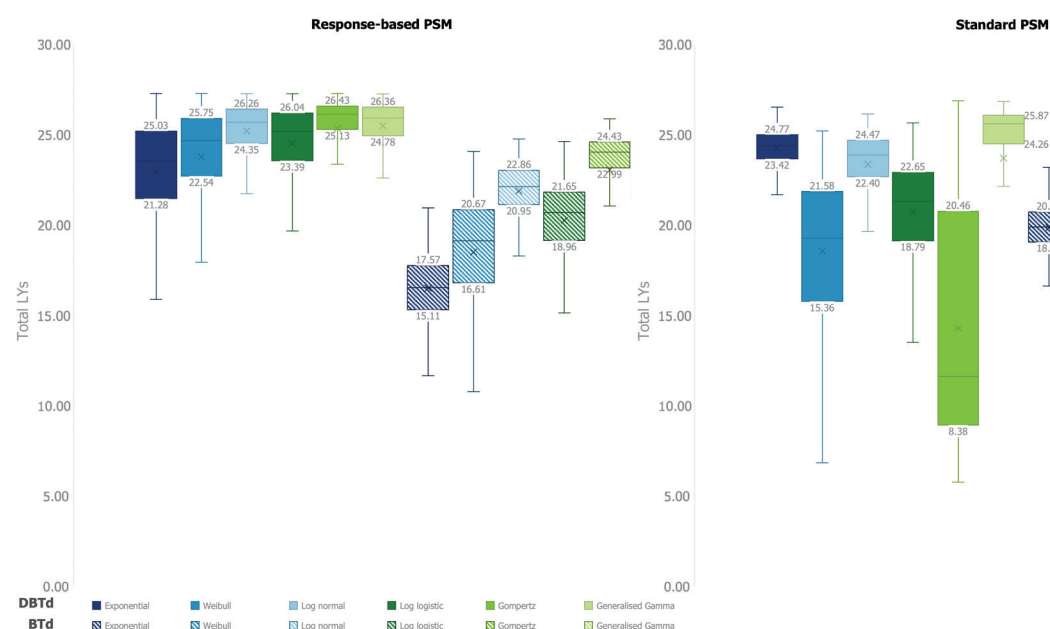
^aCurve choices refer to the extrapolation of MRD-positive BTd patients; hazard ratios are then applied to this extrapolation to generate survival estimates for the full BTd cohort and the DBTd cohort.

^bRanges have been calculated from unrounded data and therefore may differ from the range of the values presented. Blue cells indicate the curves that were chosen for the base case extrapolations.

Table 3. Validation of Base Case Model Outcomes (Percentage of Patients Alive)

Data Source	Time (mo)					
	12	18	24	36	60	120
DBTd						
Standard PSM – log logistic	99	98	97	95	89	75
Response-based PSM – exponential (BTd MRD+)	98	97	95	93	88	77
CASSIOPEIA ³²	98	98	97	—	—	—
BTd						
Standard PSM – Weibull	98	97	95	90	81	57
Response-based PSM – exponential (BTd MRD+)	97	94	92	87	78	60
CASSIOPEIA ³²	98	95	93	—	—	—
GIMEMA ³⁶	97	93	91	86	79	60
NHS Digital cohort BTd as first-line treatment (ASCT-positive) ³⁷	98	96	93	88	—	—

Abbreviations: ASCT, autologous stem cell transplant; BTd, bortezomib, thalidomide, and dexamethasone; DBTd, daratumumab, bortezomib, thalidomide, and dexamethasone; HR, hazard ratio; MRD+, minimal residual disease-positive; NDMM, newly diagnosed multiple myeloma; PSM, partitioned survival model.

Figure 3. Mean Life-Years Estimated by 2000 Simulations, by Treatment and Model Structure

Horizontal lines within the box indicate median LYs; X denotes mean LYs. For the response-based PSM, curve choices refer to the extrapolation of MRD+ BTd patients; hazard ratios are then applied to this extrapolation to generate survival estimates for the full BTd cohort and the DBTd cohort. For the standard PSM, curve choices refer to the extrapolation of DBTd and BTd patients directly from the CASSIOPEIA trial.

Abbreviations: BTd, bortezomib, thalidomide and dexamethasone; DBTd, daratumumab, bortezomib, thalidomide and dexamethasone; LY, life-year; PSM, partitioned survival model.

data from the clinical trials and real-world settings (Table 3). The base case extrapolations in both the standard and response-based PSM, and across both treatment arms, produced closely matching estimates of OS compared with the observed data from the CASSIOPEIA trial, based on 2 years of follow-up.

In addition, the base case extrapolations for BTd in both the standard and response-based PSM showed reasonably similar estimates of longer-term OS compared with observed data from the GIMEMA trial, based on 10 years of follow-up, and the observed real-world survival outcomes from NHS Digital, based on 3 years of follow-up.

Uncertainty of Long-term Outcomes

In the standard PSM, the high level of divergence in long-term estimates across different parametric extrapolations for both the DBTd and BTd treatment arms reflects the immaturity of the OS data and resultant uncertainty. In comparison, the divergence in extrapolations was found to be substantially reduced in the response-based PSM (Figure 2). This is demonstrated by the mean LYs estimated by the different extrapolation curves across the lifetime time horizon in the standard PSM, which had a range (min-max) of 14.8 (11.1-25.9) and 11.8 (8.5-20.3) LYs for DBTd and BTd, respectively, compared with 3.4 (22.5-25.9) and 7.7 (16.4-4.0) LYs for the response-based PSM (Table 2). This trend remains when removing any clinically implausible extrapolations (ie, Gompertz for DBTd and BTd in the standard PSM; Figure 2). For DBTd in the standard PSM, the range decreased from 14.8 to 7.1 LYs when excluding the Gompertz curve, which is still greater than the range observed for DBTd in the response-based PSM of 3.4 LYs. For BTd, the range decreases from 11.8 to 6.7 LYs in the standard PSM, comparable to the range observed in the response-based PSM of 7.7 LYs.

These findings were supported by probabilistic simulations; the range of mean LYs estimated in the PSA from different parametric extrapolations is narrower in the response-based PSM than the standard PSM (Figure 3). Additionally, the probabilistic analyses demonstrated that uncertainty of LY estimates *within* a given parametric extrapolation

(as indicated by the interquartile range [IQR] and range) was also reduced in the response-based PSM relative to the standard PSM for most extrapolations (Figure 3; Supplementary Table 2).

Although the uncertainty of the response-based PSM estimates was lower than that of the standard PSM between and within parametric extrapolations for both the DBTd and BTd arms, the reduction in uncertainty associated with the response-based model was greater for DBTd (Figures 2 and 3).

DISCUSSION

Interpretation of Results

The results demonstrated that this response-based PSM was associated with reduced uncertainty in long-term survival outcomes compared with the standard PSM, illustrated by the lower divergence in OS estimates and mean LYs predicted across different parametric extrapolations in the response-based PSM. This trend remained when excluding the clinically implausible extrapolations. The reduced uncertainty associated with this response-based PSM was further supported by the smaller range of LYs estimated for a given parametric model in the PSA, which was true for all parametric extrapolations across both treatment arms, with the exception of the exponential curve. The finding was particularly pronounced for DBTd due to the relative immaturity of the data used in the standard PSM for this treatment arm.

As such, the results support use of a response-based PSM, which in this case better accounts for underlying patient heterogeneity vs the standard PSM by modeling patients by MRD status. This approach also facilitates the incorporation of external data with longer follow-up and a more precise pooled estimate of effect, decreasing reliance on direct extrapolation of less mature survival data (eg, MRD- patients).

Despite the high level of divergence in long-term estimates across different parametric extrapolations for both DBTd and BTd, both the standard and response-based PSMs could generate plausible estimates of LYs compared with long-term survival estimates from GIMEMA

and observed data from CASSIOPEIA. However, the reduced divergence of extrapolations in the response-based PSM provided a greater number of extrapolations that were clinically plausible. This demonstrates that, regardless of the approach taken, there is a need to explore multiple extrapolations and assess the clinical plausibility of extrapolations using external sources of validation and clinical expert opinion.

Within the response-based PSM approach, BTd showed more divergence in long-term OS estimates and a larger range of probabilistic LYs for the majority of extrapolations than DBTd. As the response-based PSM for DBTd includes all the same parameters as BTd, with the addition of 2 HRs applied to the BTd arm (associated with additional uncertainty), the increased divergence for BTd vs DBTd may appear counterintuitive as one may expect the uncertainty to be similar or increased across all extrapolations. This discrepancy is likely because of the GPM limit applied as a cap to OS, which has a greater influence on DBTd than BTd (in both model structures), given the improved survival of DBTd. This is exemplified by the shape of the response-based extrapolations, whereby the shapes of the DBTd curves were visually more similar to each other and the GPM curve than was seen for BTd. Additionally, the upper bounds of the DBTd IQRs on the response-based PSM were all tightly aligned, suggesting that there is an upper limit in the mean predicted LYs. Applying the GPM cap on extrapolations prevents implausible extrapolations whereby patients treated with DBTd are predicted to live beyond the general population, which would not be accounted for in the extrapolations of observed data. As a result, the uncertainty in long-term survival associated with DBTd was less than BTd, despite having 2 additional parameters associated with uncertainty.

Within each model structure, particularly the standard PSM, the level of uncertainty associated with each curve changed considerably. Depending on the hazard of the observed data, different curves are associated with different levels of uncertainty in long-term survival estimates.³⁸ This is also partially explained by the GPM cap. For example, extrapolations with the smallest IQR in the standard PSM (exponential and log normal for both DBTd and BTd) had the most optimistic OS extrapolations, which were therefore more influenced by the GPM cap, limiting the extent to which their long-term extrapolations could vary. In comparison, extrapolations that had lower OS estimates (eg, Gompertz and Weibull for both DBTd and BTd) showed larger IQRs, reflective of more uncertainty.

It was also noted that when comparing between model structures, different extrapolations for the same treatment arms had different levels of associated uncertainty. This variability was expected given the underlying observed data used to inform the long-term survival estimates for each treatment were different between model structures.

Strengths and Limitations

A key strength of the study was the demonstration of uncertainty through several means: namely, by demonstrating less divergence in the extrapolations for OS and smaller ranges in the predicted mean LYs, as well as demonstrating reduced uncertainty within each extrapolation method through probabilistic simulations. Furthermore, the comparison of model estimates against the observed data from CASSIOPEIA and a range of external data sources with long-term follow-up was critical to validate the survival estimates from each model structure.³²

Both models were developed using IPD from the same source (CASSIOPEIA), supporting comparability of results and avoiding reliance on simulated IPD or aggregate data. The response-based PSM was developed in line with NICE DSU guidance for landmark analyses and leveraged a clinically meaningful landmark time point at which to assess MRD status. A major strength of utilizing a response-based model

structure is that it acknowledges underlying patient heterogeneity and facilitates incorporation of external data with longer follow-up in the form of a pooled estimate, to inform the relationship between MRD status and OS.²⁰ The response-based model structure was recently accepted by both NICE and the Scottish Medicines Consortium as part of the HTA process for daratumumab in this indication.^{39,40}

A potential limitation of this analysis is how the GPM cap was implemented in the model. As noted in the methods, the survival hazard was set to the minimum of the GPM and the hazard of the OS extrapolation. Recently, alternative methods for incorporating GPM into survival extrapolations, such as incorporating GPM hazards in the log-link function, have come to light, which might influence survival estimates.^{17,41} As such, exploring alternative methods to incorporate GPM into this type of analysis requires further investigation. A second limitation of this analysis was that GPM was not sampled in the PSA. Therefore, any associated uncertainty in the GPM would not have been accounted for. However, since the GPM comes from population-level data (ie, the Office for National Statistics), it is not expected to be associated with parameter uncertainty and is therefore not typically included in probabilistic simulations.⁴²

In taking a response-based modeling approach, the limitations associated with landmark analyses also apply.¹⁷ For example, the choice of MRD assessment time point may influence the analysis, and the reliability of the long-term survival data for each responder category can impact uncertainty. Splitting samples by response can also increase uncertainty in extrapolated estimates due to the smaller patient populations (ie, the BTd MRD+ population). We also acknowledge the appropriateness of modeling survival via the use of HRs is dependent on the assumption of proportional hazards over time. For the duration of the trial, this was assessed and confirmed via inspection of log-cumulative hazard plots (see **Supplementary Figure S1**) and was assumed to be maintained for the modeled time horizon. A further limitation of this model structure is that MRD status was assessed at a single time point (100 days post-ASCT). While recent relapsed/refractory MM studies indicate that sustained MRD negativity for 12 or more months translates to improved long-term outcomes, reported data on the significance of durable MRD response in the front-line setting is limited.⁴³

Implications

This paper demonstrates an alternative approach to long-term survival modeling for patients with NDMM following ASCT that reduces uncertainty compared with a standard PSM, by incorporating MRD status and external data in a response-based PSM framework, avoiding the need to fit standard parametric extrapolations to immature survival data.

This alternative approach to cost-effective modeling in NDMM may have wider implications for modeling survival. For example, this response-based approach can be used to reduce uncertainty in long-term survival estimates in other disease areas, where (1) there are clear response categories with a validated relationship between response and survival and (2) mature survival data are not available at the time of drug assessment. By extension, this reduces uncertainty in cost-effectiveness results assessed by HTA bodies for decision-making, enabling faster access to novel treatments. To allow the use of this approach, drug manufacturers would benefit from ensuring the necessary data are collected when designing clinical trials (eg, defining a clinically relevant landmark time point and including a response assessment).

This response-based model produced more clinically plausible results than the standard PSM, aligning with the stronger clinical rationale underpinning the modeling approach that better captures patient heterogeneity. As targeted treatments, such as precision medicine and advanced therapy medicinal products, become more common and

widely used, the use of models founded in biological plausibility in the face of immature data will be vital for estimating long-term outcomes and, ultimately, healthcare decision-making.

Furthermore, a response-based modeling approach provides a framework to assess and explore alternative pricing mechanisms, such as outcomes-based pricing.

CONCLUSIONS

In disease areas with a relatively good prognosis, such as NDMM, estimating long-term survival for novel therapies is often challenging, as immature data can lead to considerable uncertainty in long-term survival estimates. Systematic literature reviews and meta-analyses have established MRD status as the most sensitive outcome in NDMM and supported its use as a surrogate for PFS and OS.²⁰⁻²²

This paper presents an alternative approach for estimating long-term survival in NDMM, leveraging MRD status via a landmark analysis with external data to inform the relationship between MRD status and OS. This approach demonstrates that the use of MRD status in a response-based PSM reduces uncertainty in modeling long-term survival in patients with NDMM who are eligible for ASCT by providing a greater number of clinically plausible extrapolations compared with a standard PSM. By reducing uncertainty in modeling predictions, the method presented here can support faster market access for patients to more effective medicines by reducing decision-making uncertainty around cost-effectiveness estimates for HTA bodies and payers. This method has potential to be applied in other disease areas with immature survival data and a suitable marker for survival.

Author Contributions: N.v.H. participated in the conception, design and development of the economic model, substantial contributions to analysis and interpretation of the data and drafting of the manuscript. K.S. participated in the conception, design, and development of the economic model, acquisition of data, substantial contributions to analysis and interpretation of the data, and drafting of the manuscript. P.M. participated in the conception, design and development of the economic model, substantial contributions to analysis and interpretation of the data and drafting of the manuscript. N.T. participated in the conception, design and development of the economic model, acquisition of data, substantial contributions to analysis and interpretation of the data, and drafting of the manuscript.

Acknowledgments: The authors acknowledge Harry Freeman, BSc, from Costello Medical, UK, for medical writing and editorial assistance.

Disclosures: N.v.H. is an employee of Costello Medical Consulting; P.M. was an employee of Costello Medical Consulting at the time of this research, and K.S. and N.T. are employees of Janssen-Cilag Ltd.

Funding: This study was funded by Janssen-Cilag Ltd, who, along with CASSIOPEIA study cosponsors Intergroupe Francophone du Myelome and the Dutch-Belgian Cooperative Trial Group for Hematology Oncology, have proprietary rights to the clinical trial data included in the health economic model presented in this manuscript.

REFERENCES

- Food and Drug Administration (FDA). *Orphan Drug Designations and Approvals. Daratumumab*. <http://bit.ly/3W33z9t>
- European Medicines Agency (EMA). *Public Summary of Opinion on Orphan Designation: Daratumumab for the Treatment of Plasma Cell Myeloma*. 2013.
- Brenner H, Gondas A, Pulte D. Recent major improvement in long-term survival of younger patients with multiple myeloma. *Blood*. 2008;111(5):2521-6. doi:10.1182/blood-2007-08-104984
- Kumar SK, Rajkumar SV, Dispenzieri A, et al. Improved survival in multiple myeloma and the impact of novel therapies. *Blood*. 2008;111(5):2516-20. doi:10.1182/blood-2007-10-116129
- Pratt G, Jenner M, Owen R, et al. Updates to the guidelines for the diagnosis and management of multiple myeloma. *Br J Haematol*. 2014;167(1):131-3. doi:10.1111/bjh.12926
- Usmani S, Ahmadi T, Ng Y, et al. Analysis of real-world data on overall survival in multiple myeloma patients with ≥ 3 prior lines of therapy including a proteasome inhibitor (PI) and an immunomodulatory drug (IMiD), or double refractory to a PI and an IMiD. *Oncologist*. 2016;21(11):1355-1361. doi:10.1634/theoncologist.2016-0104
- Yong K, Delforge M, Driessen C, et al. Multiple myeloma: patient outcomes in real-world practice. *Br J Haematol*. 2016;175(2):252-264. doi:10.1111/bjh.14213
- Voorhees P, Fonseca R, Cote S, et al. Characterization of frontline treatment patterns and the proportion of patients reaching subsequent lines of therapy in transplant eligible patients with newly diagnosed multiple myeloma. *HemaSphere*. 2019;3:274-5.
- Moreau P, San Miguel J, Sonneveld P, et al. Multiple myeloma: ESMO clinical practice guidelines for diagnosis, treatment and follow-up. *Ann Oncol*. 2017;28(suppl 4):iv52-iv61. doi:10.1093/annonc/mdx096
- National Institute for Health and Care Excellence (NICE). NG35. Myeloma: diagnosis and management.(2016. <https://www.nice.org.uk/guidance/ng35>.
- Knauf W, Aldaoud A, Hutzschenreuter U, et al. Survival of non-transplant patients with multiple myeloma in routine care differs from that in clinical trials-data from the prospective German Tumour Registry Lymphatic Neoplasms. *Ann Hematol*. 2018;97(12):2437-2445. doi:10.1007/s00277-018-3449-8
- Nishimura KK, Barlogie B, van Rhee F, et al. Long-term outcomes after autologous stem cell transplantation for multiple myeloma. *Blood Adv*. 2020;4(2):422-431. doi:10.1182/bloodadvances.2019000524
- Palumbo A, Avet-Loiseau H, Oliva S, et al. Revised international staging system for multiple myeloma: a report from international myeloma working group. *J Clin Oncol*. 2015;33(26):2863-9. doi:10.1200/jco.2015.61.2267
- Richardson P, Mitsiades C, Schlossman R, et al. The treatment of relapsed and refractory multiple myeloma. *Hematology Am Soc Hematol Educ Program*. 2007:317-23. doi:10.1182/asheducation-2007.1.317
- European Medicines Agency. Guideline on the use of minimal residual disease as a clinical endpoint in multiple myeloma studies. Accessed August 17, 2021. https://www.ema.europa.eu/en/documents/scientific-guideline/draft-guideline-use-minimal-residual-disease-clinical-endpoint-multiple-myeloma-studies_en.pdf
- Latimer N. *NICE DSU Technical Support Document 14: Undertaking Survival Analysis for Economic Evaluations Alongside Clinical Trials—Extrapolation With Patient-Level Data*. 2011. <http://nicedsu.org.uk/>
- Rutherford M, Lambert P, Sweeting M, et al. NICE DSU Technical Support Document 21: Flexible Methods for Survival Analysis. <http://nicedsu.org.uk/>
- Kumar S, Paiva B, Anderson KC, et al. International Myeloma Working Group consensus criteria for response and minimal

- residual disease assessment in multiple myeloma. *Lancet Oncol.* 2016;17(8):e328-e346. doi:10.1016/s1470-2045(16)30206-6
19. Paiva B, van Dongen JJ, Orfao A. New criteria for response assessment: role of minimal residual disease in multiple myeloma. *Blood.* 2015;125(20):3059-68. doi:10.1182/blood-2014-11-568907
 20. Munshi N, Avet-Loiseau H, Anderson KC, et al. Expanded meta-analysis confirms the association between MRD and long-term survival outcomes in multiple myeloma (MM). Poster presented at: American Society of Hematology (ASH). 2019.
 21. Munshi NC, Avet-Loiseau H, Rawstron AC, et al. Association of minimal residual disease with superior survival outcomes in patients with multiple myeloma: a meta-analysis. *JAMA Oncol.* 2017;3(1):28-35. doi:10.1001/jamaoncol.2016.3160
 22. Munshi NC, Avet-Loiseau H, Anderson KC, et al. A large meta-analysis establishes the role of MRD negativity in long-term survival outcomes in patients with multiple myeloma. *Blood Adv.* 2020;4(23):5988-5999. doi:10.1182/bloodadvances.2020002827
 23. Landgren O, Devlin S, Boulad M, Mailankody S. Role of MRD status in relation to clinical outcomes in newly diagnosed multiple myeloma patients: a meta-analysis. *Bone Marrow Transplant.* 2016;51(12):1565-1568. doi:10.1038/bmt.2016.222
 24. Paiva B, Vidriales MB, Cerveró J, et al. Multiparameter flow cytometric remission is the most relevant prognostic factor for multiple myeloma patients who undergo autologous stem cell transplantation. *Blood.* 2008;112(10):4017-23. doi:10.1182/blood-2008-05-159624
 25. Lahuerta JJ, Paiva B, Vidriales MB, et al. Depth of response in multiple myeloma: a pooled analysis of three PETHEMA/GEM clinical trials. *J Clin Oncol.* 2017;35(25):2900-2910. doi:10.1200/jco.2016.69.2517
 26. Harousseau JL, Avet-Loiseau H. Minimal residual disease negativity is a new end point of myeloma therapy. *J Clin Oncol.* 2017;35(25):2863-2865. doi:10.1200/jco.2017.73.1331
 27. Zweegman S, Usmani SZ, Chastain K, et al. Bortezomib, lenalidomide, and dexamethasone (VRd) ± daratumumab (DARA) in patients (pts) with newly diagnosed multiple myeloma (NDMM) for whom transplant is not planned as initial therapy: a multicenter, randomized, phase III study (CEPHEUS). *J Clin Oncol.* 2019;37(15_suppl):TPS8056-TPS8056. doi:10.1200/JCO.2019.37.15_suppl.TPS8056
 28. European Medicines Agency. Reflection paper on the use of measurable residual disease as a clinical endpoint in multiple myeloma studies. Accessed September 2022. <https://www.ema.europa.eu/en/reflection-paper-use-measurable-residual-disease-clinical-endpoint-multiple-myeloma-studies>
 29. Yamamoto C, Minakata D, Koyama S, et al. Daratumumab in first-line therapy is cost-effective in transplant-eligible patients with newly diagnosed myeloma. *Blood.* 2022;140(6):594-607. doi:10.1182/blood.2021015220
 30. de Weers M, Tai YT, van der Veer MS, et al. Daratumumab, a novel therapeutic human CD38 monoclonal antibody, induces killing of multiple myeloma and other hematological tumors. *J Immunol.* 2011;186(3):1840-8. doi:10.4049/jimmunol.1003032
 31. Khagi Y, Mark TM. Potential role of daratumumab in the treatment of multiple myeloma. *Onco Targets Ther.* 2014;7:1095-100. doi:10.2147/ott.S49480
 32. Moreau P, Attal M, Hulin C, et al. Bortezomib, thalidomide, and dexamethasone with or without daratumumab before and after autologous stem-cell transplantation for newly diagnosed multiple myeloma (CASSIOPEIA): a randomised, open-label, phase 3 study. *Lancet.* 2019;394(10192):29-38. doi:10.1016/s0140-6736(19)31240-1
 33. European Medicines Agency. *Daratumumab Assessment Report.* Accessed October 8, 2021. <https://www.ema.europa.eu/en/documents/variation-report/darzalex-h-c-4077-ii-0030-epar-assessment-report-variation.pdf>
 34. Woods B, Sideris E, Palmer S, Latimer N, Soares M. *NICE DSU Technical Support Document 19: Partitioned Survival Analysis for Decision Modelling in Health Care.* 2017. Accessed September 2022. <https://www.sheffield.ac.uk/nice-dsu/tsds/partitioned-survival-analysis>
 35. de Tute RM, Rawstron AC, Gregory WM, et al. Minimal residual disease following autologous stem cell transplant in myeloma: impact on outcome is independent of induction regimen. *Haematologica.* 2016;101(2):e69-71. doi:10.3324/haematol.2015.128215
 36. Tacchetti P, Pantani L, Patriarca F, et al. Bortezomib, thalidomide, and dexamethasone followed by double autologous haematopoietic stem-cell transplantation for newly diagnosed multiple myeloma (GIMEMA-MMY-3006): long-term follow-up analysis of a randomised phase 3, open-label study. *Lancet Haematol.* 2020;7(12):e861-e873. doi:https://doi.org/10.1016/S2352-3026(20)30323-9
 37. Health Data Insight CIC. Standing cohort study of newly diagnosed multiple myeloma (NDMM) patients in England. Report v1.0. June 6, 2021.
 38. Kearns B, Stevens J, Ren S, et al. How uncertain is the survival extrapolation? A study of the impact of different parametric survival models on extrapolated uncertainty about hazard functions, lifetime mean survival and cost effectiveness. *Pharmacoeconomics.* 2019;38:193-204.
 39. Scottish Medicines Consortium (SMC). ID 927/13: In combination with dexamethasone, or with dexamethasone and thalidomide, for the induction treatment of adult patients with previously untreated multiple myeloma who are eligible for high-dose chemotherapy with haematopoietic stem cell transplantation. <https://www.scottishmedicines.org.uk/medicines-advice/bortezomib-velcade-fullsubmission-92713/>
 40. National Institute of Health and Care Excellence (NICE). *TA763. Daratumumab in Combination for Untreated Multiple Myeloma When Stem Cell Transplant Is Suitable: Final Appraisal Document.* Accessed April 7, 2022. <https://www.nice.org.uk/guidance/ta763/documents/final-appraisal-determination-document>
 41. van Oostrum I, Ouwers M, Remiro-Azócar A, et al. Comparison of parametric survival extrapolation approaches incorporating general population mortality for adequate health technology assessment of new oncology drugs. *Value Health.* 2021;doi:https://doi.org/10.1016/j.jval.2021.03.008
 42. Briggs AH, Weinstein MC, Fenwick EAL, Karnon J, Schulpher MJ, Paltiel AD. Model parameter estimation and uncertainty: a report of the ISPOR-SMDM modeling good research practices task force-6. *Value Health.* 2012;15(6):835-842. doi:https://doi.org/10.1016/j.jval.2012.04.014
 43. Avet-Loiseau H, San-Miguel J, Casneuf T, et al. Evaluation of sustained minimal residual disease negativity with daratumumab-combination regimens in relapsed and/or refractory multiple myeloma: analysis of POLLUX and CASTOR. *J Clin Oncol.* 2021;39(10):1139-1149. doi:10.1200/jco.20.01814