ScienceDirect

Contents lists available at sciencedirect.com
Journal homepage: www.elsevier.com/locate/jval

Economic Evaluation

Estimating and Extrapolating Survival Using a State-Transition Modeling Approach: A Practical Application in Multiple Myeloma

Istvan Majer, PhD, Sonja Kroep, PhD, Rana Maroun, PhD, Claire Williams, PhD, Sven Klijn, MSc, Stephen Palmer, MSc

ABSTRACT

Objectives: State-transition models (STMs) applied in oncology have given limited considerations to modeling post-progression survival data. This study presents an application of an STM focusing on methods to evaluate the postprogression transition and its impact on survival predictions.

Methods: Data from the lenalidomide plus dexamethasone arm of the ASPIRE trial was used to estimate transition rates for an STM. The model accounted for the competing risk between the progression and preprogression death events and included an explicit structural link between the time to progression and subsequent death. The modeled transition rates were used to simulate individual disease trajectories in a discrete event simulation framework, based on which progression-free survival and overall survival over a 30-year time horizon were estimated. Survival predictions were compared with the observed trial data, matched external data, and estimates obtained from a more conventional partitioned survival analysis approach.

Results: The rates of progression and preprogression death were modeled using piecewise exponential functions. The rate of postprogression mortality was modeled using an exponential function accounting for the nonlinear effect of the time to progression. The STM provided survival estimates that closely fitted the trial data and gave more plausible long-term survival predictions than the best-fitting Weibull model applied in a partitioned survival analysis.

Conclusions: The fit of the STM suggested that the modeled transition rates accurately captured the underlying disease process over the modeled time horizon. The considerations of this study may apply to other settings and facilitate a wider use of STMs in oncology.

Keywords: economic evaluation, extrapolation, multiple myeloma, multistate model, oncology, state-transition model, survival analysis.

VALUE HEALTH. 2022; 25(4):595–604

Introduction

Partitioned survival analysis has become the most commonly applied decision-modeling approach for economic evaluations in oncology.[1–3] Partitioned survival models (PSMs) are characterized by several mutually exclusive health states, such as progression-free (PF), postprogression, and death, that represent the progressive nature of the disease. In PSMs, state membership is determined by a set of survival curves fitted to time-to-event data routinely reported for clinical trials, for example, progression-free survival (PFS) and overall survival (OS). PSMs are intuitively appealing and easy to communicate and generally provide accurate estimates of OS for the within-trial period.[1] Nevertheless, the survival curves that commonly capture some of the same events are modeled independently and do not represent mutually exclusive estimates of state membership. In addition, the lack of structural relationship between progression and subsequent death can limit the validity of long-term survival extrapolations if those are based on the within-trial trend of OS only. Therefore, the use of external evidence has been generally recommended to support and validate longer-term extrapolations.[4]

The most common alternative to the PSM approach is the state-transition modeling (STM) approach, often applied in the form of Markov models. STMs describe clinical pathways in terms of mutually exclusive health states that patients move through during the course of their disease.[5–7] In STMs, the proportion of patients occupying a health state at a certain time point is determined by transition probabilities explicitly considering the relationship between clinical events; therefore, model predictions are closely linked to biological and clinical processes and are based on a more direct use of information on prognostic intermediate endpoints. Although, for the trial follow-up period, STMs and PSMs are expected to give similar survival estimates because relationships between outcomes are represented within the data, STMs may provide more plausible extrapolations and improved transparency and allow more meaningful sensitivity analyses to be conducted.[1,7]

Partitioned survival analysis has faced considerable scrutiny over recent years as a result of its lack of theoretical underpinnings in modeling the disease process and consequent issues with extrapolating survival data.[8–11] To mitigate the limitations associated with PSMs and to assess uncertainty in OS extrapolations in decision models, the National Institute for Health and Care Excellence (NICE) recommends the use of STMs alongside PSMs.[1] Earlier literature reviews reported that only a small proportion of NICE technology appraisals in oncology used STMs, and the NICE Technical Support Document 19 found that STMs were often poorly implemented.[1–3] A recent review of modeling methods and NICE technology appraisals suggested that there has been a growing number of empirical studies applying STM approaches and that applications have become more sophisticated.[7,12–15] The review also highlighted that earlier applications of STMs gave limited consideration to modeling postprogression survival data; nevertheless, if the timing of progression is related to prognosis, for example, early progressors have shorter survival than late progressors, then a naïve analysis of postprogression survival data may bias survival extrapolations. As a consequence, including a covariate in the postprogression survival model that describes the prognosis at the time of progression may improve the accuracy of STM predictions.[7]

Given the importance of extrapolation methods for clinical and policy decision making, the objective of our study was to present an application of an STM using data from a clinical trial in multiple myeloma focusing on methods to evaluate the postprogression transition as a function of time to progression and its impact on postprogression survival and OS. To illustrate the impact of differences in structural assumptions, our study also compared survival outcomes obtained by the STM and the more conventional PSM approach.

Methods

Data

Individual-level PFS, OS, and baseline characteristics data were used from the phase 3 ASPIRE randomized controlled trial that compared carfilzomib in combination with lenalidomide and dexamethasone (KRd) with lenalidomide and dexamethasone (Rd) in 792 patients with relapsed multiple myeloma. In ASPIRE, KRd significantly improved both PFS (hazard ratio [HR] 0.66; 95% confidence interval [CI] 0.55-0.78) and OS (HR 0.79; 95% CI 0.67-0.95) versus Rd.[16] For the present study, data from the Rd arm of the trial (n = 396) were used because more PFS and OS events occurred in the Rd arm. In addition, the purpose of the study was to present and discuss the considerations associated with modeling baseline transitions and ultimately to set up an STM. Incorporating the treatment effect was beyond the scope of this study.

By the time of the data cut (April 2017), 238 patients (60%) experienced progression, 34 (9%) died before progression, and 124 (31%) were censored for PFS. Of the 238 patients who experienced progression, 191 (80.3%) died after progression, whereas 42 of the 124 patients (33.9%) who were censored for PFS died. The median follow-up for PFS and OS in Rd patients were 48.0 months and 67.1 months, respectively. The median follow-up for postprogression survival was 54.9 months. Per protocol, disease assessment was performed on day 1 of each treatment cycle; nevertheless, assessment could also be done at any time in an unscheduled fashion. For the PFS analysis, the date of the actual assessment was considered. Similarly, for the OS analysis, the actual death date recorded in medical records was considered. Details on baseline characteristics of patients have been published elsewhere[16,17] and are presented in the Supplemental Materials found at https://doi.org/10.1016/j.jval.2021.09.011.

Long-term data from a retrospective analysis of the Registry of Monoclonal Gammopathies (RMG) were used to inform model selection and validate OS predictions.[18] The RMG registry was founded in 2007 by the Czech Myeloma Group and is intended to collect real-world clinical data regarding the treatment of patients in Czech Republic and Slovakia. The analysis included 880 patients who had received 1 to 3 previous therapies were treated with Rd and had information on OS. Maximum follow-up in the analysis cohort was approximately 9 years. A comparison of the ASPIRE and RMG patient population is provided in the Supplemental Materials found at https://doi.org/10.1016/j.jval.2021.09.011.

STM Approach

An STM was developed including PF, postprogression (PP), and death (D) health states using monthly cycles. As in the commonly applied illness-death structure,[1] transitions were allowed from the PF to the PP state (PF → PP), from the PF to the D state (PF → D), and from the PP to the D state (PP → D). All patients started the model in the PF health state, from where they could either progress or die. Once patients progressed, they could die and transition from the PP health state to the D health state (see Fig. 1).

By design, the STM accounted for 2 important features. First, the competing risk between the mutually exclusive progression and preprogression death events, which could occur to patients in the PF health state, was explicitly modeled. Because the occurrence of one event precludes the occurrence of the other event, the corresponding transition probabilities are not independent and cannot be directly estimated from a single transition-specific hazard alone.[19] Second, the STM included an explicit structural link between progression and subsequent death. As a consequence, the modeled OS was a function of all 3 transitions because the number of deaths was directly influenced by the mortality risk of PF and PP patients and by the evolving number of PP patients, which was affected by the risk of progression.

The STM was implemented in the R statistical software[20] and adopted a continuous-time multistate modeling framework. The model can also be labeled as a state-arrival extended semi-Markov model for the following reasons. First, for estimating postprogression transition rates, the time was measured from entry in the postprogression state; that is, the clock was reset for each simulated patient upon entering the postprogression state and so the model formed a Markov renewal or semi-Markov process. Second, the postprogression transition hazard depended on the time of arrival at the postprogression state; that is, estimation of the hazard included a parameter associated with the time until progression. The STM, using a discrete event simulation approach,

Figure 1. Model structure.

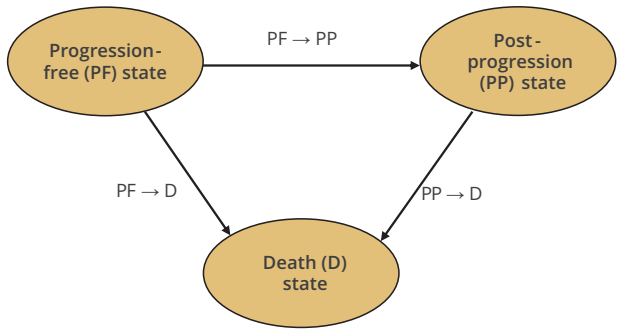
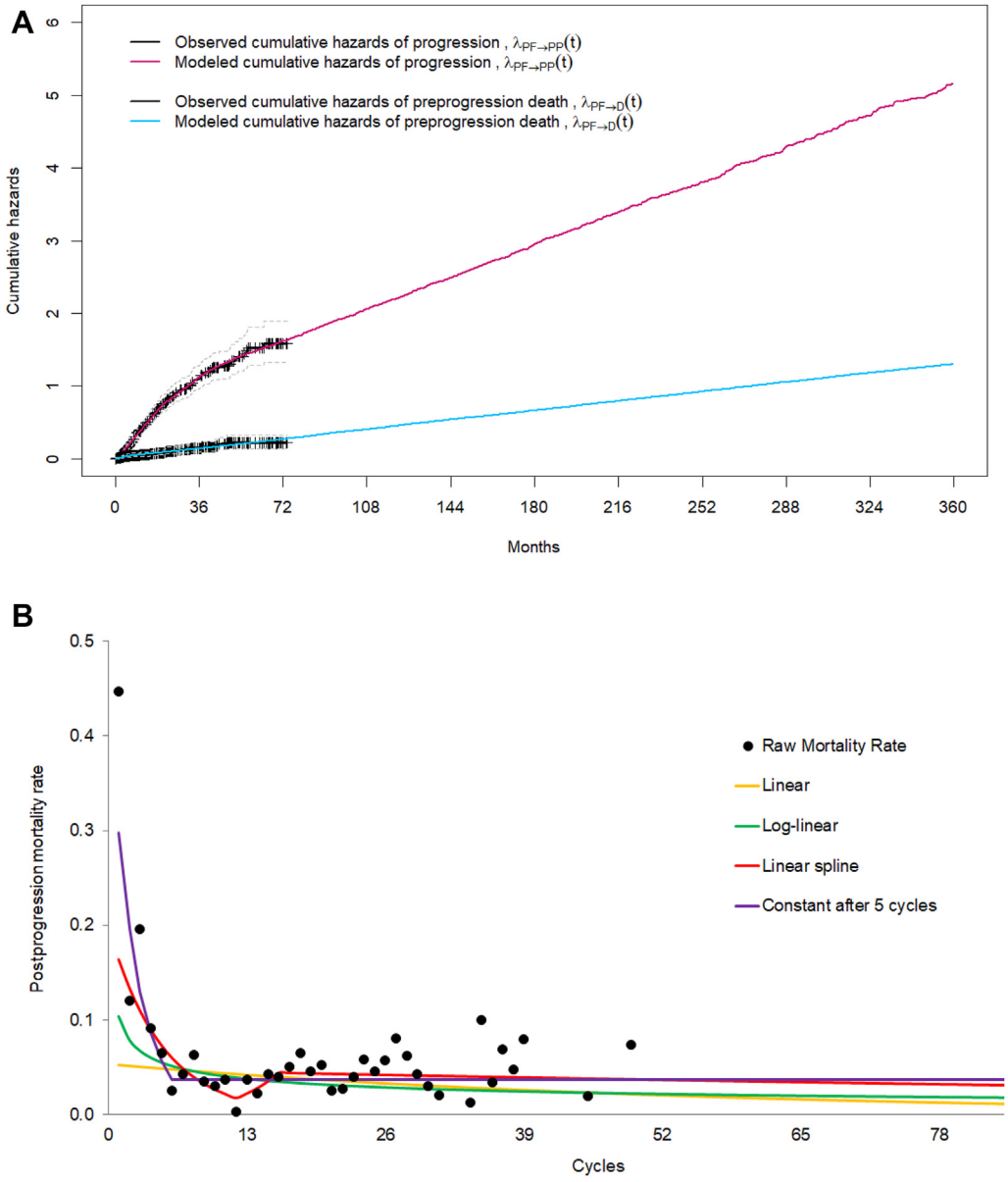


Figure 2. Modeled transition rates. (A) Cumulative rate of progression and preprogression death. (B) Cumulative rate of postprogression death. *Note*. AIC values for the different models: Linear, 1559.9; log-linear, 1548.9; linear spline, 1529.9; constant after 5 cycles, 1532.6. Cycle-specific raw mortality rate was estimated by fitting an exponential model to data from patients who progressed during a given cycle.

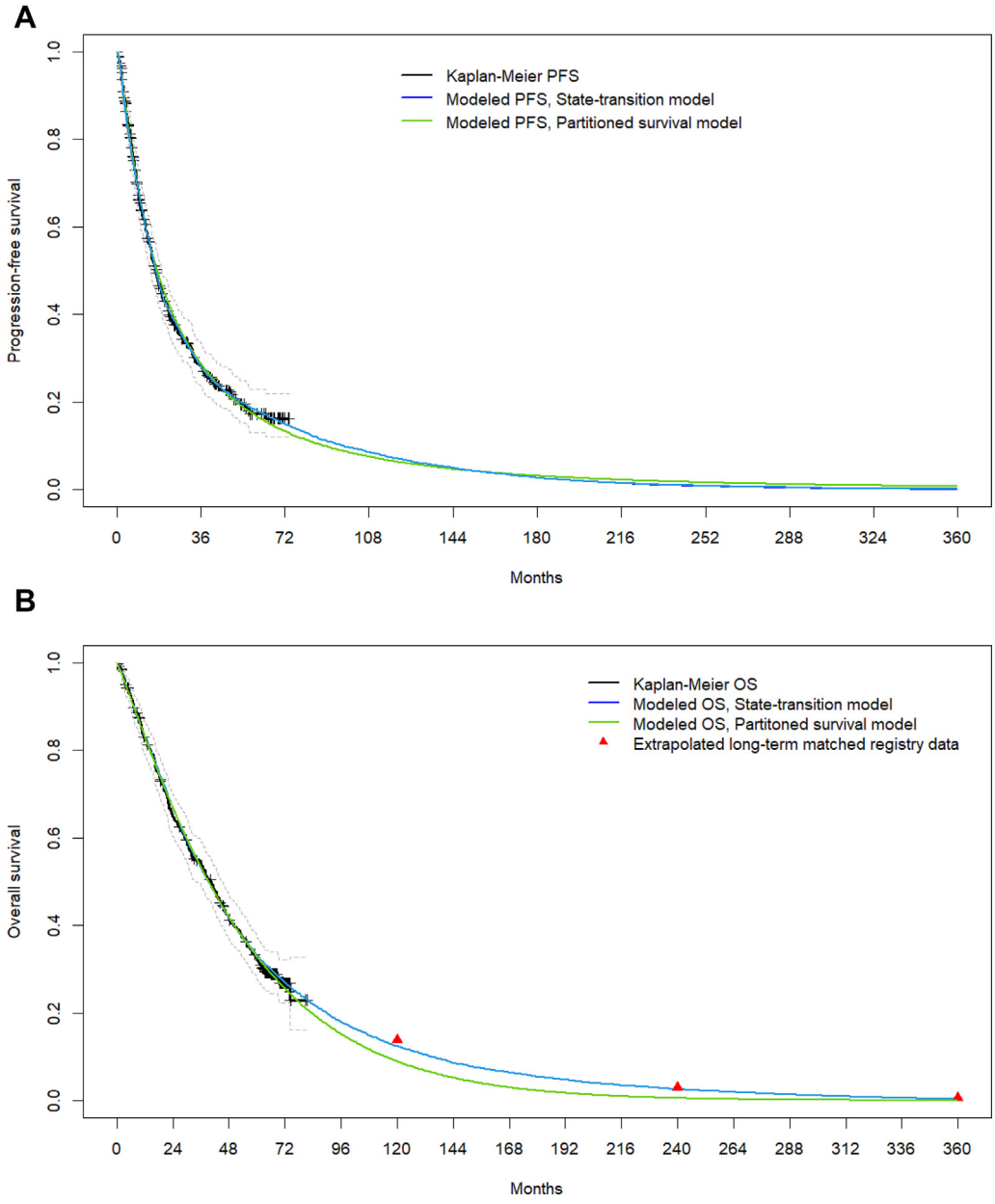


AIC indicates Akaike information criterion.

simulated a large number of individual disease trajectories for identical patients based on the transition-specific cumulative hazards and estimated state occupancy probabilities by counting the number of simulated patients in each state over time. This approach provided a high level of convenience because, for example, assessing various functional forms that captured the impact of the time to progression on postprogression survival and ultimately on OS was possible with little programming. Although using a cohort simulation approach for such analyses is possible and a transparent trace within an economic model is sometimes explicitly required,[21] it may have necessitated the use of a large number of formulas and a consequent large Excel spreadsheet size because postprogression survival should have been followed for each subgroup of newly progressing patients in a cycle. The STM simulated individual disease trajectories for a large sample of patients (N = 10 000) using transition hazards for the 3 possible transitions (ie, from PF to PP, from PF to D, and from PP to D). The hazard functions specified the rate of progression, preprogression death, and postprogression death at time t and were denoted by $\lambda_{PF\rightarrow PP}(t)$, $\lambda_{PF\rightarrow D}(t)$, and $\lambda_{PP\rightarrow D}(t)$, respectively. The hazard for each transition was determined based on event-history data from the ASPIRE trial.

Transition hazards were modeled and extrapolated using standard parametric distributions (ie, Weibull, exponential, Gompertz, generalized gamma, log-logistic, and lognormal) and piecewise exponential models. Piecewise models with a

Figure 3. Modeled PFS and OS with state-transition and partitioned survival model. (A) PFS. *Note.* For the partitioned survival model, the generalized gamma model was used. (B) OS. *Note.* For the partitioned survival model, the Weibull model was used.



OS indicates overall survival; PFS, progression-free survival.

maximum of 3 cutoff points were explored. The number of cutoff points applied was based on visual inspection of fitted models to the cumulative hazards plot and goodness-of-fit statistic using the Akaike information criterion (AIC). For estimating the hazard of postprogression death, additional models were explored because in a Cox proportional hazards model the time to progression was found to have a statistically significant association with post-progression survival (HR 0.97, $P<.001$) indicating that a patient's risk of death after progression decreased with 3% after each additional month spent PF. Although this specific Cox model was run as a preliminary analysis, it confirmed the presence of a strong association, indicated evidence against the Markov property, and suggested that including a covariate describing the time to progression may improve the accuracy of postprogression survival and ultimately OS predictions. Therefore, several models were explored that estimated the hazards of postprogression mortality as a function of the time to progression. The models postulated various relationships, describing the impact of time to progression on postprogression survival with increasing complexity including linear, log-linear, and different spline functions (ie, linear, cubic, and quadratic splines). The models were compared to identify the functional form that provided the best fit to the data considering the trade-off between the number of parameters and goodness of fit. Given the complexity of the explored postprogression death risk models in terms of the included covariate, the parametric distribution for the baseline hazard model was determined

without the covariate and, supported by goodness of fit, was preselected to be the exponential function (see Supplemental Materials found at https://doi.org/10.1016/j.jval.2021.09.011). As a result, the estimated postprogression mortality rate was constant for the postprogression lifetime of patients who progressed in a given cycle but this rate was different across patients who progressed in different cycles.

To estimate the rate of postprogression death, data from patients who progressed during the first 3 years of follow-up (224 of 238 patients) were used in the base case analysis because most patients who progressed late were censored for death at the trial data cut. In particular, 87% of patients who progressed during the first 3 years subsequently died whereas for patients who progressed after 3 years, it was only 14%. Similarly, the follow-up time for postprogression survival in patients who progressed before 3 years was considerably longer (median 4.7 years) than in those who progressed after 3 years (median 2.0 years). As a result, postprogression mortality was notably lower for patients who progressed after 3 years, but this was an artifact of the short follow-up time to observe postprogression death in these patients. Given that the time until progression was a predictor in the postprogression mortality model, including these patients would have distorted the estimated relationship between the time to progression and postprogression death and ultimately would have resulted in implausibly high OS estimates. The impact of including all patients in the postprogression mortality model was explored in a scenario analysis.

Transition hazard models were fitted to each transition separately and were compared based on their goodness of fit using the AIC and visual assessment. In particular, for $\lambda_{PF\rightarrow PP}(t)$ and $\lambda_{PF\rightarrow D}(t)$, the fitted models were visually compared with the observed cumulative hazards, whereas for $\lambda_{PP\rightarrow D}(t)$ the estimated postprogression mortality rates were visually compared with the "raw" postprogression mortality rates. The raw rates were obtained by fitting a separate exponential survival model to postprogression survival data in each subset of patients who progressed in a given cycle. Hence, these raw rates gave an indication of the postprogression mortality rate given the time of progression. This more complex approach for visualizing modeled versus raw postprogression mortality rates helped greatly to understand the relationship between time to progression and postprogression death risk. In addition, selection of transition rate models was informed and validated by the comparison of survival predictions with the STM versus long-term OS based on matched RMG data.

Individual times elapsing until an event occurred (ie, progression, preprogression death, postprogression death) were simulated based on the general inversion method using the cumulative hazard functions for each possible transition.[22] Then, based on the simulated time-to-event values, patient-level disease trajectories were constructed as follows. If for a patient the simulated time to preprogression death was shorter than the simulated time to progression ($t_{PF\rightarrow D} < t_{PF\rightarrow PP}$), the patient was considered to have died before progression. In this case, the lifetime of the simulated patient was equal to the time the patient spent in the PF state. In contrast, if for a patient the simulated time to progression was shorter than the time to preprogression death ($t_{PF\rightarrow PP} \leq t_{PF\rightarrow D}$), the patient was considered to have experienced progression. In this case, the simulated postprogression survival time was added to the simulated time to progression, and the lifetime of the patient was equal to the sum of the time spent before and after progression (see Fig. S1 in the Supplemental Materials found at https://doi.org/10.1016/j.jval.2021.09.011 that illustrates the approach).

Using the simulated patient trajectories, health state membership was determined for each month. Subsequently, PFS and OS were estimated by the proportion of patients in the PF health state and by the sum of patients in the PF and PP health states, respectively. Finally, the mean PFS and OS were determined by calculating the area under the predicted PFS and OS curves over 30 years.[23] Confidence intervals (CIs) around the mean estimates were generated by nonparametric bootstrapping method.[24] To ensure the modeled risk of death is not less than that for the general population, age- and sex-adjusted US all-cause mortality was applied.[25] All analyses and simulations were performed in R. Survival analyses (Cox proportional hazards models and parametric survival analyses) were performed using the survival and flexsurv packages,[26,27] the ASPIRE trial data were prepared for estimating transition-specific hazards models by using the mstate package,[28] and the different splines for the postprogression mortality risk models were constructed using the splines package.[29] Piecewise exponential hazards models were fitted to data split into time intervals with an indicator for each period. The R program codes are provided in the Supplemental Materials found at https://doi.org/10.1016/j.jval.2021.09.011.

PSM

PFS and OS were also estimated by the PSM approach. In particular, PFS and OS were predicted by fitting standard parametric and piecewise exponential functions to the individual patient-level PFS and OS data from the ASPIRE trial and extrapolating over a 30-year time horizon. The PFS and OS models were selected based on the AIC values, visual assessment of the fitted models versus the Kaplan-Meier curves, previously accepted OS models in relapsed/refractory multiple myeloma by health technology assessment bodies, and clinical plausibility of long-term extrapolations using external data.

Model Validation

Predictions of the STM and the PSM were compared with matched long-term OS data derived from a retrospective analysis of the RMG registry.[18] The estimated OS range was 13.7% to 15.4%, 2.9% to 5.7%, and 0.1% to 2.6% at 10, 20, and 30 years, respectively. Details on the matching procedure and how survival estimates based on the matched registry data were obtained are provided in the Supplemental Materials found at https://doi.org/10.1016/j.jval.2021.09.011.

Results

STM Approach

Several distributions were assessed to model and extrapolate the rate of progression ($\lambda_{PF\rightarrow PP}(t)$) and preprogression death ($\lambda_{PF\rightarrow D}(t)$). Regarding the rate of progression, the piecewise exponential model with cutoff points at cycle 1, 22, and 45 provided the best fit (AIC 2107.7) and captured the shape of the cumulative hazard function well. Other functions had considerably worse fit (eg, Weibull AIC 2143.8) or resulted in overly optimistic extrapolations of PFS. The generalized gamma function yielded predictions similar to the piecewise exponential model; nevertheless, it had a worse goodness of fit (AIC 2113.2). Overall, the piecewise exponential model was selected for the base case analysis. Regarding the rate of preprogression death, the piecewise exponential model with a cutoff point at cycle 4 was selected (AIC 437.7). Other functions captured the higher preprogression mortality risk during the first few cycles less accurately. Of the parametric functions, the Weibull model had the best fit (AIC 438.8).

Table 1. Comparison of overall survival estimated by state-transition models and partitioned survival models.

Model type	Model	OS, year 1 (%)	OS, year 2 (%)	OS, year 3 (%)	OS, year 4 (%)
Observed data	Kaplan-Meier estimate (ASPIRE)	82.8 (79.1-86.7)	65.0 (60.4-70.0)	54.2 (49.4-59.5)	41.6 (36.9-46.9)
State-transition model	Base case model	83.3 (80.8-86.6)	67.1 (63.5-71.6)	53.3 (48.8-57.8)	41.9 (37.3-46.8)
	PP MR constant after 5 cycles	84.1 (82.0-86.9)	67.5 (64.6-71.6)	53.5 (49.4-58.2)	42.2 (37.4-47.0)
	PP MR constant after 5 cycles, including all patients*	84.7 (82.2-87.4)	68.7 (65.3-72.9)	55.1 (50.7-60.1)	43.6 (38.9-49.1)
	PP MR has log-linear relationship with TTP	84.3 (81.9-87.6)	66.6 (63.4-71.4)	52.6 (49.1-58.0)	42.8 (38.4-47.6)
	PP MR has linear relationship with TTP	86.1 (83.7-89.1)	67.9 (63.9-73.1)	53.1 (48.4-58.5)	42.3 (37.3-47.0)
	PP MR independent from TTP	86.9 (85.0-89.6)	69.3 (65.4-73.9)	53.6 (48.7-59.0)	40.7 (36.0-46.6)
	PP MR independent from TTP (generalized gamma function)	87.1 (84.5-89.7)	69.2 (64.3-72.9)	53.1 (48.1-57.9)	41.1 (36.1-46.4)
	PP MR independent from TTP (Gompertz function)	86.7 (84.4-89.4)	68.2 (64.3-73.2)	53.4 (48.4-58.3)	41.3 (36.1-46.4)
	PP MR independent from TTP (Weibull function)	87.1 (85.0-89.9)	69.0 (65.7-73.6)	52.9 (49.3-58.6)	41.3 (36.7-46.3)
	Progression rate modeled with generalized gamma function	83.4 (80.6-87.0)	67.4 (63.8-72.2)	53.5 (49.1-58.6)	42.2 (37.7-47.5)
	Preprogression MR modeled with Weibull function	83.2 (80.5-86.6)	66.9 (63.4-71.2)	53.2 (49.1-57.7)	41.9 (37.6-46.6)
Partitioned survival model	Weibull OS (base case)	83.2 (80.0-86.4)	67.2 (63.4-71.1)	53.4 (49.4-57.7)	42.1 (37.8-46.6)
	Generalized gamma OS	83.1 (79.8-86.3)	66.7 (62.5-70.9)	53.0 (48.4-57.5)	41.9 (37.4-46.4)
	Piecewise exponential OS	81.7 (79.2-84.2)	66.7 (62.8- 70.9)	54.5 (49.7-59.7)	42.5 (38.2-47.1)

Note. Values in parentheses indicate the 95% confidence intervals. In the base case state-transition model, the rates of progression and preprogression death were modeled using piecewise exponential functions, whereas the rate of postprogression mortality was modeled using an exponential function accounting for the nonlinear impact of time to progression with splines.
LE, life expectancy; MR, mortality rate; OS, overall survival; PP, postprogression; TTP, time to progression.
*Including all progressed patients in the postprogression mortality model. Progression time was made equal to 3 years if the true progression time was larger than 3 years.

The observed versus modeled cumulative hazards for the 2 transitions in the base case analysis are presented in Figure 2.

To model the rate of postprogression death ($\lambda_{PP\rightarrow D}(t)$), the exponential model was selected (AIC 1561.8) (see Supplemental Materials found at https://doi.org/10.1016/j.jval.2021.09.011). Given that the time to progression was a strong and significant predictor of postprogression survival, in the next steps, the rate of postprogression death was modeled with a series of functions that included a covariate indicating the cycle in which the patients progressed. The first model assumed a linear relationship between the cycle of progression and the log-rate of postprogression mortality. Compared with the first model, the fit was improved (AIC 1559.9) and the time covariate was highly significant, indicating that patients who progressed in earlier cycles had a higher risk of postprogression death. To explore possible further improvements, instead of a linear relationship between the time to progression and postprogression death, the log-linear relationship, spline models of various degrees, and models assuming constant effect beyond a certain cycle were assessed. Of these, the linear spline model with knots placed at 12 and 16 cycles had the best fit based on the AIC (AIC 1529.9) and visual assessment. Hence, this latter model was selected for the base case analysis. The raw versus predicted mortality rates based on various postprogression mortality rate models are presented in Figure 2. The parameters of the different postprogression mortality hazard models are provided in the Supplemental Materials found at https://doi.org/10.1016/j.jval.2021.09.011.

The STM provided PFS estimates that fitted the PFS Kaplan-Meier curve remarkably well (see Fig. 3) and yielded plausible long-term predictions (ie, 7.1%, 1.0%, and 0.0% at 10, 20, and 30 years, respectively). In terms of OS, the base case STM seemed to have a very good fit to the Kaplan-Meier curve (see Fig. 3) and provided OS rates of 12.5%, 2.6%, and 0.3% at 10, 20, and 30 years, respectively. These were generally below the OS rates indicated by

Table 1. Continued

OS, year 5 (%)	OS, year 6 (%)	OS, year 7 (%)	OS, year 10 (%)	OS, year 20 (%)	OS, year 30 (%)	LE (years)
32.3 (27.8-37.4)	26.8 (22.3-32.2)	-	-	-	-	-
33.2 (28.6-38.6)	27.0 (21.9-32.5)	22.1 (16.7-28.0)	12.5 (6.7-19.8)	2.6 (0.3-9.1)	0.3 (0.0-1.7)	4.9 (4.2-6.4)
33.6 (28.6-38.2)	26.8 (22.0-31.1)	21.8 (17.0-26.0)	11.5 (7.7-15.9)	1.7 (0.4-4.5)	0.1 (0.0-0.7)	4.8 (4.1-5.5)
34.8 (30.0-40.3)	27.9 (23.2-33.5)	22.6 (18.0-28.0)	12.2 (8.3-17.2)	1.6 (0.4-4.9)	0.1 (0.0-0.8)	4.9 (4.3-5.8)
35.2 (30.3-39.9)	29.1 (24.4-33.7)	24.3 (19.7-28.9)	15.1 (10.9-19.4)	3.6 (1.4-6.9)	0.5 (0.1-1.2)	5.3 (4.6-6.2)
34.4 (29.2-39.0)	28.2 (23.4-33.3)	23.8 (18.9-29.2)	15.4 (10.2-21.6)	6.0 (1.5-11.1)	1.1 (0.1-2.1)	5.6 (4.6-6.9)
31.8 (26.9-37.3)	25.2 (20.4-30.3)	20.0 (15.4-24.9)	10.9 (6.6-15.7)	1.5 (0.3-5.1)	0.1 (0.0-0.9)	4.7 (4.0-5.6)
32.2 (27.4-37.7)	26.0 (21.1-31.2)	21.0 (16.1-26.4)	11.7 (7.5-16.9)	1.6 (0.5-4.8)	0.1 (0.0-0.7)	4.8 (4.1-5.7)
32.1 (27.7-37.6)	25.7 (21.2-31.3)	20.5 (16.4-26.3)	10.7 (7.5-17.5)	1.6 (0.5-6.2)	0.1 (0.0-1.0)	4.7 (4.2-5.9)
32.4 (27.5-37.3)	26.3 (20.8-30.4)	21.4 (15.8-25.3)	11.6 (7.0-15.7)	1.9 (0.3-4.6)	0.1 (0.0-0.7)	4.8 (4.1-5.6)
33.3 (28.6-38.8)	26.7 (21.4-32.6)	21.7 (16.1-28.0)	12.1 (7.3-19.5)	3.1 (1.1-9.0)	0.5 (0.1-1.7)	5.0 (4.2-6.4)
33.4 (28.8-38.4)	27.4 (22.3-32.6)	22.6 (17.0-28.2)	13.1 (7.4-20.3)	3.0 (0.5-10.2)	0.4 (0.0-1.9)	5.0 (4.2-6.6)
33.0 (28.5-37.5)	25.7 (21.2-30.2)	19.9 (15.6-24.2)	9.0 (5.8-12.6)	0.5 (0.1-1.4)	0.0 (0.0-0.2)	4.4 (3.9-4.9)
33.1 (28.6-37.6)	26.1 (21.4-30.6)	20.5 (15.7-25.3)	10.0 (5.3-15.0)	0.9 (0.0-3.7)	4.5 (4.0-5.3)	4.5 (4.0-5.3)
33.2 (28.9-37.8)	26.0 (21.8-30.6)	20.3 (16.2-25.1)	9.7 (6.7-13.7)	0.8 (0.3-1.9)	0.1 (0.0-0.3)	4.5 (4.0-5.1)

the matched RMG data (−1.2% point, −0.3% point, and 0.2% point at 10, 20, and 30 years, respectively) and were considered clinically plausible. The predicted life expectancy with the base case model was 4.9 years (95% CI 4.2-6.4).

Other STMs had worse fit to the Kaplan-Meier OS curve and provided higher survival estimates at 20 and 30 years. When the impact of the time to progression was excluded from the postprogression mortality rate model, OS predictions had a considerably worse fit to the Kaplan-Meier curve particularly during the first years. This was also true when alternative parametric models (generalized gamma, Gompertz, Weibull) without the covariate for time to progression were explored, supporting the finding that adding the event history was key to achieving a better fit to OS (see Supplemental Materials found at https://doi.org/10.1016/j.jval.2021.09.011). Although to a lesser extent, OS was still overestimated with postprogression death rate models that captured the impact of the time to progression in a linear or log-linear fashion. The life expectancy estimated with these models was 5.6 (95% CI 4.6-6.9) and 5.3 (95% CI 4.6-6.2) years, respectively. Other scenarios (ie, assuming constant postprogression death risk beyond 5 cycles, generalized gamma model for the rate of progression, Weibull model for the rate of preprogression death) had a smaller impact on the results although these yielded worse fit to the OS data. The results for the different STMs in terms of the predicted OS rate at different years and the estimated life expectancy are detailed in Table 1.

PSM

Of the survival models fitted directly to the PFS data, the generalized gamma function provided the most plausible long-term PFS estimates (6.3%, 1.8%, and 0.1% at 10, 20, and 30 years, respectively) and had the third-lowest AIC value. Although the log-logistic and lognormal models had somewhat better fit to the data, these models predicted implausibly high long-term PFS rates and so were discarded. Among the standard survival regression models directly fitted to the OS data, the Weibull was considered to be the most suitable model because it had the lowest AIC value and it was accepted in recent health technology appraisals by the NICE.[30,31] The Weibull model provided lower long-term OS rates than the STM and the matched RMG data (9.0%, 0.5%, and 0.0% at 10, 20, and 30 years, respectively) and an estimated life expectancy of 4.4 years (95% CI 3.9-4.9).

Discussion

This study presented an application of an STM using data from the control arm of the ASPIRE trial in relapsed multiple myeloma. The analysis focused on the modeling and extrapolation of transition hazards and assessed various functions to accurately capture the relationship between the time of progression and postprogression mortality. Transition rates from the PF health state were estimated with piecewise exponential functions whereas the transition rate from the PP health state was modeled with an exponential function accounting for the impact of time to progression. The modeled transition rates were used to simulate individual disease trajectories in a discrete event simulation framework based on which PFS and OS were derived.

The STM predicted the observed trial data remarkably well. In particular, the modeled survival had excellent fit to the Kaplan-Meier curve, to longer-term trial OS data, and to matched external data, indicating that the modeled transition rates appropriately captured the underlying disease process. The mean OS predicted by the STM was 4.9 years, which was approximately half year longer than that estimated by the best-fitting Weibull model applied directly to the OS data in a PSM. Scenario analyses suggested that predictions of the mean OS were robust (range 4.8-5.6 years) although the estimated survival had worse fit to the Kaplan-Meier curve when alternative postprogression survival modeling assumptions were used.

Our study found that piecewise exponential models approximated the shape of the progression and preprogression death rate profile most closely. This is consistent with a previous analysis that fitted a 3-state multistate model to Rd data from the pooled MM-009/010 trials in multiple myeloma.[32] The analysis found that piecewise models with cutoff points determined at 6 months and 2 years estimated PFS and OS reasonably well and provided clinically plausible survival extrapolations. Regarding the association between the time to progression and postprogression survival in our study, a linear spline model fitted the data best. The spline model suggested that patients who experience progression later have longer postprogression survival. This positive correlation is supported by previous clinical studies and to some extent explained by the heterogeneity of the patient populations observed (eg, cytogenetic abnormalities) and not observed (response to treatment) at baseline.[33–36] In ASPIRE, early relapse was primarily associated with worse stage of disease, worse performance status, and refractoriness to the last previous line of therapy. An overview of patient characteristics in ASPIRE by the time of relapse is presented in the Supplemental Materials found at https://doi.org/10.1016/j.jval.2021.09.011.

Several published studies have compared survival estimates obtained by STMs and PSMs suggesting that these 2 approaches can give considerably different survival predictions.[12–14,37–39] Although STMs were considered to outperform PSMs,[7] they often had implementation challenges. In addition, the risk of progression was sometimes derived directly from time to progression data,[14,40] even though in the presence of competing events the hazard of progression itself does not determine the risk of progression.[19] Such an approach may provide biased number of progressing patients and predictions for OS. The relationship between time to progression and postprogression survival has been rarely explored. We are aware of one study where survival predictions were done by stratifying patients into groups based on time to progression.[14] The current study represents an improvement compared with previous applications in this respect because several postprogression mortality models were explored to accurately capture the relationship between time to progression and postprogression survival and ultimately to improve the within-trial fit to the observed data and extrapolations. Recently, Williams[15,39] published a series of tutorials on the use of STM for cost-effectiveness analysis in oncology and encouraged the adoption of the multistate modeling approach. Similarly to the present study, the authors proposed the use of a simulation approach when the underlying structural assumptions are complex.

Our study has relevance for economic evaluations in oncology because cost-effectiveness analyses typically require clinical endpoints to be extrapolated beyond the observed trial data. Because the method chosen for the extrapolation is often a key driver of the estimated cost-effectiveness, progress in survival modeling methods have received increased attention.[4,11,15,41] For the present study, by explicitly modeling the underlying transitions between health states, understanding the causal relationship between progression and postprogression death was achieved by assessing several functional forms. This in turn provided more plausible survival predictions and allowed more meaningful scenario analyses to be performed than the PSM. Although assessing the treatment effect on the transitions was not in the scope of this study, it should be noted that estimating the postprogression hazards for the KRd arm of the ASPIRE trial may be limited by insufficient and immature data. In such a situation, the modeled relationship between progression and subsequent death should be subject to careful consideration including a comparison of patient characteristics at progression and subsequent treatments across the treatment arms and whether the treatment itself affects postprogression survival.

The findings of this study may be also relevant for situations when adjustment for treatment switching or reflecting local treatment patterns is required. With the increasing availability of novel treatments in multiple myeloma, patients in clinical trials may go through different treatment pathways depending on the study drug they receive.[42] As a consequence, postprogression treatments may be imbalanced, which in turn can confound OS results. In addition, treatments that clinical trial patients receive over the course of the disease may differ from the treatments available in a given country. In these situations, estimating postprogression survival with models that include covariates indicating specific postprogression therapies may provide helpful insights.

There are several limitations associated with the use of STMs in oncology that are applicable to our study. First, analysis of postprogression data may be subject to informative censoring; therefore, the estimated relationship between time to progression and postprogression survival may be biased and the trend observed in the trial period for postprogression death may misinform the extrapolation period. Given the relatively long follow-up in ASPIRE and the high number of patients with documented progression event, such risk was deemed to be small for the current study. Second, it has been recognized that achieving a satisfactory fit to the observed survival is often challenging because OS is determined by the combined effect of all modeled transitions. Although, in the present study, the observed and estimated survival were remarkably aligned, it required a postprogression survival model with a nonlinear covariate to be set up. Estimating a complex relationship for the postprogression survival may be particularly challenging because the model should capture the changing relationship between the time to progression and postprogression survival and the potential change in the shape of the postprogression survival itself. For the current study, the shape of postprogression survival was assumed to remain the same, which allowed focusing on a simpler problem and facilitated visual comparison of the raw and modeled mortality rates. Further

research may explore the use of alternative methods, such as those applied for population mortality forecasts,[43] that model changes in both the scale and the shape of mortality profiles over time. Finally, it has been recognized that relying solely on AIC or other goodness-of-fit measures to select transition hazard models may have limited validity in the presence of competing risks.[39] For the present study, goodness of fit to the observed data was considered together with both a visual assessment of the fitted and extrapolated outcomes and the validity of predictions compared with external registry data. Nevertheless, in general, better fit to the data does not necessarily translate to better predictive outcomes in extrapolation.

Conclusions

STMs applied in oncology have given limited consideration to modeling postprogression survival data. This study presented an STM applied to data in multiple myeloma and assessed various models to capture the relationship between time to progression and subsequent death. Early relapse has been established as a risk factor for survival in other hematologic malignancies and solid tumors[14,44]; therefore, the considerations of this study may apply to other settings, facilitate a wider use of STMs in oncology, and ultimately improve clinical and policy decision making.

Supplemental Materials

Supplementary data associated with this article can be found in the online version at https://doi.org/10.1016/j.jval.2021.09.011.

Article and Author Information

Accepted for Publication: September 2, 2021

Published Online: November 3, 2021

doi: https://doi.org/10.1016/j.jval.2021.09.011

Author Affiliations: Global Value and Access, Health Economics and Outcomes Research, Amgen (Europe) GmbH, Rotkreuz, Switzerland (Majer, Maroun); OPEN Health, Modeling and Meta-Analysis, Rotterdam, the Netherlands (Kroep, Klijn); Health Economics Research Centre, University of Oxford, Oxford, UK (Williams); Centre for Health Economics, University of York, York, UK (Palmer).

Correspondence: Istvan Majer, PhD, Global Value and Access, Health Economics and Outcomes Research, Amgen (Europe) GmbH, Suurstoffi 22, CH-6343 Rotkreuz, Switzerland. Email: imajer@amgen.com

Author Contributions: *Concept and design:* Majer, Kroep, Maroun, Williams, Klijn, Palmer
Analysis and interpretation of data: Majer, Kroep, Maroun, Klijn
Drafting of the manuscript: Majer, Kroep, Palmer,
Critical revision of the paper for important intellectual content: Majer, Kroep, Maroun, Williams, Klijn, Palmer
Statistical analysis: Majer, Kroep
Administrative, technical, or logistic support: Williams

Conflict of Interest Disclosures: Dr Majer is employed by Amgen (Europe) GmbH. Dr Kroep is employed by Pharmerit, an OPEN Health Company. Dr Maroun was employed by Amgen at the time of the study and is currently employed by Roche. Dr Williams reported that this work was a collaboration with Amgen on a personal basis rather than being connected to her position at the University of Oxford during the conduct of this study. Dr Klijn is employed by Pharmerit, an OPEN Health Company, and reported receiving payment from Amgen to Pharmerit, an OPEN Health Company, for conducting the analyses during the conduct of the study. Dr Palmer reported receiving personal payment from Amgen for being a member of an advisory board during the conduct of the study.

Funding/Support: The study was supported by Amgen.

Role of the Funder/Sponsor: The funder had no role in the design of the study, the conduct of the analysis, interpretation of the data, and preparation of the manuscript.

Acknowledgment: The authors thank Carla Kraan of Pharmerit International for providing writing support funded by Amgen.

REFERENCES

1. Woods BS, Sideris E, Palmer S, et al. NICE DSU technical support document 19. Partitioned survival analysis for decision modelling in health care: a critical review. Decision Support Unit, ScHARR, University of Sheffield. http://nicedsu.org.uk/wp-content/uploads/2017/06/Partitioned-Survival-Analysis-final-report.pdf.
2. Bullement A, Cranmer HL, Shields GE. A review of recent decision-analytic models used to evaluate the economic value of cancer treatments. *Appl Health Econ Health Policy*. 2019;17(6):771–780.
3. Gallacher D, Auguste P, Connock M. How do pharmaceutical companies model survival of cancer patients? A review of NICE single technology appraisals in 2017. *Int J Technol Assess Health Care*. 2019;35(2):160–167.
4. Jackson C, Stevens J, Ren S, et al. Extrapolating survival from randomized trials using external data: a review of methods. *Med Decis Making*. 2017;37(4):377–390.
5. Siebert U, Alagoz O, Bayoumi AM, et al. State-transition modeling: a report of the ISPOR-SMDM modeling good research practices task Force-3. *Value Health*. 2012;15(6):812–820.
6. Srivastava T, Latimer NR, Tappenden P. Estimation of transition probabilities for state-transition models: a review of NICE appraisals. *Pharmacoeconomics*. 2021;39(8):869–878.
7. Woods BS, Sideris E, Palmer S, Latimer N, Soares M. Partitioned survival and state transition models for healthcare decision making in oncology: where are we now? *Value Health*. 2020;23(12):1613–1621.
8. Coyle D, Coyle K. The inherent bias from using partitioned survival models in economic evaluation. *Value Health*. 2014;17(3):A194.
9. Beca J, Husereau D, Chan KKW, Hawkins N, Hoch JS. Oncology modeling for fun and profit! Key steps for busy analysts in health technology assessment. *Pharmacoeconomics*. 2018;36(1):7–15.
10. Masucci L, Beca J, Sabharwal M, Hoch JS. Methodological issues in economic evaluations submitted to the Pan-Canadian oncology drug review (pCODR). *Pharmacoecon Open*. 2017;1(4):255–263.
11. Hawkins N, Grieve R. Extrapolation of survival data in cost-effectiveness analyses: the need for causal clarity. *Med Decis Making*. 2017;37(4):337–339.
12. Gao L, Li SC. Modelled economic evaluation of nivolumab for the treatment of second-line advanced or metastatic squamous non-small-cell lung cancer in Australia using both partition survival and markov models. *Appl Health Econ Health Policy*. 2019;17(3):371–380.
13. Pan F, Reifsnider O, Zheng Y, et al. Modeling clinical outcomes in prostate cancer: application and validation of the discrete event simulation approach. *Value Health*. 2018;21(4):416–422.
14. Smare C, Lakhdari K, Doan J, Posnett J, Johal S. Evaluating partitioned survival and markov decision-analytic modeling approaches for use in cost-effectiveness analysis: estimating and comparing survival outcomes. *Pharmacoeconomics*. 2020;38(1):97–108.
15. Williams C, Lewsey JD, Briggs AH, Mackay DF. Cost-effectiveness analysis in R using a multi-state modeling survival analysis framework: a tutorial. *Med Decis Making*. 2017;37(4):340–352.
16. Siegel DS, Dimopoulos MA, Ludwig H, et al. Improvement in overall survival with carfilzomib, lenalidomide, and dexamethasone in patients with relapsed or refractory multiple myeloma. *J Clin Oncol*. 2018;36(8):728–734.
17. Stewart AK, Rajkumar SV, Dimopoulos MA, et al. Carfilzomib, lenalidomide, and dexamethasone for relapsed multiple myeloma. *N Engl J Med*. 2015;372(2):142–152.
18. Maisnar V, Stefanikova Z, Spicka I, et al. Lenalidomide and dexamethasone in treatment of patients with relapsed and refractory multiple myeloma - analysis of data from the Czech myeloma Group Registry of monoclonal gammopathies. *Neoplasma*. 2019;66(3):499–505.
19. Putter H, Fiocco M, Geskus RB. Tutorial in biostatistics: competing risks and multi-state models. *Stat Med*. 2007;26(11):2389–2430.
20. R Development Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. https://www.R-project.org.
21. Canadian Agency for Drugs and Technologies in Health. Procedures for CADTH Reimbursement Reviews. https://cadth.ca/sites/default/files/Drug_Review_Process/CADTH_Drug_Reimbursement_Review_Procedures.pdf.
22. Crowther MJ, Lambert PC. Simulating biologically plausible complex survival data. *Stat Med*. 2013;32(23):4118–4134.
23. Majer IM, Kumar SK, Palmer S. Time horizons in cost analyses. *JAMA*. 2019;322(6):582–583.
24. Rizzo ML. *Statistical Computing With R*. 2nd ed. Boca Raton: CRC Press; 2019.
25. Human Mortality Database. Germany: University of California BU, and Max Planck Institute for Demographic Research. www.mortality.org or www.humanmortality.de. Accessed February 24, 2020.
26. Therneau T. A Package for Survival Analysis in R. R Package. https://cran.r-project.org/web/packages/survival/vignettes/survival.pdf.

27. Jackson CH. flexsurv: A Platform for Parametric Survival Modeling in R. *J Stat Softw*. 2016;70:i08.
28. de Wreede LC, Fiocco M, Putter H. mstate: an R package for the analysis of competing risks and multi-state models. *J Stat Softw*. 2011;38(7):1–30.
29. Douglas MB, William NV. Regression Spline Functions and Classes. R Core Team. https://stat.ethz.ch/R-manual/R-devel/library/splines/html/splines-package.html.
30. National Institute for Health and Care Excellence. Carfilzomib for previously treated multiple myeloma [TA657]. https://www.nice.org.uk/guidance/ta657/documents/committee-papers.
31. National Institute for Health and Care Excellence. Ixazomib with lenalidomide and dexamethasone for treating relapsed or refractory multiple myeloma [TA505]. https://www.nice.org.uk/guidance/ta505/documents/committee-papers.
32. Saunders O, Gregory J, Lee D, Farrell J. Lenalidomide for treating multiple myeloma after 1 prior treatment; an application of multi-state markov modelling for the extrapolation of patient level survival data. *Value Health*. 2016;19(7):A382.
33. Majithia N, Rajkumar SV, Lacy MQ, et al. Early relapse following initial therapy for multiple myeloma predicts poor outcomes in the era of novel agents. *Leukemia*. 2016;30(11):2208–2213.
34. Kumar S, Mahmood ST, Lacy MQ, et al. Impact of early relapse after auto-SCT for multiple myeloma. *Bone Marrow Transplant*. 2008;42(6):413–420.
35. Mateos MV, Goldschmidt H, San-Miguel J, et al. Carfilzomib in relapsed or refractory multiple myeloma patients with early or late relapse following prior therapy: A subgroup analysis of the randomized phase 3 Aspire and Endeavor trials. *Hematol Oncol*. 2018;36(2):463–470.
36. Avet-Loiseau H, Attal M, Moreau P, et al. Genetic abnormalities and survival in multiple myeloma: the experience of the Intergroupe Francophone du Myélome. *Blood*. 2007;109(8):3489–3495.
37. Gibson EJ, Begum N, Koblbauer I, et al. Modeling the economic outcomes of immuno-oncology drugs: alternative model frameworks to capture clinical outcomes. *Clinicoecon Outcomes Res*. 2018;10:139–154.
38. Lee D, Amadi A, Sabater J, et al. Can we accurately predict cost effectiveness without access to overall survival data? The case study of nivolumab in combination with ipilimumab for the treatment of patients with advanced melanoma in England. *Pharmacoecon Open*. 2019;3(1):43–54.
39. Williams C, Lewsey JD, Mackay DF, Briggs AH. Estimation of survival probabilities for use in cost-effectiveness analyses: A comparison of a multi-state modeling survival analysis approach with partitioned survival and markov decision-analytic modeling. *Med Decis Making*. 2017;37(4):427–439.
40. Goeree R, Villeneuve J, Goeree J, Penrod JR, Orsini L, Tahami Monfared AA. Economic evaluation of nivolumab for the treatment of second-line advanced squamous NSCLC in Canada: a comparison of modeling approaches to estimate and extrapolate survival outcomes. *J Med Econ*. 2016;19(6):630–644.
41. Latimer NR. Survival analysis for economic evaluations alongside clinical trials–extrapolation with patient-level data: inconsistencies, limitations, and a practical guide. *Med Decis Making*. 2013;33(6):743–754.
42. Moreau P, Dimopoulos MA, Mikhael J, et al. Isatuximab plus carfilzomib and dexamethasone vs carfilzomib and dexamethasone in relapsed/refractory multiple myeloma (IKEMA): interim analysis of a Phase 3, randomized, open-label study. European Hematology Association virtual congress. https://library.ehaweb.org/eha/2020/eha25th/303392/philippe.moreau.isatuximab.plus.carfilzomib.and.dexamethasone.vs.carfilzomib.html.
43. Lee RD, Carter LR. Modeling and forecasting U.S. Mortality. *J Am Stat Assoc*. 1992;87(419):659–671.
44. Locatelli F, Schrappe M, Bernardo ME, Rutella S. How I treat relapsed childhood acute lymphoblastic leukemia. *Blood*. 2012;120(14):2807–2816.