ORIGINAL RESEARCH

Survival Analysis in Patients with Metastatic Merkel Cell Carcinoma Treated with Avelumab

Tereza Lanitis · Irina Proskorovsky · Apoorva Ambavane ·

Matthias Hunger · Ying Zheng · Murtuza Bharmal · Hemant Phatak

ABSTRACT

Introduction: Complex underlying risk functions associated with immuno-oncology treatments have led to exploration of different methods (parametric survival, spline, landmark, and cure-fraction models) to estimate long-term survival outcomes. The objective of this study was to examine differences in estimated short- and long-term survival in previously treated metastatic Merkel cell carcinoma (mMCC) patients receiving avelumab, when using alternative extrapolation approaches.

T. Lanitis (✉) · A. Ambavane
Evidera, London, UK
e-mail: Tereza.Lanitis@evidera.com

I. Proskorovsky
Evidera, Montreal, QC, Canada

M. Hunger
Mapi, an ICON plc company, Munich, Germany

Y. Zheng · H. Phatak
EMD Serono Inc., Rockland, MA, USA

M. Bharmal
Merck KGaA, Darmstadt, Germany

Methods: Efficacy data from the phase 2 JAVELIN Merkel 200 trial (part A) with at least 12 months of follow-up were analyzed. Standard parametric survival analyses and analyses of overall survival (OS) as a function of surrogate outcomes comprised of response (landmark analyses) and progression-free survival plus post-progression survival (PFS + PPS) were used to project OS. Overall survival throughout lifetime was projected and compared with the observed OS data with at least 24 months of follow-up.
Results: Estimated OS from all three approaches provided a good fit to the observed OS curve from at least 12 months of follow-up. However, performance compared with OS data from at least 24 months showed that the landmark approach followed by PFS + PPS provided a better fit to the data as compared to standard parametric analysis. Mean life expectancy estimated with avelumab was 2.48 years with best-fitting parametric function (a log-normal distribution), 3.15 years with the landmark approach, and 3.54 years with PFS + PPS.
Conclusion: Although standard parametric survival analysis may provide a good fit to short-term survival, it appears to underestimate the long-term survival benefits associated with avelumab in mMCC. Extrapolations based on surrogate outcomes of response or progression predict OS outcomes from longer follow-up better and appear to provide more clinically plausible projections.

INTRODUCTION

Determining the long-term survival associated with a new intervention is an essential component of the health economic evaluation of oncology treatments to understand whether material differences in costs and health outcomes may be produced by new treatments and determine cost-effectiveness [1–3]. Data on long-term survival are typically absent when marketing authorization or reimbursement is sought, as the outcome of interest is rarely observed for all individuals in clinical trials as a result of censoring [4, 5]. This challenge becomes even more prominent when a new therapy is investigated for use in rare tumors where small patient populations that limit the speed and sample of clinical trial enrollment [6] can lead to even more limited data on survival.

Survival analysis and extrapolation of survival curves (progression-free survival [PFS] and overall survival [OS]) using parametric models has been the dominant approach used to project long-term survival outcomes of oncology treatments. A systematic review of technology appraisals (TAs) submitted to the National Institute for Health and Care Excellence (NICE), completed by December 2009, reported that 71% of TAs for advanced or metastatic cancer used standard parametric models to obtain estimates of long-term survival [4].

In recent years this trend appears to be shifting [7–9], likely because of an increasing number of oncology therapies with novel mechanisms of action receiving regulatory approvals based on short-term data and/or surrogate endpoints [10, 11]. Specifically, the novel patterns of response observed with immuno-oncology (IO) treatments, including durable response for a fraction of patients and a plateau in survival curves [12–14], have led to

"unconventional" survival curves, inadequately captured with standard parametric approaches [7–9, 15–18].

This has led researchers to explore the use of more flexible approaches to ensure accurate estimation of clinical value to inform reimbursement decisions and consequently timely patient access to treatment [15–20]. More flexible statistical techniques such as piecewise fits have been proposed to enhance fit by separating the time-axis into different time intervals [15–18]. Such approaches allow model parameters to change in different periods, which can aid in fitting the observed pattern more closely. However, the shape of the long-term projection may remain implausible [19, 20].

Alternative modeling methods rely on capturing the heterogeneity between patients by using response (landmark analysis) and progression as surrogate outcomes (PFS + post-progression survival [PPS] approach) or modeling cured vs. non-cured patients (cure fraction models) [15–20]. In landmark analyses, the population is stratified by response status at an appropriate time point (i.e., the landmark) following initiation of treatment. Survival is then modeled separately for responders and non-responders. The OS curve is reconstructed by combining aggregated estimates of survival beyond the landmark, with survival estimates prior to the landmark [18–20]. Though this approach may provide an improved fit by capturing alternative survival patterns in the landmark groups, fitting survival in these groups may be challenging. This is particularly true when a small subset of patients respond/do not respond or where sustained remission is observed for responders [19, 20]. One alternative involves modeling OS as the sum of PFS and PPS, under the assumption that patients who progress may be at a greater risk of death, thus making PPS easier to fit. To reconstruct the OS curves, projections are required for PFS and PPS. The OS curve is then estimated by applying PPS to progressed patients with PFS informing survival among non-progressed patients. Whilst this approach may improve fit by capturing alternative survival patterns by progression status, PFS may be difficult to project, particularly among patients receiving IO therapy, where

long-term remission may manifest as a plateau in the curve [19, 20]. In this instance assumptions would be necessary to inform long-term survival among patients who do not progress.

Despite the methodological discussions of alternative OS extrapolation approaches for IO therapies, there are a limited number of studies that have systematically evaluated the short-term and long-term performance of these methods, especially in rare tumor types. The aims of this study were therefore to contribute to the evidence base by evaluating various survival analysis approaches for estimation of long-term OS in patients with a rare immunogenic tumor treated with an IO therapy.

The case of avelumab, a human monoclonal antibody approved for treatment of metastatic Merkel cell carcinoma (mMCC) [21, 22], represents a suitable case study to highlight the challenges associated with extrapolating outcomes with IO treatments. The initial approval for avelumab was based on a clinically meaningful and durable overall response rate (ORR) at a minimum of 12-month follow-up [21]. In addition, clinical trial data suggests heterogeneity in patient outcomes similarly to patterns observed with IOs in other indications [7–9]; 33% of patients showed durable antitumor response and changing hazards over time for PFS [22, 23], whilst OS data were immature to observe any plateaus. The requirement for accurate estimation of survival outcomes for avelumab in mMCC is heightened by lack of treatment options in this rare and aggressive cancer [24] and highly encouraging clinical trial results for avelumab [22, 23], necessitating timely patient access based on an assessment of long-term survival potential. Until March 2017, there were no US Food and Drug Administration (FDA)-approved treatments for patients with MCC with treatment primarily limited to cytotoxic chemotherapy or investigational therapies associated with short-lived response with a median OS of 10.2 months after first-line treatment and 4.4 months after second-line treatment or later [24].

As such, the objectives of this study were to evaluate the predictive ability and clinical plausibility of various survival modeling approaches for estimation of both the short-term and long-term OS in previously treated patients with mMCC who were receiving avelumab treatment.

METHODS

The individual patient-level data from 88 participants in part A of the phase 2, single-arm, multicenter trial JAVELIN Merkel 200 (NCT02155647), with at least 12 months of follow-up (median 16.4 months [range 12.1–25.4]) were used for analyses [22]. The trial was performed in accordance with the Helsinki declaration, had institutional review board (IRB) approval, and informed consent was received [23]. The three statistical approaches that were used to extrapolate the OS included standard parametric survival analysis, landmark analysis stratified by objective response, and the PFS + PPS approach with progression as a surrogate outcome. Although analysts have also proposed exploring the use of a cure fraction model [18, 20, 25, 26], this was not considered because OS data were not sufficiently mature to apply this method. Similarly, as the study focused on exploring alternatives to standard parametric analysis based on surrogate outcomes, piecewise fits and spline-based OS projections were not examined.

Standard Parametric Survival Analysis

Standard parametric survival analyses consisted of fitting exponential, Weibull, log-normal, log-logistic, generalized gamma, and Gompertz to the observed OS data. The best-fitting distribution was then selected using visual assessment of fit, goodness-of-fit statistics (i.e., Akaike information criterion [AIC]/Bayesian information criterion [BIC]), diagnostic plots (i.e., log cumulative hazard and parametric plots), and clinical plausibility of long-term OS projections [5, 27].

Response-Based Landmark Approach

As the first step, a time point that would maximize the number of patients who achieve

response while minimizing the number of patients who died or were censored prior to the landmark time point was selected. Median time to response was 42.7 days (range 42–252 days), with 75.9% of responses observed by 42 days [22]. Using a landmark point of 47 days, which would maximize follow-up for responders and non-responders, identified 20 patients with objective response, whilst using a landmark point of 89 days identified 27 (out of 29) patients achieving objective response. The remaining two patients who responded to treatment achieved this at 127 and 253 days [22]. Therefore, to maximize the number of patients who achieved response, whilst maximizing follow-up, the landmark time point of 89 days was chosen. After the landmark time point was selected, patients were stratified by objective response (i.e., patients with objective response vs. patients with no objective response). At the landmark point, 27 patients had achieved response with avelumab and 48 patients had not achieved response.

For patients who did not achieve response by the landmark time point, parametric survival analyses were used to project OS from landmark, similarly to the approach outlined above. However, there were very limited OS data available for patients with objective response to extrapolate survival using parametric fits (i.e., only 3 out of 27 patients with response died after the landmark point of 89 days, following disease progression). As the feasibility of parametric survival analysis is driven by the number of events, the observed data were considered too limited to reliably attempt distribution fitting. To aid in the decision-making process, observed Kaplan–Meier (KM) estimates with at least 18 months of follow-up [28, 29] which became available during the study were consulted. The data indicated a flattening of the OS curve among responders, which could not be adequately captured using parametric survival analysis. Therefore, the OS patterns in patients who respond to treatment were assumed to follow mortality patterns in the US general population [30] (assuming an average age at enrollment of 70 years [31] and that 74% [23] of patients were male). Use of general population mortality estimates without adjustment would

have resulted in higher survival estimates as compared to the observed estimates. Thus, a hazard ratio (HR) was applied to general population mortality rates, calibrated to a value of 4.5 to match the observed KM OS curves for patients with response, based on a graphical assessment of the fit.

Estimates of OS beyond the landmark for responders and non-responders were combined using the proportion of patients with/without response at the landmark point. The full OS curve was reconstructed using mortality estimates prior to the landmark. Combined OS estimates beyond the landmark were multiplied by the proportion of patients alive at the landmark.

PFS + PPS Approach

A third approach used progression as a surrogate endpoint for OS [19, 20]. Spline models and parametric distributions were fitted to project the time to progression (TTP) and PFS whilst standard parametric survival analysis was used to project PPS. OS was then estimated as a function of incident progression, PFS, and PPS. Incident progression was determined by applying risks of progression (derived from the TTP curves) to the proportion of patients that were alive and progression-free (i.e., PFS). The OS curve was reconstructed by applying PPS to progressed patients with PFS informing survival among non-progressed patients.

Nonetheless, as the PFS curve was characterized by a "plateau", use of PFS projections to determine the proportion of patients who were alive and progression-free in the long-term was considered inappropriate. From the events informing the PFS curve, eight were deaths observed in the first 4 months of follow-up, adjudicated to be attributed to progression, with the remaining being progression events. After 4 months of trial follow-up, the PFS and TTP curves converged, as both curves were informed by progression events. Consequently, use of projected PFS curves would have resulted in higher life expectancy estimates versus the general population. To correct for this, the risk of pre-progression death beyond the first

4 months of the follow-up was assumed to be equivalent to the risk of death in the general population in the United States [30]. This assumption incorporated risk of mortality in the patient population that was not projected to experience disease progression based on the plateau observed in the TTP curve.

Model Validation and Clinical Plausibility

The predicted OS from each method was used to calculate average life expectancy (assuming a time horizon of 30 years given that the average age of the population at trial enrollment was 70 years) [31] and the proportion of patients alive at 2, 3, 5, and 10 years. Both deterministic and probabilistic estimates, considering parameter uncertainty were computed (see "Probabilistic analysis" in the Electronic Supplementary Material, ESM). Best modeling practices for transparency and validation were utilized to determine goodness of fit for OS and clinical plausibility [32]. Technical validity was assessed by a graphical assessment of fits of predicted OS vs. data with a minimum follow-up of 12 months. Diagnostic plots and diagnostic statistics (AIC and BIC) were evaluated for each fit. For each method, predictive validity was assessed by comparing estimated OS vs. observed OS data with a minimum follow-up of 24 months [33], which became available after the analyses had been completed. Clinical plausibility (or face validity) of long-term projections was assessed by a practicing MCC oncologist, prior to data with at least 24 months follow-up becoming available. Data with at least 18 months of follow-up were available during the assessment of clinical plausibility.

RESULTS

Standard Parametric Survival Analysis

Among candidate parametric distributions, based on fit statistics (i.e., AIC/BIC), diagnostic plots, and clinical plausibility of long-term projection, the log-normal distribution provided the best fit to the observed trial data (see Fig. 1, ESM Table 1). Visual comparison of observed and predicted OS from at least 12 months of follow-up data suggested that the log-normal model provided a good fit to the data over the observed trial follow-up (Fig. 1). The estimated mean OS from log-normal model was 2.48 years.

Response-Based Landmark Approach

Patients Who Did Not Achieve Response
The log-normal distribution was selected as the best-fitting distribution on the basis of the fit statistics (AIC/BIC), graphical assessment of fit, and plausibility of the long-term projections. The estimated mean OS (beyond the landmark point) from the log-normal model for patients who did not achieve response was 1.2 years (Fig. 2a) (see ESM Table 2).

Patients Who Achieved Response
As a result of limited OS data available for patients who achieved response with avelumab, estimation of OS had to be based on assumptions, with an HR of 4.5 calibrated and applied to US general population mortality estimates (Fig. 2b). Under this scenario, the mean estimated OS in patients who achieved response (after the landmark point) was 7.53 years.

Combined Estimates
The observed OS data prior to landmark and estimated OS from landmark for patients who achieved response and patients who did not achieve response were combined. The predicted OS was slightly underestimated between 6 and 15 months; however, the observed tail of OS was captured well (Fig. 2c). The projected mean OS with this approach was 3.15 years.

PFS + PPS Approach

The observed PFS and TTP outcomes exhibited standard behavior observed with IO treatments across many tumor types (i.e., a sharp decline in the first few months but then stabilization and plateau over the long-term follow-up) [14, 15]. None of the parametric functions provided a
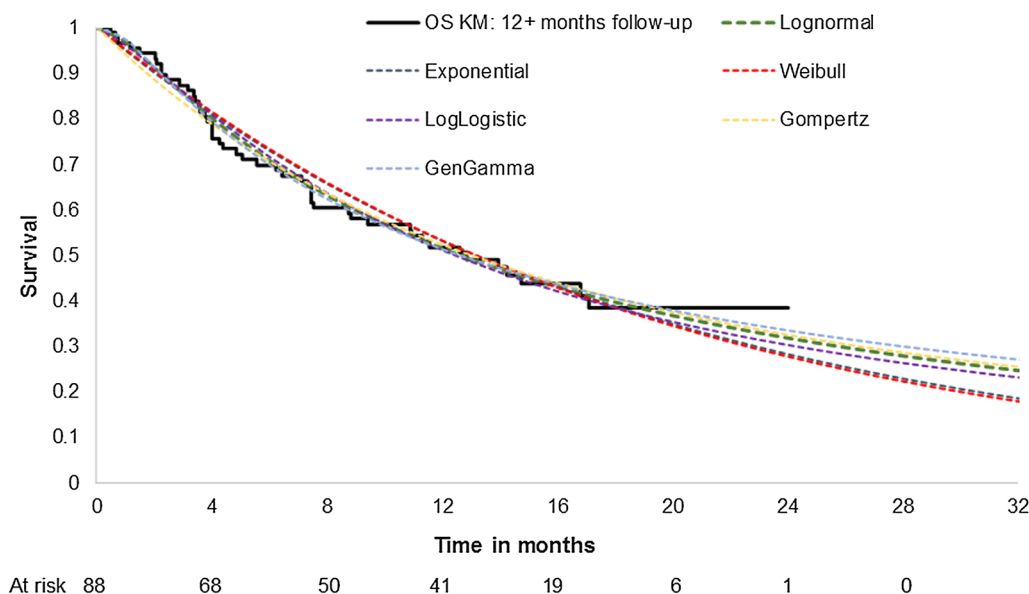
Fig. 1 Projected OS using standard parametric survival analysis

good fit to the observed PFS and TTP curves (see ESM Figs. 1 and 2, Tables 4 and 5). Spline-based models with four knots were selected and provided a good fit to the observed PFS and TTP (Fig. 3a). The TTP and PFS curves converged after approximately 4 months and the risk of progression declined with eventual plateau of the curves, which implies that approximately 11% of patients do not progress. Using the risk of progression based on TTP and pre-progression mortality based on US life tables, we estimated the mean PFS to be 2.9 years.

Results from parametric fitting of PPS suggested that the log-normal distribution provided the best fit (Fig. 3b) (see ESM Table 3). The mean PPS was estimated to be 0.78 years (9.3 months). The estimated OS from the PFS + PPS approach replicated observed OS from 12 or more months of follow-up data very closely (Fig. 3c). The mean OS using this approach was estimated to be 3.54 years.

Comparison of Approaches

The projected OS rates at different time points are shown in Table 1. The 3-year OS rate was 22% for the standard parametric survival approach, 30% for the response-based approach, and 26% for the PFS + PPS approach.

The 10-year OS rate was 5% for the standard parametric survival approach and 11% for both the response-based and PFS + PPS approaches. Average life expectancy estimated with avelumab and 95% percentiles from the probabilistic analysis was 2.48 (1.58–3.75) years with a log-normal distribution, 3.15 (2.34–4.17) years with the landmark approach, and 3.54 (2.43–5.16) years with the PFS + PPS approach (see ESM Figs. 3–5). Mean estimates from the probabilistic analyses were similar to deterministic estimates; however, the credible intervals between the approaches overlapped because of the parameter uncertainty associated with projections based on a limited sample.

Figure 4 displays the projected OS curves for avelumab among previously treated mMCC patients using a standard parametric approach, a response-based approach, and a PFS + PPS approach compared with observed data from a minimum follow-up of 24 months. Over the observed at least 12 months of follow-up of the trial used to estimate OS, all three approaches provided a good fit to the observed OS (Figs. 1, 2c, 3c). Standard parametric analysis and PFS + PPS provided a better fit between 6 and 15 months, with the landmark approach slightly underestimating OS in the same interval. However, compared with the observed data

with at least 24 months of follow-up, the landmark approach appeared to capture the tail of the OS curve better compared to the standard parametric and PFS + PPS approaches. Long-term projections varied substantially, and the log-normal model that was applied directly to the observed OS data provided the most conservative estimates. At the time of the analysis, these results were reviewed by a practicing oncologist who suggested that the PFS + PPS method represented the more plausible clinical approach to estimating OS.

DISCUSSION

Several modeling methods have been suggested to extrapolate short-term clinical outcomes associated with IO therapies; however, there is no specific consensus on the appropriate choice of method [15–19]. In the current analysis, three methods were applied to understand the differences in short-term and long-term OS associated with avelumab. There was noticeable variation in long-term OS projections between the three methods used in these analyses.

The use of standard parametric analyses directly applied to OS and PFS, which is commonly used in oncology indications to estimate long-term outcomes, was evaluated [4, 5]. When using the best-fitting curves (log-normal distribution for OS and spline-based model for PFS) in this analysis, the projected OS and PFS curves intersected, indicating that approximately 25% of patients die without progressing. Clinical opinion suggested that estimates of OS calculated using standard parametric analysis were below those expected with avelumab. This is consistent with findings for IOs in other indications, which noted that single parametric curves were insufficient to model the changes in hazards over time [7–9]. Therefore, use of standard parametric survival analysis to project long-term OS was deemed to underestimate the survival benefit associated with avelumab.

To address the underlying heterogeneity in patient responses to treatment, alternative modeling methods were considered. Using a response-based landmark analysis, projected survival estimates in patients who did not

achieve response were found to be as expected by a practicing MCC oncologist. However, the estimated OS for patients who achieved response (assuming an excess HR of 4.5 vs. the general population) was higher than expected. Projections implied that 91% and 79% of patients who achieve response are alive at 1.5 and 3 years, which was deemed to be an overestimate of the potential benefits associated with avelumab at that specific time point, given that data with a minimum follow-up of 18 months indicated that overall survival was approximately 85% among responders at 1.5 years [28, 29]. However, further follow-up will provide insight into whether the assumptions adopted were appropriate. Similarly, the plausibility of approximately 20% of patients not progressing, based on the projected OS and PFS curves intersecting after 74 months, requires further evaluation. The application of this approach was limited by sparse OS data available for patients who achieved response. If more events were observed in patients who achieved response or if longer follow-up data were used to estimate OS, it is possible that parametric survival distributions could be fitted, and more reasonable long-term projections could be obtained for patients who respond to avelumab. As the landmark point was chosen to maximize follow-up, two patients who achieved response beyond the landmark point were considered as non-responders in our analyses, which may partially explain the poorer fit using this approach between 6 and 15 months. Though modeling time to response and survival as a function of time-dependent predictors of response [34] may overcome these misclassifications imposed by the choice of a specific landmark point, such analyses remain an area for future research.

A third approach was based on estimation of the OS as a function of TTP, PFS, and PPS. This was the only approach where projected OS and PFS curves did not intersect, as the curves were not independently estimated. To some extent, use of progression as a surrogate endpoint considered improved survival patterns in patients who achieved response, since the majority of patients who did not progress in the JAVELIN Merkel 200 trial had achieved response
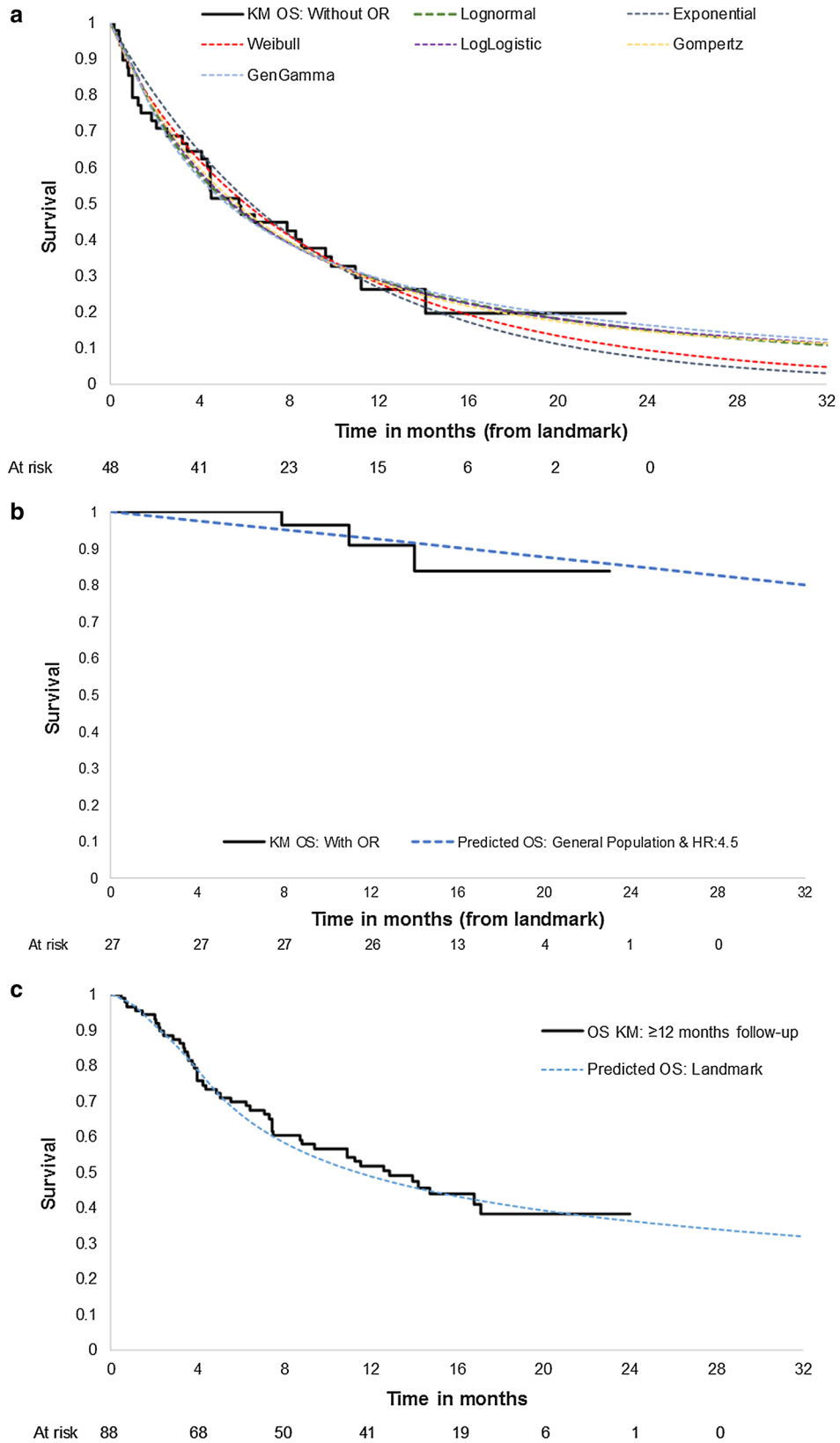
◄**Fig. 2** Projected OS in patients **a** without response and **b** with response using landmark/response-based approach. **c** Projected OS using landmark/response-based approach

(approximately 76%) [22]. This may be a plausible scenario in highly immunogenic tumors where response achieved with IO therapies is deemed durable [35]. The implicit assumption that was adopted when using this approach was that patients with response or stable disease follow mortality patterns observed in the general population unless they progress. The PFS + PPS approach implied that 85% and 66% of patients who achieved response would be alive at 1.5 and 3 years, lower than estimates of the response-based approach but more aligned with clinical expectations. As a result, at the time the analysis was conducted, on the basis of the data available, the PFS + PPS approach was considered most plausible from a clinical perspective. Interestingly, this approach resulted in almost identical projections to a published analysis that fitted spline models directly to OS on the data cut of the JAVELIN Merkel 200 trial with at least 18 months of follow-up (see ESM Fig. 6), indicating that the assumptions adopted in our analysis were aligned with the changes in hazard of OS over time [36]. These spline-based estimates, used within the recent assessment of avelumab by NICE [37], were considered to be appropriate and subsequently accepted by the committee. Nonetheless, longer-term estimates were noted to be highly uncertain.

Despite the PFS + PPS approach being considered most clinically plausible at the time the analyses were conducted, and closely resembling the projections using alternative methods receiving acceptance by NICE [37], comparison with OS curves with minimum follow-up of 24 months [33] indicated a poorer fit at the tail of the curve. This may be explained by the relationship between time of progression and PPS, which was not captured in our analysis. For example, patients experiencing progression earlier or experiencing hyper-progression may have alternative PPS as compared to patients who progress at a later point in time. We considered this to be an area of further research, as

further data cuts and follow-up become available to aid in informing PPS among patients who progress at later time points.

The response-based landmark approach, which made assumptions regarding the long-term mortality associated with patients with response, appeared to capture the tail of the curve better compared to the PFS + PPS approach. Although the landmark approach predicted higher survival in the first 8 years, the PFS + PPS approach resulted in the highest estimates of mean OS due to higher projected survival beyond 8 years, driven by the reduction in the risk of progression over time. This resulted in a declining excess mortality vs. the general population over time in contrast to the response-based approach, which assumed that excess mortality vs. the general population was constant. Although the landmark approach appears to fit the tail of the curve better, it remains unclear if increased mortality vs. the general population would apply over a patient's lifetime or whether a decline would be observed because of the plateau in PFS. After a minimum follow-up of 2 years in the JAVELIN Merkel 200 study, patients with mMCC continue to respond to treatment with avelumab and have meaningful survival outcomes [33]. However, additional longer-term follow-up data or real-world evidence are required to validate these long-term projections and determine which approach results in the most accurate estimates.

Limitations

Results from our analyses need to be evaluated while taking into consideration limitations associated with a modeling exercise based on single-arm trial data. Our analyses focused on projecting OS for avelumab based on a single-arm phase II study. Therefore, in absence of head-to-head data, the incremental value of avelumab versus standard of care and how this may differ when utilizing alternative approaches to project OS were not examined. As reimbursement decisions are formed on the basis of incremental value, further studies are required to ascertain how the use of alternative
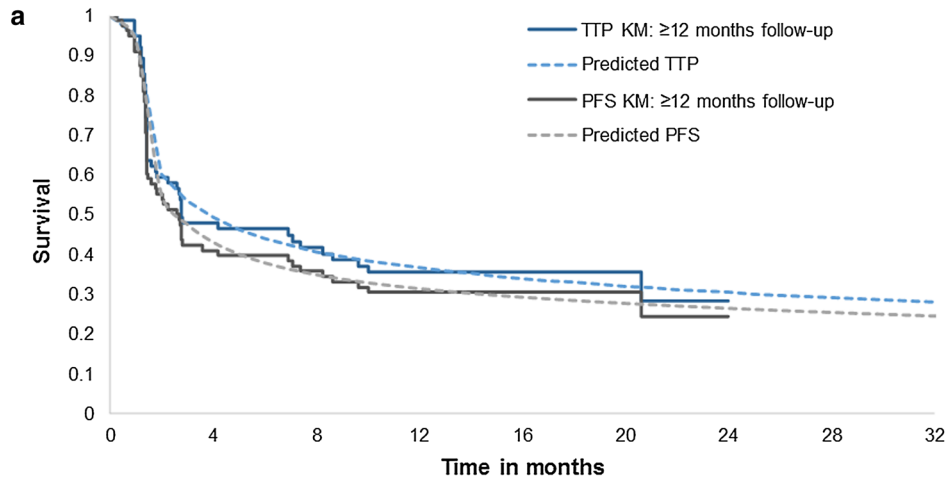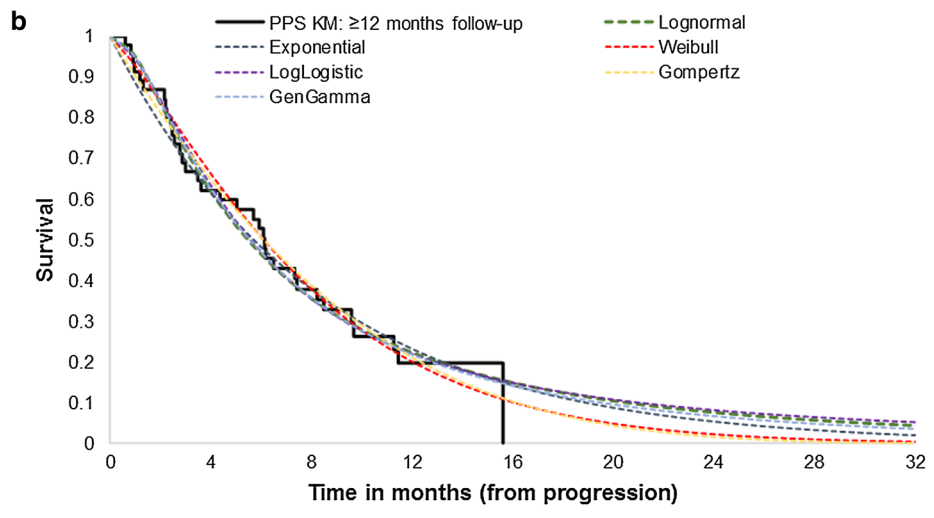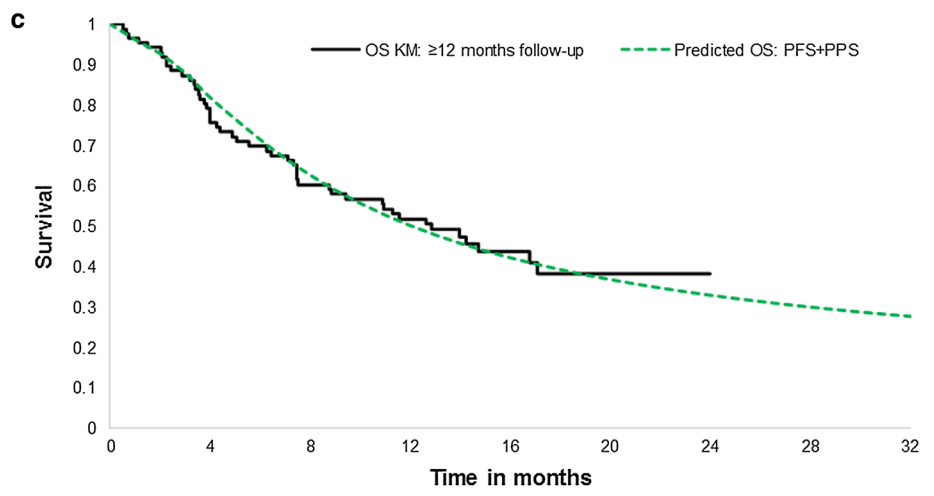
◄**Fig. 3 a** Projected PFS and TTP using a spline-based model. **b** Projected PPS using standard parametric survival analysis. **c** Projected OS using the PFS + PPS approach

projection methods may alter estimates of incremental value.

Though the primary analyses were conducted on the largest trial in MCC patients to date, analyses were based on a relatively small cohort of 88 patients resulting in considerable uncertainty in model parameters and projections. Though predictive validation was conducted, this was limited to a comparison versus the same data set used to derive projections, albeit with longer follow-up. To fully ascertain the accuracy of projections, validation versus an independent cohort would have been preferable; however, such data from the real-world or another single-arm trial were not available at the time of this study.

To alleviate uncertainties associated with long-term projections, expert opinion was solicited to comment on the plausibility of the projections. Within such a rare condition, there are a limited number of physicians who see

enough patients to be able to provide validation on long-term projections. Therefore, projections were evaluated with one practicing oncologist, which may not be sufficient. In the absence of longer-term data, physicians' perceptions on the long-term survival potential of treatments are inevitably formed on the basis of experiences within their practices/centers, thus the distributional choices informed on the basis of clinical opinion in our analyses would be limited by potential subjective judgements. However, clinical opinion alongside visual assessment of fit and goodness-of-fit statistics was utilized to select the distributions; where these were similar, there was limited evidence to preferentially select one approach over the other.

Future Directions

The application of alternative approaches to project OS in this case study is important in the context of innovative oncology drugs coming to market earlier, often based on short-term data and/or surrogate endpoints (sometimes from single-arm clinical trials) [38]. At the same time, insights or estimation of long-term

Table 1 Projected OS rates over 10 years with avelumab

Approaches to OS projection	Alive at different time points (%)				Average life expectancy, years	Probabilistic mean	95% percentiles
	2 years	3 years	5 years	10 years			
Standard parametric survival analysis[a]	32%	22%	12%	5%	2.48	2.53	1.58–3.75
Response-based landmark[b]	36%	30%	23%	11%	3.15	3.17	2.34–4.17
Patients with response	87%	79%	64%	32%	7.53[c]	7.60	6.46–9.04
Patients without response	17%	10%	5%	2%	1.22[c]	1.26	0.70–2.10
PFS + PPS[d]	33%	26%	19%	11%	3.54	3.64	2.43–5.16

OS overall survival, *PFS* progression-free survival, *PPS* post-progression survival

[a] Extrapolated for overall population using a log-normal distribution

[b] Extrapolated for responders using general population mortality and an HR of 4.5, and non-responders using a log-normal distribution

[c] Reflects life expectancy beyond landmark point of 89 days

[d] Time to progression and progression free-survival extrapolated for overall population using spline-based models. Post-progression survival extrapolated using log-normal distribution
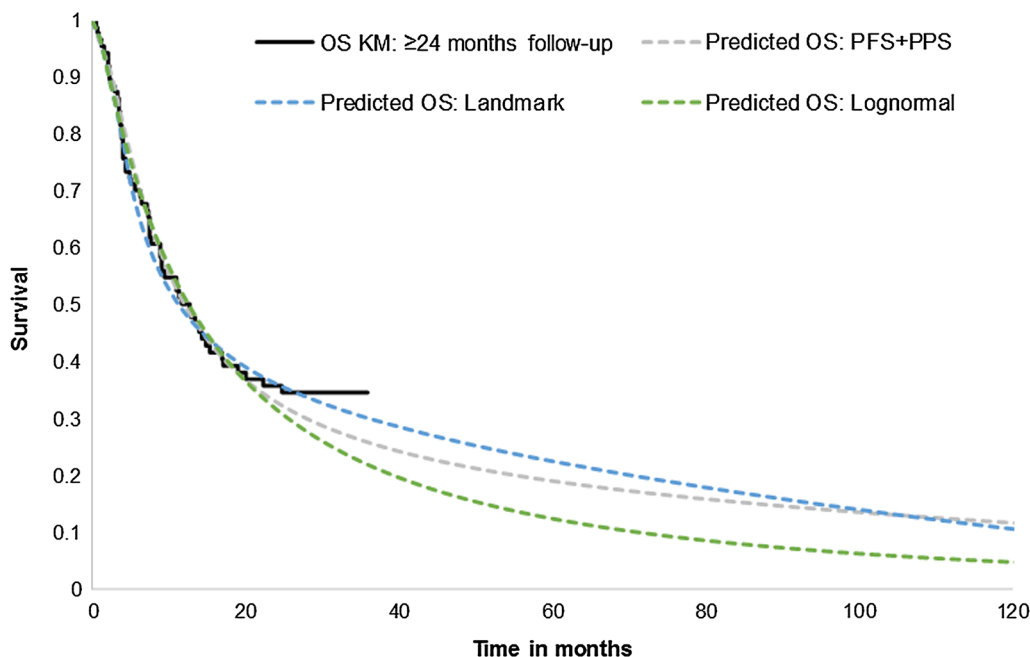
Fig. 4 Projected OS using alternative approaches

survival benefits are important for those making population healthcare decisions. As demonstrated in this case study the inadequacy of traditional parametric survival approaches can underestimate survival, which can result in negative consequences including misinformed reimbursement decisions or a delay in patient access to treatment. Therefore, it is important to explore alternative methods and follow the International Society for Pharmacoeconomic and Outcomes Research guidelines to identify the most appropriate approach based on the best information available at the time [32].

On the basis of this case study, analysis based on surrogate outcomes (i.e., response and progression) appeared to provide more plausible long-term extrapolations as compared to standard parametric analysis. An additional advantage of these approaches is in interpretation and understanding of the results. For example, the choice of a log-normal distribution to project OS directly may be more challenging to justify and interpret from a clinical perspective. Extrapolations using surrogate outcomes provide a rationale on survival projections. In our case study, survival patterns were considerably different among patients who responded to

treatment and among patients who had not progressed, thus enhancing our understanding of results and supporting regulatory and early access decisions based on surrogate endpoints. Nonetheless, there is substantial uncertainty regarding whether the results are generalizable to other indications and treatments. Further research and applications of these approaches across other indications and treatments are required to determine the generalizability of our findings.

CONCLUSIONS

Three different approaches to predicting OS in patients with mMCC treated with avelumab demonstrated good short-term predictive accuracy. However, estimates for long-term survival varied considerably. Therefore, attempting various extrapolation approaches can aid in understanding survival patterns and sensitivity of conclusions from economic analyses. Results from such analyses can help early decision-making based on immature data. The extrapolation results need to be interpreted on the basis of the advantages, limitations, and

appropriateness of alternative methods, as well as the disease and treatment context. In the context of an immunotherapy for a rare disease, where sample size and maturity will be inevitably low, an analysis based on surrogate outcomes may provide a good alternative, pending further evidence generation to ascertain generalizability across indications and treatments.

ACKNOWLEDGEMENTS

REFERENCES

1. Hawkins N, Grieve R. Extrapolation of survival data in cost-effectiveness analyses: the need for causal clarity. Med Decis Making. 2017;37(4):337–9.

2. Pannier A, Dunn JD. AMCP Format for Formulary Submissions, Version 40. J Manag Care Spec Pharm. 2016;22(5):448.

3. National Institute for Health and Care Excellence (NICE). Guide to the methods of technology appraisal 2013. https://www.nice.org.uk/process/pmg9/chapter/introduction. Accessed 29 October 2018.

4. Latimer N. 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):743–54.

5. Ishak KJ, Kreif N, Benedict A, Muszbek N. Overview of parametric survival analysis for health-economic applications. Pharmacoeconomics. 2013;31(8):663–75.

6. Logviss K, Krievins D, Purvina S. Characteristics of clinical trials in rare vs common diseases: a register-based Latvian study. PLoS One. 2018;13(4):194494.

7. National Institute for Health and Care Excellence (NICE). Nivolumab for treating advanced (unresectable or metastatic) melanoma. NICE technology appraisal guidance TA384. 2016. https://www.nice.org.uk/guidance/ta384. Accessed 8 Aug 2018.

8. National Institute for Health and Care Excellence (NICE). Pembrolizumab for advanced melanoma not previously treated with ipilimumab. NICE technology appraisal guidance TA366. 2015. https://www.nice.org.uk/guidance/ta366. Accessed 8 Aug 2018.

9. National Institute for Health and Care Excellence (NICE). Pembrolizumab for treating advanced melanoma after disease progression with ipilimumab. NICE technology appraisal guidance TA357. 2015. https://www.nice.org.uk/guidance/ta357. Accessed 8 Aug 2018.

10. Kepplinger E. FDA's expedited approval mechanisms for new drug products. Biotechnol Law Rep. 2015;34(1):15–37.

11. U.S. Food and Drug Administration, Center for drug evaluation and Research. Advancing health through innovation—2017 new drug therapy approvals. January 2017. https://www.fda.gov/downloads/AboutFDA/CentersOffices/OfficeofMedicalProductsandTobacco/CDER/ReportsBudgets/UCM591976.pdf. Accessed 8 Aug 2018.

12. Ades F, Yamaguchi N. WHO, RECIST, and immune-related response criteria: is it time to revisit pembrolizumab results? Ecancermedicalscience. 2015;9:604.

13. Wolchok JD, Hoos A, O'Day S, et al. Guidelines for the evaluation of immune therapy activity in solid tumors: immune-related response criteria. Clin Cancer Res. 2009;15(23):7412–20.

14. Chen TT. Statistical issues and challenges in immuno-oncology. J Immunother Cancer. 2013;1:18.

15. Gibson E, Koblbauer I, Begum N, et al. Modelling the survival outcomes of immuno-oncology drugs in economic evaluations: a systematic approach to data analysis and extrapolation. Pharmacoeconomics. 2017;35(12):1257–70.

16. Briggs A, Haddad R, Muszbek, Zhang Y. Developing cost-effectiveness models to assess value of immuno-oncology therapies: challenges and approaches. International Society for Pharmacoeconomics and Outcomes Research, 22nd Annual International Meeting. May 20–24, 2017; Boston, MA: Workshop 3.

17. Zhang Y, Mukhopadhyay P, Briggs A, Latimer N, Huang Min. Determining the value of long term outcomes associated with immuno-oncology therapies—challenges and approaches for OS extrapolations. International Society for Pharmacoeconomics and Outcomes Research 20th Annual European Congress. November 4–8, 2017; Glasgow, Scotland: Workshop 11.

18. Ouwens M, Mukhopadhyay P, Zhang Y, Huang M, Latimer N, Briggs A. Estimating lifetime benefits associated with immuno-oncology therapies: challenges and approaches for overall survival extrapolations. PharmacoEconomics. 2019. https://doi.org/10.1007/s40273-019-00806-4.

19. Ishak KJ, Muszbek N, Proskorovsky I, Zhang Y. Projecting survival with immuno-oncology

therapies: new challenges and new approaches. International Society for Pharmacoeconomics and Outcomes Research, 19th Annual European Congress. 29 October–2 November 2016, Austria, Vienna.

20. Ishak KJ, Proskorovsky I, Muszbek N. Projecting overall survival with immuno-oncology treatments. The Evidence Forum. 2016. https://www.evidera.com/wp-content/uploads/2016/10/Projecting-Overall-Survival-with-Immuno-Oncology-Treatments.pdf. Accessed 25 June 2019.

21. U.S. Food and Drug Administration. Avelumab (Bavencio). https://www.fda.gov/drugs/informationondrugs/approveddrugs/ucm547965.htm. Accessed 29 Oct 2018.

22. Kaufman HL, Russell JS, Hamid O, et al. Updated efficacy of avelumab in patients with previously treated metastatic Merkel cell carcinoma after ≥ 1 year of follow-up JAVELIN Merkel 200, a phase 2 clinical trial. J Immunother Cancer. 2018;6(1):7.

23. Kaufman HL, Russell J, Hamid O, et al. Avelumab in patients with chemotherapy-refractory metastatic Merkel cell carcinoma: a multicentre, single-group, open-label, phase 2 trial. Lancet Oncol. 2016;17(10):1374–85.

24. Cowey CL, Mahnke L, Espirito J, Helwig C, Oksen D, Bharmal M. Real-world treatment outcomes in patients with metastatic Merkel cell carcinoma treated with chemotherapy in the USA. Future Oncol. 2017;13(19):1699–710.

25. Othus M, Bansal A, Koepl L, Wagner S, Ramsey S. Accounting for cured patients in cost-effectiveness analysis. Value Health. 2017;20(4):705–9.

26. Ishak J, Villalobos CF, Proskorovsky I. Projecting survival with cure mixture models: when are the data mature enough for reliable analysis? Value Health. 2018;21(Suppl 1):PCP24.

27. Latimer NR. National Institute for Health and Care Excellence (NICE); Mar 2013. NICE DSU Technical Support Document No. 14. NICE Decision Support Unit Technical Support Documents. Survival analysis for economic evaluations alongside clinical trials—extrapolation with patient-level data. http://nicedsu.org.uk/technical-support-documents/technical-support-documents. Accessed 8 Aug 2018.

28. D'Angelo SP, Hunger M, Brohl AS, et al. Early objective response to avelumab treatment is associated with improved overall survival in patients with metastatic Merkel cell carcinoma. Cancer Immunol Immunother. 2019;68(4):609–18.

29. D'Angelo SP, Russell JS, Bhatia S, et al. 18-month efficacy and safety update from JAVELIN Merkel 200 part A. a phase II study of avelumab in metastatic Merkel cell carcinoma progressed on chemotherapy. J Clin Oncol. 2018;36(5_suppl):192.

30. The Human Mortality Database. United States Mortality Database (2014). http://www.mortality.org/. Accessed 10 Oct 2017.

31. Bharmal M, Fofana F, Barbosa CD, et al. Psychometric properties of the FACT-M questionnaire in patients with Merkel cell carcinoma. Health Qual Life Outcomes. 2017;15(1):247.

32. Eddy DM, Hollingworth W, Caro JJ, Tsevat J, McDonald KM, Wong JB. Model transparency and validation: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force-7. Value Health. 2012;15(6):843–50.

33. Nghiem P, Bhatia S, Brohl AS, et al. Two-year efficacy and safety update from JAVELIN Merkel 200 part A: a registrational study of avelumab in metastatic Merkel cell carcinoma progressed on chemotherapy. J Clin Oncol. 2018;36(5_suppl):9507.

34. Deniz B, Altincatal A, Ambavane A, et al. Application of dynamic modeling for survival estimation in advanced renal cell carcinoma. PLoS One. 2018;13(8):e0203406.

35. Hamid O, Robert C, Daud A, et al. Five-year survival outcomes for patients with advanced melanoma treated with pembrolizumab in KEYNOTE-001. Ann Oncol. 2019;30(4):582–8.

36. Bullement A, Amin A, Stapelkamp C, et al. Modelling overall survival in immunotherapy using parametric techniques: avelumab in previously treated metastatic Merkel cell carcinoma. International Society for Pharmacoeconomics and Outcomes Research 21st Annual European Congress. November 1014, 2018, Barcelona, Spain.

37. National Institute for Health and Care Excellence (NICE). Avelumab for treating metastatic merkel cell carcinoma. NICE technology appraisal guidance TA517. 2018. https://www.nice.org.uk/guidance/ta517/. Accessed 8 Aug 2018.

38. Alexiou D, Chatzitheofilou I, Pi Blanque A. A review of NICE technology appraisals in oncology using single arm trials evidence. Value Health. 2018;21(Suppl 1):S224. https://doi.org/10.1016/j.jval.2018.04.1516.