Article:
Cranmer, H., Shields, G.E. and Bullement, A. (2020) A comparison of partitioned survival analysis and state transition multi-state modelling approaches using a case study in oncology. Journal of Medical Economics, 23 (10). pp. 1176-1185. ISSN: 1369-6998

https://doi.org/10.1080/13696998.2020.1796360

JME 10

Journal of Medical Economics

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/ijme20

A comparison of partitioned survival analysis and state transition multi-state modelling approaches using a case study in oncology

Holly Cranmer , Gemma E. Shields & Ash Bullement

To cite this article: Holly Cranmer , Gemma E. Shields & Ash Bullement (2020) A comparison of partitioned survival analysis and state transition multi-state modelling approaches using a case study in oncology, Journal of Medical Economics, 23:10, 1176-1185, DOI: 10.1080/13696998.2020.1796360

To link to this article: https://doi.org/10.1080/13696998.2020.1796360

JOURNAL OF MEDICAL ECONOMICS
2020, VOL. 23, NO. 10, 1176–1185
https://doi.org/10.1080/13696998.2020.1796360
Article 0102-RT.R1/1796360

Taylor & Francis
Taylor & Francis Group

ORIGINAL RESEARCH

A comparison of partitioned survival analysis and state transition multi-state modelling approaches using a case study in oncology

Holly Cranmer[a], Gemma E. Shields[b] and Ash Bullement[c,d]

[a]Takeda Pharmaceuticals International Co., London, UK; [b]Faculty of Biology, Medicine, and Health, Division of Population Health, Health Services Research, and Primary Care, School of Health Sciences, Manchester Centre for Health Economics, University of Manchester, Manchester, UK; [c]Delta Hat Limited, Nottingham, UK; [d]School of Health and Related Research, University of Sheffield, Sheffield, UK

ABSTRACT

Aims: To construct and compare a partitioned-survival analysis (PartSA) and a semi-Markov multi-state model (MSM) to investigate differences in estimated cost effectiveness of a novel cancer treatment from a UK perspective.

Materials and Methods: Data from a cohort of late-stage cancer patients ($N > 700$) enrolled within a randomized, controlled trial were used to populate both modelling approaches. The statistical software *R* was used to fit parametric survival models to overall survival (OS) and progression-free survival (PFS) data to inform the PartSA (package "flexsurv"). The package "mstate" was used to estimate the MSM transitions (permitted transitions: (T1) "progression-free" to "dead", (T2) "post-progression" to "death", and (T3) "pre-progression" to "post-progression"). Key costs included were treatment-related (initial, subsequent, and concomitant), adverse events, hospitalizations and monitoring. Utilities were stratified by progression. Outcomes were discounted at 3.5% per annum over a 15-year time horizon.

Results: The PartSA and MSM approaches estimated incremental cost-effectiveness ratios (ICERs) of £342,474 and £411,574, respectively. Scenario analyses exploring alternative parametric forms provided incremental discounted life-year estimates that ranged from +0.15 to +0.33 for the PartSA approach, compared with −0.13 to +0.23 for the MSM approach. This variation was reflected in the range of ICERs. The PartSA produced ICERs between £234,829 and £522,963, whereas MSM results were more variable and included instances where the intervention was dominated and ICERs above £7 million (caused by very small incremental QALYs).

Limitations and conclusions: Structural uncertainty in economic modelling is rarely explored due to time and resource limitations. This comparison of structural approaches indicates that the choice of structure may have a profound impact on cost-effectiveness results. This highlights the importance of carefully considered model conceptualization, and the need for further research to ascertain when it may be most appropriate to use each approach.

ARTICLE HISTORY
Received 21 May 2020
Revised 3 July 2020
Accepted 9 July 2020

KEYWORDS
Cost-effectiveness; multi-state model; partitioned survival; decision-analytic model; oncology

JEL CLASSIFICATION CODES
I00; D61; H51

Introduction

In 2017, there were an estimated 24.5 million incident cases of cancer globally and 9.6 million cancer deaths[1]. It is the second leading cause of death globally and cancer deaths are predicted to rise globally to 16.3 million by 2040[2,3]. Cancer can be severely debilitating (particularly for those with progressed disease), often with a profound impact on patient and carer quality of life[4,5]. In addition, cancer is often associated with substantial financial burden causing distress for patients, caregivers, and dependents[6,7].

Since the turn of the century, there has been a rapid development in the range of innovative treatments available to treat patients with cancer by improving survival and quality of life. However, such treatments often come at a high cost. With an ever-increasing demand for new, effective treatments within the constraints of a finite healthcare budget, decision modelling plays an important role in the estimation of the value of these new cancer treatments.

Cost-effectiveness analysis (CEA) provides decision makers with an objective basis from which decisions may be informed. Typically, a CEA involves the development of an economic model to synthesize the available evidence concerning costs and effects in order to compare alternative treatment strategies. The most common model structures constructed in the field of oncology are partitioned survival analyses (PartSA) and state transition models (STMs), which are frequently based on three health states relevant to cancer: pre-progression, progressed disease and death[8,9].

In the three-state structure, PartSA requires two outcomes (progression-free survival (PFS) and overall survival (OS)) to

CONTACT Holly Cranmer holly.cranmer@takeda.com Health Economics Lead, Global Patient Value & Access, Oncology Business Unit, Takeda Pharmaceuticals International Co, 61 Aldwych, London, WC2B 4AE, United Kingdom

Supplemental data for this article can be accessed at https://doi.org/10.1080/13696998.2020.1796360.

inform health state occupancy, with time in the progressed disease state inferred through the difference between the two outcomes. PFS and OS outcomes are often readily available from the literature and are widely understood by clinicians and other stakeholders. This makes the PartSA a practical model choice, which is likely to be one reason that this structure was found to be the most common structure applied for cancer treatments in health technology assessments (HTAs) submitted to the National Institute for Health and Care Excellence (NICE) – the HTA body for England and Wales[8,9].

Conversely, STMs require specific information relating to the three transitions possible between the health states. When patient-level data are available this requirement is simple to fulfil. However, often these data are unavailable for comparators outside of a pharmaceutical company's clinical trial. As an additional complexity, STMs can be further divided based on discrete-time/continuous-time, Markov/semi-Markov or cohort/patient level. Multi-state modelling (MSM) falls under the STM bracket and could be considered when there are a series of competing events and when these events occur sequentially. In the three-state cancer model, progression and death are competing events and could occur sequentially, thus MSM is a suitable modelling method for consideration in cancer. The MSM approach models each of the transitions of interest simultaneously and uses a continuous-time framework.

A review of published CEAs in cancer found that despite the modelling structures available, typically only one is presented with limited explanation as to the justification and validation of this choice[8]. This is despite guidance from the Decision Support Unit (DSU) supporting NICE in UK HTAs stating that "state transition modelling should be used alongside the PartSA approach to assist in verifying the plausibility of PartSA's extrapolations and to address uncertainties in the extrapolation period"[9].

There is limited research considering the impact of different structural assumptions within economic models of cancer treatments. Studies that have been published highlight a discrepancy in results across the different approaches, suggesting that model structure may influence conclusions of clinical- and cost-effectiveness:

- Williams et al. (2017) considered a case study in first-line chronic lymphocytic leukemia comparing outcomes from a PartSA, a discrete-time semi-Markov STM and a continuous-time semi-Markov (i.e. MSM) STM – results approximated ICERs of: £16,000, £13,000 and £29,000 respectively[10].
- Degeling et al. (2018) compared the use of a cohort discrete-time STM with a discrete event simulation STM in patients with metastatic colorectal cancer; estimated ICERs were €172,443 and €168,383, respectively[11].
- Gibson et al. (2019) and Gibson et al. (2018) considered a PartSA, a discrete-time Markov STM and a patient level simulation STM to assess the value of immuno-oncology therapies in metastatic melanoma[12,13]. Both studies also explored extending the standard three state oncology model by including an immune-specific health state. Results were discrepant between the different model structures with incremental cost-effectiveness ratios (ICERs) varying between £6,474 and £49,000
- Smare et al. (2020) compared the use of a PartSA to two variations of a semi-Markov STM for a treatment for renal cell carcinoma and found that model structure varied estimated survival benefit by up to 14%[14].

The studies published to date emphasize the importance of justifying and validating the choice of model structure. Ideally, for each CEA, all suitable model structures should be considered, and results presented with an explanation as to which best reflects the dynamics of the disease and treatment pathway, with clinician input and external data sources serving as a critical source of validation. However, there are many reasons why this may not be standard practice; such as, feasibility constraints (e.g. time and funding), challenges acquiring patient-level data, and a lack of understanding relating to the structural assumptions underpinning different model structures.

The aim of this paper is to compare a PartSA and a semi-Markov MSM STM approach as methods for estimating the cost-effectiveness of a novel treatment compared to the standard of care within the context of late-stage cancer. This paper aims to add to the growing body of literature emphasizing the importance of justifying model structure and to explore why these differences occur.

Data and methods

Data used for extrapolation

To inform both modelling approaches, data were sought to populate the model transitions. While data could have been developed using simulation methods, data collected as part of a clinical trial were preferred in order to test the approaches using "true" data. Data to compare modelling approaches were provided to the authors under the proviso that the treatments compared were anonymized.

A case study comprising of data from a randomized controlled trial (RCT) comparing two treatments (TX1 vs. TX2) for a type of late-stage cancer was used to compare the two modelling frameworks (PartSA vs. MSM) – the average age of patients was between 60 and 70. The RCT considered a large cohort of patients ($N > 700$) and had a median follow-up of approximately 15 months. At the end of follow-up between 50–70% patients had progressed and between 20–30% had died across both the TX1 and TX2 arms, respectively.

Economic models

Both economic models were developed to compare the total costs and quality-adjusted life-years (QALYs) associated with TX1 and TX2 from a UK National Health Service (NHS) perspective, ultimately providing an ICER for TX1 vs. TX2. Costs and QALYs were discounted using a rate of 3.5%, in line with UK HTA requirements[15]. A 15-year time horizon was selected

such that all patients had died in each model framework and a monthly cycle length, across both modelled strategies.

Models were constructed in the statistical software R version 3.01[16]. The code used is provided in the Supplementary Material.

PartSA structure

The PartSA was characterized by three health states (pre-progression, progressed disease and death); this structure is the most commonly seen in cancer submissions to NICE in the UK[8,9]. State membership was determined by two independent survival curves (PFS and OS) that allow sub-division (or "partitioning") of the OS curve. Time dependency (i.e. the relationship between time spent in a health state and the probability of leaving that health state) is implicitly captured within the PartSA framework. The model structure is presented in Figure 1.

Joint parametric models were fitted to the independent PFS and OS outcomes using the *phreg, aftreg* and *flexsurvreg* functions in R (using *eha* and *flexsurv* packages, respectively); the proportional hazard assumption was considered appropriate following inspection of the log-cumulative hazard plots and the Schoenfeld residual plots (Supplementary Material). As per DSU guidance, six parametric distributions were fit using the individual patient-level data for each trial outcome: exponential, generalized gamma, Weibull, lognormal, log-logistic and Gompertz[17]. Goodness of fit was based on Akaike's Information Criterion (AIC), Bayesian Information Criterion (BIC) and visual comparison with Kaplan-Meier estimates.

Statistically, the generalized gamma and the lognormal provided plausible fits to the PFS data. Visual interpretation indicated the lognormal to predict an implausibly wide tail. The generalized gamma appeared to provide estimates that better aligned with other literature in this disease area. Therefore, the generalized gamma curve was selected to model PFS outcomes in the base case. The goodness of fit statistics indicated that the lognormal, log-logistic and Weibull parametric curves provided plausible fits to the OS data. Based on visual interpretation of the extrapolated curves and comparison with other literature in this disease area, the Weibull curve was applied in the base case. Out of the three highlighted by the goodness-of-fit statistics, the Weibull was the most pessimistic curve in terms of mean survival predicted. Therefore, this is considered a conservative assumption. Alternative parametric curves are explored in scenario analyses.

MSM structure

The MSM was also characterized by the same three health states: progression-free, progressed disease and death. The model structure is presented in Figure 2.

The MSM requires assessment of the Markovian assumption; this assumption refers to the memoryless feature of a Markov model i.e. transitions from a health state are independent of the duration of time spent in the currently-occupied or any previously-occupied health state(s). To assess the applicability of this assumption, a Markov Cox proportional hazards model was constructed. The model considered the transition from progression to death explained by the time spent in the previous health state. This covariate (time in the

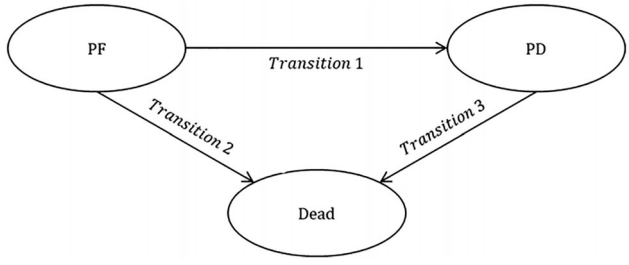

Figure 2. MSM model structure.

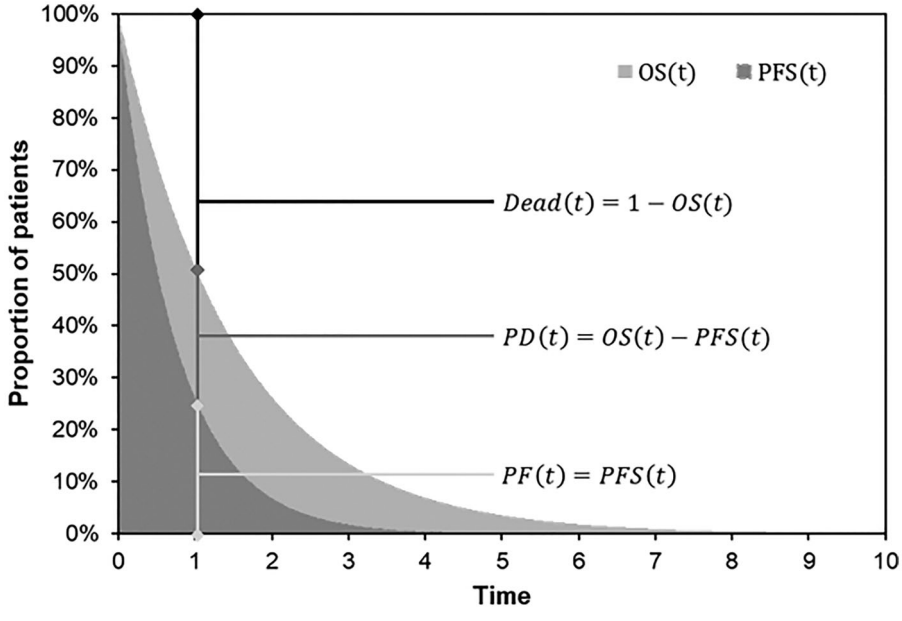

Figure 1. PartSA model structure.

previous state) was shown to be statistically significant ($p = .022$); results indicated a longer duration spent in the progression health state would increase the risk of death. Therefore, a semi-Markov approach was undertaken, wherein the time spent in the progression health state depends on time spent in the pre-progression health state.

In line with the PartSA approach, joint parametric models were fit to the data for each transition. Three survival models were estimated (transition 1 [T1]: progression-free to progressed disease; transition 2 [T2]: progression-free to death and transition 3 [T3]: progressed disease to death). The published code from Williams et al. (2017) was adapted to this dataset to estimate the transitions relevant to this MSM structure[18]. The aforementioned standard six distributions were considered for each transition. These methods are similar to the PartSA approach with one key difference: any observation where an event occurs which is not the event of interest for a specific transition is treated as a censored observation i.e. patients that experience competing events are treated in the same way as a patient that was lost to follow-up. For example, for the transition from progression-free to progressed [T1] any deaths reported in the data are not the event of interest and so they are censored.

The MSM approach considered a continuous time structure using the exact timing of transitions. However, for the purposes of estimating mean survival using the area under the curve approach, a monthly cycle length was applied. Due to computational issues with the generalized gamma and Gompertz distributions when fitting the transitions from progression-free to progressed disease, the calculation of transition probabilities with the MSM used cycle increments shorter than one month (up to 1/72 of a month) for these distributions up to the 15-year time horizon. This shortening of the cycle length was needed to overcome a difficulty in meeting the requirement that differences in cumulative hazards between consecutive time points were below one (as is aligned with the approach proposed by Williams et al. [2016])[18].

The AIC and BIC provide us with information as to how well the parametric curves fit the individual transitions. However, the transitions do not correspond to the state occupancy probabilities, which are defined by the competing risks of progression and death. Therefore, the AIC and BIC measures need to be interpreted with caution in an MSM framework as they do not account for the underlying relationships between the transitions; transitions defined by AIC/BIC score may not produce health state occupancies that provide a good fit to the data. Parametric curves were selected based on AIC, BIC and visual comparison with Kaplan–Meier estimates. The generalized gamma was selected for T1 and the Weibull was selected for T2 and T3. These base case parametric forms broadly align with those applied in the PartSA structure for PFS and OS outcomes, respectively. Alternative parametric curves are explored in scenario analyses.

Cost and utility inputs

Table 1 presents the cost and quality of life inputs informing the economic models. Costs applied within the models included: drug, adverse event, concomitant medication, hospitalization and subsequent therapy costs. It was assumed that all patients remained on treatment until progression. Therefore, all costs associated with treatment (drug, adverse event, concomitant medication and hospitalization costs) were accrued by patients in the pre-progression health states. Only drug and hospitalization costs were dependent on type of treatment. It was assumed that receipt of TX1 or TX2 did not impact choice of subsequent therapy. This is a simplification of the treatment pathway; in real-world clinical practice subsequent therapies may differ between the treatment arms. However, for the purposes of focusing on differences in model structure driving results, these costs have been assumed to be equal. All patients in the progressed disease health states accrued a weekly cost of subsequent therapy.

Table 1. Cost and utility inputs.

Parameter	Value*	Frequency	Health state applied
Drug costs			
TX1	£6,000	Every 4-weeks	Pre-progression
TX2	£4,000	Every 4-weeks	
Adverse events	£30	Weekly	
Concomitant medications	£50	Weekly	
Hospitalizations			
TX1	£100	Weekly	Pre-progression
TX2	£130	Weekly	
Subsequent therapy	£500	Weekly	Progressed disease
Utility values			
Pre-progression	0.8	Constant	Pre-progression
Progressed disease	0.6	Constant	Progressed disease

Note: Cost and utility inputs are arbitrary parameters that were included to broadly resemble values seen in a range of late-stage cancer models. Simple assumptions informed these parameters such that this analysis could focus on the comparison of model structures rather than model inputs.

In reality, some costs may differ within a given health state; for example: patients may discontinue treatment before documented disease progression or toxicity profiles may differ. However, these differential impacts were not explored so that the impact of model structure on outcomes could be clearly identified – without introducing additional differences from cost inputs. In addition, both of these examples are expected to predominantly affect costs and effects related to pre-progression disease which should be captured reasonably well by both the PartSA and MSM approaches.

Quality of life was captured through the application of health state specific utility values: 0.80 for pre-progression and 0.60 for progressed disease. These inputs were applied identically across both model structures. While inputs are based on an approximation from the literature, these inputs were considered to reflect patients with late-stage cancer treated with TX1 and TX2.

Comparison of approaches

To understand the differences in the model approaches, the occupancy of the model health states projected using each approach were compared using Markov traces. Through inspection of Markov traces, the proportion of patients residing within each health state may be established (regardless of how long patients have been within a given state). This approach allows for a more in-depth inspection of the

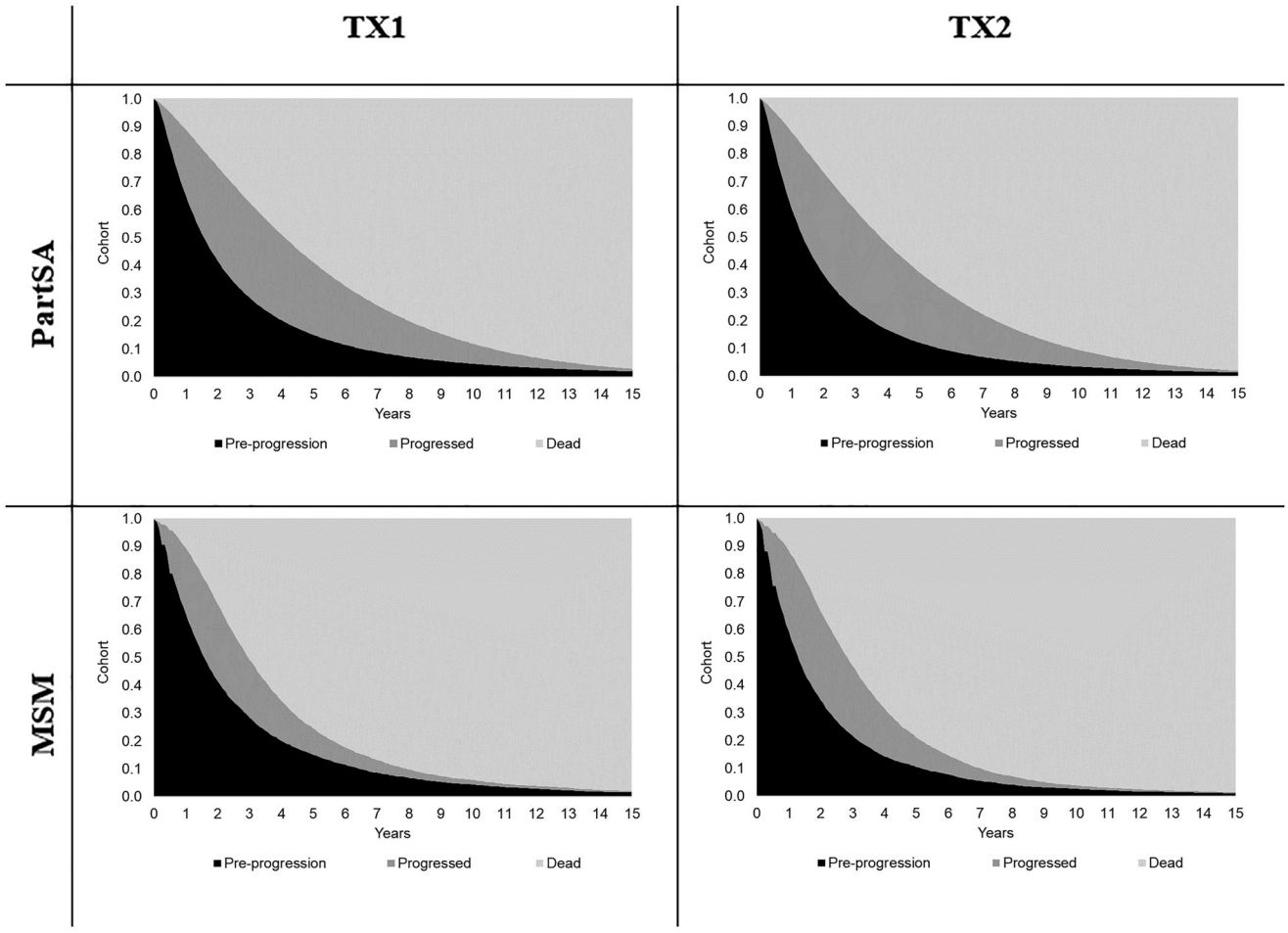

Figure 3. Markov traces for TX1 and TX2 from the PartSA (3a and 3b, respectively) and MSM (3c and 3d, respectively) approaches.

accrued life-years (LYs) in each model health state. After comparing health state occupancy, the probability of residing within a given state was plotted over time to establish differences in this outcome.

To contextualize the differences in the modelling approaches, the costs and utility inputs described above were used to produce CEA results. The values used to populate these results are informed by simple assumptions rather than robust data and are used to simply demonstrate how the modelling approaches will affect modelled outcomes.

Results

The fitted PFS and OS curves for TX1 and TX2 under the PartSA approach based on the generalized gamma and Weibull functions, respectively, are presented in the Supplementary Material. The Supplementary Material also presents the fitted transitions for TX1 and TX2 under the semi-Markov MSM approach.

Health state occupancy

Figure 3 presents the Markov traces for TX1 and TX2 for each model structure (PartSA and MSM). Figure 4 then presents the probability of residing in each of the health states over time for each model structure. Table 2 presents the undiscounted LYs accrued in each health state for each structure.

Each of the model structures show similar predictions for the within-trial period. However, after the end of follow-up, there are clear differences between the results associated with each structure; both in absolute terms (i.e. affecting estimates for each treatment individually) and relative terms (affecting the estimated incremental benefit).

The probability of residing in the pre-progression health state is similar between the two model structures for each treatment – this is partly explained by the same parametric form which has been assumed for PFS and for T1 under the two approaches (i.e. generalized gamma). While similar, the MSM predicts slightly fewer LYs for both TX1 and TX2 and a larger resulting difference in progression-free LYs (PFLYs) of 0.49 compared with 0.34 for the PartSA.

As seen for the progression-free state, the probability of residing in the progressed disease health state is similar from model baseline to 13 and 20 months, for TX1 and TX2, respectively. These timepoints coincide with the approximate times until which data are available from the trial. However, after these time points (i.e. in the extrapolation period), the probability of residing in this health state continues to increase under the PartSA structure and begins to decline under the MSM structure. The probability of residing in the progressed health state does eventually decline under the

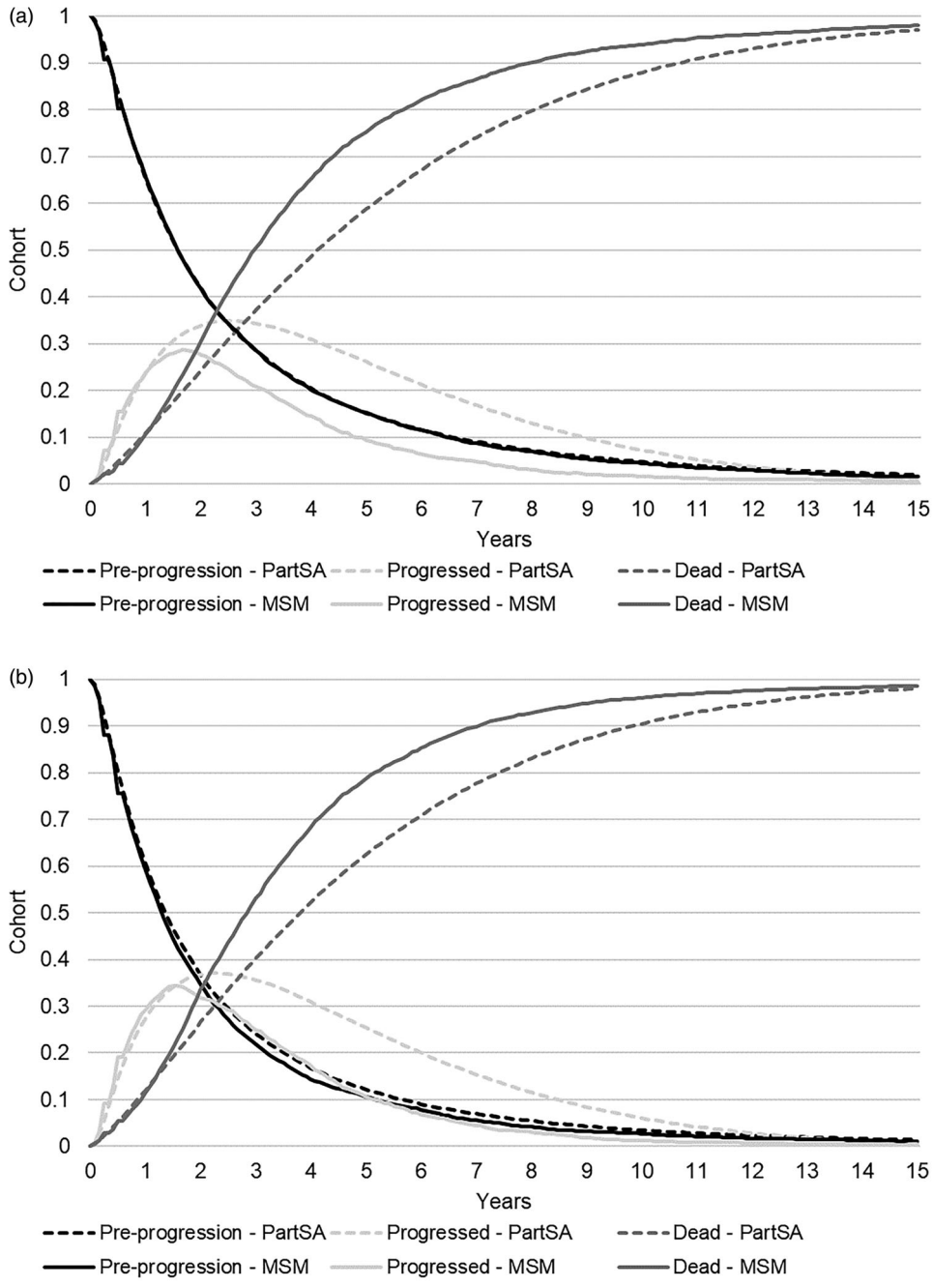

Figure 4. Probability of residing in each health state over time for TX1 (4a) and TX2 (4b) from the PartSA and MSM approaches, respectively.

Table 2. Undiscounted life-years.

	Pre-progression		Progressed disease		Total	
	PartSA	MSM	PartSA	MSM	PartSA	MSM
TX1	2.71	2.67	2.31	1.19	5.02	3.86
TX2	2.37	2.18	2.30	1.37	4.66	3.55
Incremental	0.34	0.49	0.01	−0.18	0.36	0.31

Abbreviation. MSM, multi-state model; PartSA, partitioned survival analysis.

PartSA but at a later time point and at a slower rate than the MSM structure.

The progressed disease LYs (PDLYs) are very different under the two different model structures; the PartSA predicts 2.31 and 2.30 PDLYs for TX1 and TX2, respectively; whereas the MSM predicts 1.19 and 1.37, respectively. The absolute values are smaller under the MSM approach and the direction is reversed i.e. more PDLYs are accrued for TX2 than TX1 under the MSM approach, whereas fewer PDLYs are accrued for TX2 than TX1 under the PartSA approach.

Finally, as would be expected, due to the differences arising in the progressed disease health state, the probability of remaining alive over time is higher under the PartSA approach compared with the MSM approach – as described by the total LYs: 5.02 and 4.66 for TX1 and TX2, respectively derived from the PartSA approach and 3.86 and 3.55 for TX1 and TX2, respectively derived from the MSM approach. The probability of being alive begins to diverge at 15 and 16 months for TX1 and TX2, respectively (again, approximately in line with the end of the observed data period).

Cost-effectiveness analysis results

Headline CEA results are presented in Table 3. Incremental costs are estimated as £78,045 and £78,199 for the PartSA and MSM approach, respectively; with incremental QALYs of 0.23 and 0.19, respectively. This results in ICERs of £3,42,474 (PartSA) and £4,11,574 (MSM).

The differences in the ICER are mostly driven by differences in incremental QALYs, which in turn are affected primarily by differences in the overall LYs, and the split of LYs between PFLYs and PDLYs estimated from the two model structures (as described above). Health state specific costs and QALYs are presented in Table 4.

While the incremental costs are similar, there are differences in terms of absolute total costs which will impact other economic outcomes (including budget impact) depending on which model structure is considered. The absolute costs estimated with the MSM structure are lower than the PartSA structure. This is due to a slightly smaller proportion of patients residing in the progression-free health state over time (accruing a lower total treatment cost), as well as a smaller proportion of patients who are alive over time (accruing fewer costs associated with any treatment and disease management).

Table 3. Headline CEA results.

Model approach	Total		Incremental		ICER
	Costs	QALYs	Costs	QALYs	
PartSA					
TX1	£265,693	3.14			
TX2	£187,648	2.91	£78,045	0.23	£3,42,474
MSM					
TX1	£239,499	2.57			
TX2	£161,300	2.38	£78,199	0.19	£4,11,574

Abbreviations. CEA, cost-effectiveness analysis; ICER, incremental cost-effectiveness ratio; MSM, multi-state model; PartSA, partitioned survival analysis; QALY, quality adjusted life year.

Table 4. Disaggregated model costs and outcomes.

Discounted outcomes	Pre-progression		Progressed disease		Total	
	PartSA	MSM	PartSA	MSM	PartSA	MSM
Costs						
TX1	£2,14,316	£2,11,674	£51,377	£27,825	£2,65,693	£2,39,499
TX2	£1,36,146	£1,28,387	£51,502	£32,913	£1,87,648	£1,61,300
QALYs						
TX1	1.96	1.93	1.18	0.64	3.14	2.57
TX2	1.73	1.63	1.18	0.76	2.91	2.38

Abbreviaions. CEA, cost-effectiveness analysis; ICER, incremental cost-effectiveness ratio; MSM, multi-state model; PartSA, partitioned survival analysis; QALY, quality adjusted life year.

Sensitivity analysis

A total of 36 scenarios were explored within the PartSA structure, based on all possible combinations of the six "standard" parametric curve choices applied to the OS and PFS data (i.e. 6×6). The incremental costs and QALYs derived from these scenarios are presented in Figure 5 – these scenarios yielded ICERs ranging from £2,34,829 to £522,963. PFLYs ranged from 1.63–2.59 for TX1 and 1.44–2.31 for TX2; PDLYs ranged from 0.78–4.78 for TX1 and 0.91–4.82 for TX2.

For the MSM structure, 216 scenarios were explored based on different parametric curve choices applied to T1, T2 and T3 (i.e. $6 \times 6 \times 6$). The incremental costs and QALYs derived from these scenarios are presented alongside the PartSA scenarios in Figure 5. These scenarios yielded results ranging from TX2 being dominant, to TX1 being associated with an ICER of £7,695,487. High ICERs are caused by very small incremental QALYs being estimated in some scenarios.

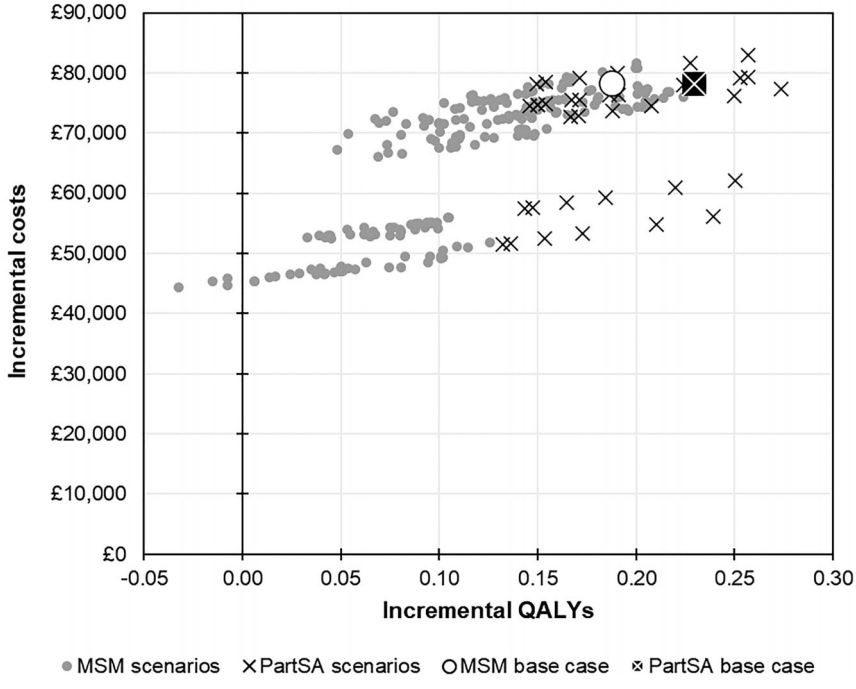

Figure 5. Scenario analyses associated with parametric forms.

PFLYs ranged from 1.58-2.51 for TX1 and 1.38-2.21 for TX2; PDLYs ranged from 1.05-1.97 for TX1 and 1.24-2.15 for TX2.

Figure 5 demonstrates a higher spread of uncertainty in terms of parametric curve selection from the MSM structure compared with the PartSA structure – this may be explained by the fact that the MSM structure specifies the model based on more specific data for each transition (three transitions vs. two endpoints in the PartSA structure using the same data source). Additionally, varying T1 has a "knock-on" effect within the model due to the embedded structural links; this transition directly impacts the proportion of patients eligible for T2 and T3. Therefore, varying the parametric curve informing T1 will likely have an amplified effect on the model results. The scenarios associated with the MSM structure are largely to the left of the scenarios associated with the PartSA structure, indicating that MSM scenarios predicted fewer incremental QALYs.

Summary of key differences

The key differences between the model structures are driven by differences in estimated outcomes beyond the duration of follow up. The MSM structure predicts a lower probability of being in the progressed disease state over time from 13 to 20 months for TX1 and TX2, respectively and a higher probability of being in the death health state from 15 to 16 months, respectively. These differences have implications for both the total costs and total QALYs accrued, which then go onto impact the ICER. The sensitivity analysis results illustrated a broad spread of estimated costs and QALYs, with the MSM scenarios generally predicting fewer incremental QALYs.

Discussion

The PartSA is the most commonly applied model structure in oncology within the UK. However, the limitations associated with this structure are often not acknowledged or explored thoroughly[8]. Use of MSMs in HTA are less common, a recent example considering the structural link between progression and death through MSM is described in the UK HTA NICE submission TA587[19]. The comparison of modelling approaches (PartSA vs. semi-Markov MSM) indicates that the choice of structure can have a profound impact on predicted outcomes and cost-effectiveness results which may subsequently impact reimbursement decisions made by HTA bodies. Given these differences it is important to understand the assumptions underpinning each structure.

The PartSA extrapolates PFS and OS independently; and so, mortality in this structure is only determined by time to death data and is not explicitly linked to earlier progression events. The assumption that the modelled survival endpoints are structurally independent is potentially problematic as there are a number of dependencies between the survival endpoints, for example: (1) they include some of the same events (e.g. PFS and OS curves include the same pre-progression deaths); (2) events are structurally dependent (e.g. death cannot be followed by progression and time spent progression-free contributes to time spent alive); and (3) intermediate events are often of prognostic importance for later events (e.g. progression is generally considered a negative prognostic factor for mortality)[9]. For the within-trial period, these dependencies are reflected in the data and should be closely reflected in the PartSA results. However, for analyses that model beyond the trial period, dependencies between endpoints are ignored with potentially important implications for extrapolation. Around 60% of patients in the dataset informing this research had progression events at data cut-off. Therefore, ignoring the dependences between endpoints is likely to impact the validity of extrapolated outcomes.

Conversely, the MSM approach models clinical events such that they are explicitly related. Note: there remains uncertainty within the MSM associated with extrapolating outcomes from immature data, for example: the probability of transitioning from progressed disease to death will encompass uncertainty if not all patients have progressed within the data set. Additionally, the MSM has the potential to model counterfactuals regarding the patterns of treatment post-progression which may offer a better reflection of the outcomes observed in clinical practice. In the three-state example presented in this paper, external data sources of post-progression survival could directly inform transition 3 of the MSM. Whereas, these data would have to be combined with the OS outcomes informing the PartSA.

The MSM approach is not without limitations. Currently, available analytical methods rely on access to individual-level data for the treatments of interest, which are unlikely to be available for published clinical studies. Furthermore, as with standard STMs, the MSM approach requires sufficient data to inform the transitions from progression-free to death and progressed disease to death. Data on these transitions specifically are often limited if the majority of deaths occur after disease progression or when follow-up is limited.

Within the context of a clinical trial, data following disease progression are often restricted due to limited follow-up. This limitation is sometimes used in defense of a model approach that does not require these data specifically (i.e. a PartSA). However, while a PartSA can (in theory) be fitted to any dataset where PFS and OS are available, this does not mean that the outcomes are robust.

From a feasibility stance, MSMs are objectively more complex and time-consuming to develop. However, the analysis and code relating to the *msm* R package is explained in detail by Williams et al. alongside a worked example[18]. Once this code is understood (requiring working knowledge of R programming), it is relatively simple to adapt and apply to other settings. To further encourage use and transparency in this area, we have also published our code in the Supplementary Material for both the PartSA and MSM frameworks.

A PartSA informed by sufficiently robust data should yield very similar outcomes to the MSM framework. However, where data are limited (e.g. due to administrative censoring), it is likely that each approach will yield different estimates of post-progression survival. Rigorous model calibration and validation from clinical experts can help to align the

post-progression survival with real world experiences. However, without more data, there is only so much model calibration can achieve. There are published guidelines available for choosing a model structure based on key requirements (including: output requirements, population size and system complexity)[20]. However, to date, there is no published validation tool to assess the relative fit of two or more model structures – this is an unmet need for model developers which should be addressed in future research.

The key differences between the model structures are driven by the estimated outcomes beyond the trial follow-up; the example presented in this paper considers a 15-year time horizon, given that the median follow-up of the trial is 15-months, predicted outcomes inform the majority of the model time horizon. The 15-year time horizon aligns with HTA guidelines for life-extending treatment i.e. the time horizon should be long enough to reflect all important differences in costs or outcomes between the technologies being compared[15]. Extrapolating beyond the trial follow-up introduces uncertainty into the model estimates. Therefore, shorter time horizons, with less uncertainty, may be considered in scenario analyses. However, it is important to note that these shorter time horizons would likely not reflect all the benefits or costs that would be accrued by the treatment and, as such, would not be the "true" ICER.

This analysis presents deterministic results. Probabilistic results are often important for decision making to assess the impact of uncertainty on the dispersion of results. However, these were outside of the scope of this example. In practice, is important that probabilistic analyses conducted within both the PartSA and MSM framework account for the correlation between endpoints and transition probabilities such that clinically implausible results are not generated (e.g. PFS curves crossing OS curves).

There is insufficient information available to inform our analysis in order for us to conclusively make a recommendation as to which of the presented models is the "least wrong", with the understanding that no health economic model is "right". It is likely that there is uncertainty introduced within the PartSA structure due to the immature survival data yet understanding the extent of this is difficult without further information on long-term outcomes. Similarly, transitions from the progressed disease to death health state in the MSM framework are based on limited follow-up from a subgroup of patients who have progressed within the clinical trial. Therefore, these data are also likely to lead to somewhat uncertain estimates of OS.

In the absence of longer follow-up from the trial, external data sources may be considered (where available). These sources can be used as a validation tool or directly built into the modelling framework. The MSM structure lends itself to the implementation of external data sources for the T2 transition, from progressed disease to death. However, given that outcomes are not modelled explicitly for patients with progressed disease in the PartSA approach, it is not possible to incorporate such external data in this framework.

Our analysis has used "real" data from a late-stage cancer clinical trial, simulating a "real-world" scenario where only immature data are available. The methods and results have been clearly explained such that the analyses can be easily repeated, additionally, the R code has been made available to encourage this. However, the research has limitations. Firstly, directional findings from our study are specific to the setting in which the clinical trial is set. Limited information has been provided on the disease area and the clinical trial due to confidentiality requirements from the pharmaceutical company. However, this does not impact the conclusions of the study in relation to model structures. Additionally, the research is limited to exploring a three-state oncology model structure for both PartSA and MSM – MSMs may be considered as a better option when the causal pathway is more complicated. The clinical data used in this study are limited. Having a longer follow-up from the clinical trial would provide more information as to which model is predicting the outcomes closest to reality. Future research could explore the impact of follow-up time on the robustness of the results; for example, censoring patients at shorter follow-up times or limiting the analyses to subgroups with different follow-up times. It would also be beneficial to re-visit these analyses with longer term data to see which model was predicting outcomes in line with observed data. In addition, we have only considered relatively "simple" parameterizations for each transition/curve used in both modelling approaches due to currently-available packages for the MSM approach.

Further research is required to understand which model should be used and when, in terms of different contexts and settings[8]. However, it is important that for now, we understand the implications of the different modelling methods and sufficiently explore these such that decisions remain evidence based and allow for the most efficient allocation of resources. In relation to MSM structures specifically, research should consider how to incorporate relative efficacy for treatments which only have OS and PFS outcomes reported in the literature. In terms of deciding the most appropriate model structure, it is important to acknowledge that no one approach will be without limitations. Therefore, we recommend that researchers state and discuss the assumptions and drawbacks featured for their chosen model structure(s) – the NICE Technical Support Documents provide a useful starting point as to what should be presented for each modelling approach[9]. Ideally, multiple model structures should be developed, and the relative advantages and limitations associated with each approach stated and explored in scenario analyses. Recommendations for decision makers using health economic models to inform allocative decisions are to explore how the assumptions underpinning the model structure may be influencing results and, where there are insufficient data to support the underlying structures, to be cautious when interpreting results.

Conclusions

This analysis adds to the growing literature demonstrating the importance of justifying the underlying model structure and exploring structural assumptions within scenario analyses. We recommend that where feasible a comparison of

model structures is made and further research to ascertain when it may be most appropriate to use each approach. In terms of contemporary HTA, we urge relevant parties to present a detailed account of the approach used alongside its associated strengths and limitations, acknowledging that other structures could have been considered.

Transparency

Declaration of funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Takeda Pharmaceuticals funded the fees associated with publication.

Declaration of financial/other interests

HC is a full-time employee at Takeda. No other significant or relevant relationships to declare.

JME peer reviewers on this manuscript have no relevant financial or other relationships to disclose.

Author contributions

HC conceived the idea for the commentary, with input from AB and GS. HC conducted the analysis and AB checked the analysis. HC, AB and GS co-wrote the manuscript. All authors read and approved the final manuscript.

Acknowledgements

No assistance in the preparation of this article is to be declared.

Previous presentations

We have previously presented a subset of our findings at the International Society for Pharmacoeconomics and Outcomes Research (ISPOR) conference in Copenhagen, November 2019 (Cranmer et al., 2019) [PCN103. A Comparison of Partitioned-Survival and State Transition Modelling Approaches – Findings from a Case Study in Oncology]) and at the International Health Economics Association (iHEA) conference in Basel, 2019.

ORCID

Holly Cranmer http://orcid.org/0000-0002-9067-3422
Gemma E. Shields http://orcid.org/0000-0003-4869-7524
Ash Bullement http://orcid.org/0000-0001-7091-0972

References

[1] Fitzmaurice C, Abate D, Abbasi N, Global Burden of Disease Cancer Collaboration, et al. Global, regional, and national cancer incidence, mortality, years of life lost, years lived with disability, and disability-Adjusted life-years for 29 cancer groups, 1990 to 2017: a systematic analysis for the global burden of disease study. JAMA Oncol. 2019;5(12):1749–1768.

[2] Bray F, Shield KD. Cancer: global burden, trends, and projections. In: International encyclopedia of public health. Elsevier Inc.; 2016. p. 347–368.

[3] American Cancer Society. Global cancer facts & figures 4th edition [Internet]. Am Cancer Soc. 2018. Available from: https://www.cancer.org/content/dam/cancer-org/research/cancer-facts-and-statistics/global-cancer-facts-and-figures/global-cancer-facts-and-figures-4th-edition.pdf.

[4] Goren A, Gilloteau I, Lees M, et al. Quantifying the burden of informal caregiving for patients with cancer in Europe. Support Care Cancer. 2014;22(6):1637–1646.

[5] Rutherford C, Mercieca-Bebber R, King M. Health-related quality of life in cancer. In: Ian Olver, editor. The MASCC textbook of cancer supportive care and survivorship. Cham: Springer; 2018. p. 109–125.

[6] Carrera PM, Kantarjian HM, Blinder VS. The financial burden and distress of patients with cancer: Understanding and stepping-up action on the financial toxicity of cancer treatment. CA Cancer J Clin. 2018;68(2):153–165.

[7] Gilligan AM, Alberts DS, Roe DJ, et al. Death or debt? National estimates of financial toxicity in persons with newly-diagnosed cancer. Am J Med. 2018;131(10):1187–1199.e5.

[8] 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.

[9] Woods B, Sideris E, Palmer S, et al. NICE DSU technical support document 19: partitioned survival analysis for decision modelling in health care: a critical review [Internet]. NICE Decis. Support Unit; 2017. Available from: http://scharr.dept.shef.ac.uk/nicedsu/wp-content/uploads/sites/7/2017/06/Partitioned-Survival-Analysis-final-report.pdf.

[10] Williams C, Lewsey JD, Mackay DF, et al. 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.

[11] Degeling K, Franken MD, May AM, et al. Matching the model with the evidence: comparing discrete event simulation and state-transition modeling for time-to-event predictions in a cost-effectiveness analysis of treatment in metastatic colorectal cancer patients. Cancer Epidemiol. 2018;57:60–67.

[12] 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.

[13] Gibson EJ, Begum N, Koblbauer I, et al. Cohort versus patient level simulation for the economic evaluation of single versus combination immuno-oncology therapies in metastatic melanoma. J Med Econ. 2019;22(6):531–544.

[14] Smare C, Lakhdari K, Doan J, et al. 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] NICE. Guide to the methods of technology appraisal 2013; 2013. [cited 2019 Jun 17]. Available from: https://www.nice.org.uk/process/pmg9/chapter/foreword.

[16] R Core Team. 2018. R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.

[17] Latimer N. NICE DSU technical support document 14: undertaking survival analysis for economic evaluations alongside clinical trials – extrapolation with patient-level data. [Internet]; 2011; [cited 2020 Feb 2]. Available from: http://www.nicedsu.org.uk.

[18] Williams C, Lewsey JD, Briggs AH, et al. Cost-effectiveness analysis in R using a multi-state modeling survival analysis framework: a tutorial. Med Decis Making. 2017;37(4):340–352.

[19] NICE. Lenalidomide plus dexamethasone for previously untreated multiple myeloma. TA587. [Internet]; 2019. [cited 2020 Jul 3]. Available from: https://www.nice.org.uk/guidance/TA587.

[20] Brennan A, Chick SE, Davies R. A taxonomy of model structures for economic evaluation of health technologies. Health Econ. 2006;15(12):1295–1310.