



This is a repository copy of *Partitioned survival and state transition models for healthcare decision making in oncology : where are we now?*.

White Rose Research Online URL for this paper:
<https://eprints.whiterose.ac.uk/167383/>

Version: Accepted Version

Article:

Woods, B.S., Sideris, E., Palmer, S. et al. (2 more authors) (2020) Partitioned survival and state transition models for healthcare decision making in oncology : where are we now? Value in Health, 23 (12). pp. 1613-1621. ISSN 1098-3015

<https://doi.org/10.1016/j.jval.2020.08.2094>

Article available under the terms of the CC-BY-NC-ND licence
(<https://creativecommons.org/licenses/by-nc-nd/4.0/>).

Reuse

This article is distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs (CC BY-NC-ND) licence. This licence only allows you to download this work and share it with others as long as you credit the authors, but you can't change the article in any way or use it commercially. More information and the full terms of the licence here: <https://creativecommons.org/licenses/>

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.



eprints@whiterose.ac.uk
<https://eprints.whiterose.ac.uk/>

Abstract

Objectives: Partitioned Survival Models (PSM) are routinely used to inform reimbursement decisions for oncology drugs. We discuss the appropriateness of this method compared to the most commonly used alternative, state transition models (STMs).

Methods: In 2017, we published a NICE Technical Support Document (TSD 19) describing and critically reviewing PSM. In this article we summarise key findings from TSD 19, review new evidence comparing PSM and STM, and review recent NICE appraisals to understand current practice.

Results: PSM evaluates state membership differently from STM and does not include a structural link between intermediate clinical endpoints (e.g. disease progression) and survival. PSM directly considers clinical trial endpoints and can be developed without access to individual patient data, but limits the scope for sensitivity analyses to explore clinical uncertainties in the extrapolation period. STMs facilitate these sensitivity analyses but require development of robust survival models for individual health state transitions. Recent work has shown PSM and STM can produce substantively different survival extrapolations and that extrapolations from STMs are heavily influenced by the specification of the underlying survival models. Recent NICE appraisals have not generally included both model types, reviewed individual clinical event data, or scrutinised life year predictions for individual health states.

Conclusions: The credibility of survival predictions from PSMs and STMs including life years accrued in individual health states, should be assessed using trial data on individual clinical events, external data and expert opinion. STMs should be used alongside PSMs to support assessment of clinical uncertainties in the extrapolation period such as uncertainty in post-progression survival.

Highlights

- i. Partitioned Survival Models (PSM) are routinely used to inform reimbursement decisions for oncology drugs.
- ii. PSM differs from state transition models as it does not include a structural link between intermediate clinical endpoints (e.g. disease progression) and survival. Survival extrapolations have been shown to differ markedly between the two modelling approaches. Although PSMs have some practical advantages, their structure limits the extent to which sensitivity analyses can be used to explore clinical uncertainties in the extrapolation period.
- iii. Decision makers should more carefully scrutinise survival predictions from PSMs and other modelling approaches to assess whether model predictions are credible. This should include an assessment of life years accrued in individual health states, and should take account of trial data on individual clinical events, external data and expert opinion. Due to the challenges in robustly estimating transition probabilities, for the time being we recommend STMs are used alongside PSMs to support the assessment of clinical uncertainties in the extrapolation period, for example by allowing exploration of different survival outcomes following disease progression.

Introduction

Decision modelling has an established role in informing clinical decision making and policy decisions relating to cost-effectiveness. When using decision models to estimate the costs and health effects of interventions, a range of alternative approaches can be used, such as decision trees, Markov models and individual sampling models. Most of these have been subject to detailed discussion within the economic evaluation literature.¹⁻⁴ This has not been the case for one approach, the partitioned survival model (PSM), which has been used extensively in the National Institute for Health and Care Excellence (NICE) Technology Appraisal (TA) Programme and is now the most commonly used approach for NICE appraisals of interventions for advanced or metastatic cancers.⁵

In 2017 we published a NICE Technical Support Document (TSD) describing and critically reviewing PSM as a decision modelling approach to inform policy decisions relating to cost-effectiveness.⁵ This work highlighted that PSM uses an approach to predicting the distribution of individuals across health states (state membership) that is distinct from other commonly used methods such as state transition models (STMs) and the implications of this for: the data required to develop models, model-generated extrapolations, and the degree to which analysts and decision makers can use models to explore uncertainties in the extrapolation period. A set of recommendations was proposed to aid those developing and reviewing decision models and making decisions based on their outputs. In this article we summarise the key findings from NICE TSD 19. We then examine new evidence relating to the appropriateness of PSM and alternative approaches, and review recent NICE appraisals to understand current practice in the use of modelling approaches in oncology.

This work was developed with a focus on the NICE TA programme in the UK, but is relevant for any context where PSM is used to inform assessments of cost-effectiveness. This includes

the broader NICE programme (for example work done to inform NICE Clinical Guidelines has used the approach) and work done by or for reimbursement agencies internationally which has also used PSMs.⁶

In this article we begin by summarising TSD 19. This part of the paper includes a description and critique of PSM, a discussion of the feasibility of alternative methods focusing on the most widely used alternative - STM, and a summary of the TSD recommendations. We then review the current state of play focusing on empirical studies that have compared PSM and STM and then reviewing recent NICE appraisals to assess current practice. We summarise these findings before returning within the discussion to assess the implications of the original work and updated findings for the future of oncology modelling.

Summary of Technical Support Document 19

The Partitioned Survival Modelling method

Decision models aim to describe key biological or clinical processes, and the way in which interventions affect these processes.⁷ Many conditions can be described in terms of a series of distinct clinical states that individuals experience and move between. STMs are often therefore used to model this process. In STMs, movements between health states are referred to as transitions, and the speed at which these transitions occur as transition probabilities or rates. State values are used to reflect the costs and health-related quality of life (HRQoL) implications of residing in, or transiting between, each health state.⁴

PSMs are also characterised by a series of health states with associated state values. However, PSMs do not use transitions between states to determine the proportion of patients in each health state at each time point (state membership) and instead use a different approach and set of information.

The way in which state membership is determined in PSM can be illustrated using a model structure commonly applied in economic evaluations of treatments for advanced or metastatic cancer. This model includes three states: progression-free, progressed and dead where progression implies a worsening or spreading of the cancer. The PSM derives state membership for this model using two survival curves. The Overall Survival (OS) curve describes time from model entry to death and is used to directly determine the proportion of patients alive (and dead) over time. OS is therefore determined independently of other clinical endpoints and independently of whether individuals reside in the progression-free or progressed health states. This assumption of independence and its credibility is rarely discussed, despite being the central assumption PSM. The Progression-Free Survival curve (PFS) describes time from model entry to exiting the progression-free state via progression or death and provides the state membership for the progression-free state over time. For the progressed health state, state membership is derived as the difference between the OS and the PFS curve at each time point as this represents the proportion of patients who are alive but not progression free. This process is shown in Figure 1. PSMs therefore directly uses standard survival analysis of clinical time-to-event endpoints to derive state membership. The approach can be applied to models with any number of health states as long as patients only move progressively through health states (i.e. no “backwards” transitions such as from progressed to progression-free are allowed).

Figure 1

In a PSM the survival curves that inform the estimates of state membership (e.g. PFS and OS) are modelled completely independently. This is the fundamental difference from STMs where clinical events are explicitly related. Figure 2 shows the STM that corresponds to the PSM for

advanced cancer described above. Estimation of state membership in this STM requires three transition probabilities: the probability of disease progression observed prior to death in a model cycle ($p_{pf,p}$), the probability of death from the progression-free state ($p_{pf,d}$) and the probability of death from the progressed state ($p_{p,d}$). Patients start the model in the progression-free state. State membership at the end of each model cycle is estimated by applying the transition probabilities to the state membership at the end of the previous cycle. OS is therefore determined by all three transition probabilities and reflects the evolving proportion of patients in the progressed state and the differences in mortality between progression-free and progressed patients. This is typically referred to as cohort simulation. The structural link between OS predictions and intermediate endpoints such as progression is the fundamental difference from PSM which considers OS to be independent of other clinical events. STMs may use constant transition probabilities (Markov models), time-dependent transition probabilities (semi-Markov models) and may be implemented using the cohort simulation approach described above or patient level simulation approaches described elsewhere.⁸

Figure 2

In the context of a within-trial analysis or a case in which data have been fully observed, PSM and STM approaches are expected to produce similar results if modelling and statistical analyses of survival data have been done appropriately, as relationships between endpoints are reflected within the data. However, when data are incomplete and parameters derived from the observed data are used for extrapolation, as is commonly the case, the approaches are expected to differ. In PSM, OS extrapolation reflects only the OS evidence and not PFS, whereas in a

STM, OS extrapolation is influenced by the model structure and each transition probability estimate.

The way in which the impact of different interventions is modelled also differs. In PSM differences between interventions are modelled independently for each survival endpoint. In STMs treatment effects on individual transition probabilities combine to produce treatment effects on OS.

A review of NICE oncology appraisals conducted to inform TSD 19 found that applications of the PSM share a number of commonalities. For comparators included within the manufacturers' pivotal trial, the Kaplan Meier estimator or a parametric model is typically used to model the survival curves for the within-trial period. A parametric model is typically used to extrapolate within-trial time trends in the rate of events over the remainder of the model time-horizon. External evidence can also quantitatively inform the survival extrapolations.⁹ A separate TSD^{10,11} provides guidance on selection between alternative parametric model specifications emphasising that this should be based on statistical fit to observed data supplemented with external empirical evidence and expert judgement to assess the plausibility of extrapolations.¹² Information on other endpoints, such as the evolving proportion of patients who have progressed, is not explicitly reflected in the extrapolation.¹³ For additional comparators not considered within the pivotal trial, survival curves are often generated by applying the relevant hazard ratio from a single trial or (network) meta-analysis to the modelled pivotal trial data. Our review found there was very limited discussion of PSM in terms of its assumptions or justification for the selection of the PSM approach.⁵

Critique of PSM as an aid to decision making

PSM directly uses survival analysis of time-to-event endpoints to derive state membership estimates. This makes the models easy to construct and communicate. The direct use of PFS and OS survival analyses also means that the approach typically validates well against these endpoints for the within-trial period. Perhaps most importantly, PSM can be implemented using published analysis of the PFS and OS endpoints which are frequently reported in clinical studies. For example, pseudo individual patient data (IPD) for the PFS and OS endpoints can be generated from published graphs.^{14,15} PSM can also incorporate treatment effect estimates for PFS and OS endpoints such as hazard ratios from individual trials, meta-analyses or network meta-analyses. This often provides a way to model the PFS and OS for comparators not included within the pivotal trial. Long-term estimates of mortality rates from historical clinical trials and observational studies have also been incorporated within the approach. That PSM can incorporate commonly reported clinical endpoint data without access to IPD is an important advantage of the approach.

The limitations of the PSM arise when there is uncertainty around long-term OS extrapolations and this leads to uncertainty around cost-effectiveness. Uncertainty around OS extrapolations is pervasive in technology appraisals which typically rely on immature trial data, and limited external data which may not directly relate to the patient group of interest and is typically only available for older comparator treatments. The limitations of PSM for extrapolating OS in these contexts stem from its structural assumption that OS is estimated and extrapolated independently of intermediate endpoints.

The NICE methods guidance recommends that the “*clinical and biological plausibility*” of extrapolations should be assessed and alternative scenarios routinely considered for the extrapolation period.¹⁶ When decision models are underpinned by a structure reflecting biological and clinical processes, it is possible to explicitly consider the mechanisms underpinning extrapolations and to subject these to scrutiny and sensitivity analyses. The

structural independence of the endpoints in a PSM makes this difficult. Rates of individual transitions, or treatment effects on individual transitions can't be varied within a PSM. Conducting these sensitivity analyses would often be insightful, particularly where there is uncertainty about outcomes post-progression and whether these differ across comparators.

For example, an important driver of cost-effectiveness in the majority of cost-effectiveness models of oncology products is the long-term treatment effect on OS. This is often assumed to take the form of a hazard ratio estimated from the within-trial period which is then extrapolated for the duration of the model. The PSM allows alternative specifications of the hazard ratio over time, for example no further treatment effect (hazard ratio=1.0) or waning of the treatment effect over time, or for treatment to result in a certain survivorship at a specified time point (e.g. 10% of patients alive at year 5). It does not, however, allow for important clinically plausible scenarios including no further treatment effect following disease progression, or a reduced treatment effect on mortality following disease progression.

PSMs can generate estimates of mean time spent within each health state and how these differ between treatments. However, it is not possible to review individual transitions which can make the plausibility of results difficult to assess.

For example, if a PSM predicts that treatment extends mean time spent in the progressed state then this indicates that treatment is associated with reductions in the rate of post-progression mortality. However, as this transition is not modelled in a PSM, it is not possible to establish from the model whether this effect was supported by the trial data.

There are instances where a PSM can predict a PFS curve that lies above the OS curve and therefore estimate that a higher proportion of patients are progression-free than alive. This reflects the lack of structural relationship between PFS and OS endpoints. There is no obvious

solution to this problem and this has been addressed in NICE TAs through the use of ad hoc adjustments or restrictions.⁵

Alternative approaches

The main alternative to the PSM in oncology modelling is the STM which uses an explicit disease model to generate survival predictions.¹⁷ In principle, this allows the natural history of the disease and treatments effects on this to be explicitly reflected when extrapolating beyond the trial data, and can improve transparency around the mechanisms and processes underpinning results generated using extrapolation, and facilitate meaningful sensitivity analyses. Until recently, application of STMs to evaluate cancer drugs have often used strong assumptions, inappropriate statistical methods for estimating transition probabilities and have not adequately reflected observed trial survival outcomes.⁵ For example, models have often assumed a constant risk of post-progression death that does not vary between comparators, despite this resulting in OS predictions which are not reflective of observed OS data.

STMs require estimates of the individual transition probabilities, as shown in Figure 2. These can be estimated using survival analytic methods that account for competing and sequential events such as multi-state survival analysis.^{18,19}

Application of STMs in contexts in which PSMs are used raises a number of implementation challenges. The first relates to the fact that survival analyses corresponding to the individual transitions shown in Figure 2 are not typically reported in clinical publications. This is not problematic in cases where modellers have access to the IPD for trials including all relevant comparators. If modellers do not have this access, as is often the case, then estimating the required transition probabilities is more challenging though methods literature is emerging to address this.^{19,20} If summary level data such as Kaplan Meier curves are available for PFS, OS

and PPS then it is possible to estimate the transition probabilities.¹⁹ If only PFS and OS is available transition probabilities can still be estimated²⁰ though the available methods are restricted to estimating constant transition probabilities which may not be appropriate in many clinical contexts.

An alternative approach is to use STM only for those comparators for which IPD is available. This would allow PFS and OS predictions to be generated for these treatments from the STM. Estimates of the relative treatment effects for the other comparators (e.g. hazard ratios for PFS and OS from other trials or a (network) meta-analysis) could then be applied to these “baseline” event rates. This would require a careful consideration of how comparator treatment effects should be applied in the extrapolation period as these treatment effects are disconnected from the disease process within the STM.

Even with the required survival data there are a number of considerations regarding how to appropriately analyse this data to inform the required transition probabilities. This requires multi-state survival analytic methods.^{18,19,21,22} A recent tutorial paper by Williams *et al.*¹⁹ provides a step-by-step guide to using these methods to estimate transition probabilities for STMs. Nonetheless, there remain a number of challenges to obtaining robust transition probability estimates. Firstly, as key outcomes such as OS are determined by a combination of survival models, achieving a satisfactory fit to the observed data can be challenging and should be assessed carefully. Secondly, there is limited research and guidance about how to specify models (e.g. inclusion of patient history covariates); model selection; and generalising evidence relating to PPS to the extrapolation period (Box 1)). The third challenge is in reflecting time-varying transition probabilities. A range of methods are available to address this issue including tunnel states,⁷ semi-Markov approaches²³ and patient level simulation.⁸ These methods are frequently employed in STMs in oncology and other therapeutic areas, though will result in models that are more complex than PSMs.

Box 1: Considerations in analysis and extrapolation of post-progression survival data

Features of post-progression survival (PPS) data

- PPS data only reflects the experience of those patients who progress within the trial period. These patients may have a different prognosis from patients who progress beyond the trial period. Using within-trial PPS data to represent PPS for patients who progress beyond the trial may therefore result in inaccurate extrapolations.
- The duration of follow up for PPS depends on time of progression. Patients who progress earlier are followed up for longer than those who progress later. If timing of progression is related to prognosis then PPS estimates will be subject to bias due to informative censoring.

Potential issues with a naïve analysis of post-progression survival

- If early progressors have a worse prognosis than late progressors, a naïve analysis of within-trial PPS may underestimate PPS in the extrapolation period. This is likely to be particularly acute in the tail of the PPS curve which is informed only by the earliest progressors.

Potential issues with a naïve analyses describing treatment effect on post-progression survival

- If the active treatment is found to extend PFS then a higher proportion of individuals will have progressed in the control arm compared to the active treatment arm. For example, if 10% have progressed on active treatment and 15% on control treatment then the PPS data is potentially informed by the 10% sickest patients in the active arm and the 15% sickest patients in the control arm. This may make PPS shorter in the active arm (i.e. a hazard ratio for active treatment in excess of 1.0), even if the treatment itself doesn't affect PPS. This is not a causal effect of treatment, but reflects a real difference. Care is required when considering if this would continue in the extrapolation period.

Recommendations in TSD 19

Table 1 summarises the recommendations from TSD 19. These recommendations aim to improve the quality of submissions including PSMs and STMs in oncology by improved reporting, using a wider set of information when generating and validating extrapolations and more carefully considering appropriate sensitivity analyses around these extrapolations. Due to the challenges in robustly estimating transition probabilities, for the time being we recommend STMs are used alongside PSMs to support the assessment of their extrapolations.

Table 1: Summary of TSD 19 recommendations

Recommendation focus	Recommendation
Selecting and reporting of modelling approach	<ol style="list-style-type: none"> 1. Report model conceptualisation and rationale for modelling approach selected 2. Use appropriate terminology to describe PSM and do not describe as a STM 3. Report and justify structural assumptions of PSM
Representing uncertainties in extrapolation within PSMs and STMs	<ol style="list-style-type: none"> 4. Recognise limitations of PSM for extrapolation resulting from lack of explicit structural link with intermediate endpoints 5. Report within-trial survival curves for individual clinical events (e.g. progression, pre-progression deaths and post-progression deaths), to allow evidence relating to disease process and treatment effects on this to impact on judgements regarding extrapolation. Holders of IPD should be encouraged to report individual clinical events to support implementation of this recommendation 6. Reflect all relevant evidence in extrapolation including additional information on individual clinical events (see recommendation 5), external empirical data and expert opinion where appropriate 7. Perform appropriate sensitivity analyses to reflect uncertainty in baseline risk and treatment effects over the extrapolation period
Use of STMs	<ol style="list-style-type: none"> 8. STMs should use transition probabilities estimated using appropriate forms of survival analysis

	<p>9. (research recommendation) Further research should explore appropriate approaches to statistical modelling when using multi-state survival analysis to inform STMs</p> <p>10. (research recommendation) Further research is required to support incorporation of summary data (e.g. PFS and OS) in to STMs</p>
<p>Understanding and validating outputs of PSM and STMs</p>	<p>11. Use STM alongside PSM to verify plausibility of PSM extrapolations and explore key clinical uncertainties in the extrapolation period</p> <p>12. Use tabulations to show the health states in which life year and QALY differences between interventions occur and discuss the mechanisms behind these differences and their clinical plausibility</p> <p>13. (research recommendation) Further research to identify biases associated with PSM and STM</p>

PSM=partitioned survival model; STM=state transition model.

Current state of play

Published literature

At the time TSD 19 was developed there was little research evaluating PSM as a modelling approach. We therefore returned to the literature to identify relevant papers. The review methods are summarised in Appendix A.

A recent review has updated and extended the review conducted in TSD 19.¹⁷ This included 100 NICE oncology appraisals published in 2013-2018 and found that PSMs were the most popular structure and used in 54% of TAs, followed by STMs which were used in 41%. In the wider published literature STMs were more dominant comprising 82% of 124 reviewed studies. PSMs comprised only 12% of studies though this may be influenced by the exclusion of country-adaptations of existing models, and by PSMs being incorrectly reported as STMs.¹⁷ The authors confirm findings from TSD 19 that there has been limited justification of modelling approach and exploration of uncertainties relating to model structure. Additionally, a recent tongue-in-cheek discussion of PSM highlighted concerns with its implementation.²⁴ In particular, that applications of PSMs often ignore good practice guidance on model conceptualisation, parametric model selection, indirect comparisons and use of appropriate sensitivity analyses to explore structural uncertainties.

There are now a number of published studies comparing the impact on model predictions and cost-effectiveness of using PSM and STMs.²⁵⁻³⁰ Only one³⁰ was published prior to TSD 19.

Four studies used the same data cut across methods, and allowed all transitions probabilities within the STM to depend on time in state and treatment.^{25,28-30} These are therefore considered the most comparable in structural assumptions to the PSM. Three of the studies^{25,29,30} used three-state model structures as shown in Figures 1 and 2. Pan 2018²⁸ compared a three-state PSM to a more complex discrete event simulation model reflecting different lines of treatment.

In this model time to event estimates included time-dependent predictors (e.g. ECOG at the start of treatment) as well as time spent in previous treatment.²⁸ A scenario analysis examined the impact of excluding these predictors. Smare 2020²⁹ explored two approaches for the STM, one where PPS was independent of patient history, and a second where PPS depended on time taken to progress. The second approach allowed patients with early progression to have a larger PPS benefit from immunotherapy, the authors provided additional clinical data and commentary to support this. Gao 2019²⁵ developed models using only summary data on PFS, OS and PPS whereas the other studies used IPD.

Mean survival estimates from these models are presented in Table 2. There are large differences in absolute and incremental survival predictions between PSM and STM, and between STMs using different specifications for transition probabilities. All studies reported that both model types were able to adequately represent the within-trial data. The differences between approaches therefore occur in the extrapolation period, though differences may be attributable to factors other than model structure such as the selection of survival models. In addition, no study directly addressed the considerations in PPS estimation outlined in Box 1. The papers provided little discussion of the drivers of differences in OS predictions between models, with only Williams 2016 reporting mean survival time by health state.³⁰

Two models compared model extrapolations to longer term OS follow-up from the underlying trials. Pan 2018 compared model predictions based on 27.1 month data to data observed at 49.2 months. The STM provided a better fit to the longer-term data than the PSM for the active trial arm (the comparison for the control arm was not shown). When predictors other than treatment were excluded from the time-to-event equations the STM less accurately predicted long-term OS though still outperformed the PSM. Smare 2020 compared model predictions based on 14-month follow-up to data observed at 38 months. The STM including time to progression as a predictor of PPS outperformed the PSM, which outperformed the simpler STM.

Table 2: Mean OS (life year) estimates from papers comparing partitioned survival models to STMs

Paper	Partitioned survival model			STM		
	Active	Comparator	Incremental	Active	Comparator	Incremental
Gao 2019	1.51	0.86	0.65	1.32	0.91	0.41
Pan 2018	2.99	NR	NR	3.61	NR	NR
Pan 2018 STM no biomarkers	2.99	NR	NR	3.37	NR	NR
Smare 2020	2.77	2.22	0.55	2.58	1.94	0.64
Smare 2020 STM-TTP	2.77	2.22	0.55	3.00	2.39	0.61
Smare 2020 38 mnth¹	4.23	3.48	0.76	3.24	2.65	0.59
Smare 2020 STM-TTP 38 mnth¹	4.23	3.48	0.76	3.43	2.80	0.63
Williams 2016	5.96	5.31	0.65	5.29	4.97	0.32

NR = not reported; STM=state transition model; STM no biomarkers indicates the analysis where all covariates other than treatment were excluded from the time-to-event equations; STM-TTP indicates analysis where post-progression survival was stratified according to time to progression (TTP).

¹ Note that as well as comparing predictions from the model developed using 14-month data to the observed 38-month data, Smare 2020 also presented results based on fitting the model to the 38-month data as shown here.

Other studies that compared the methods used different data for the PSM and STM or implemented a more restrictive type of STM assuming constant transition probabilities.^{26,27}

This makes it difficult to determine whether the differences observed are attributable to the differences in the data, model assumptions or the modelling approach itself.

Finally, one methodological paper was identified focused on estimating transition probabilities when only PFS or OS is available.²⁰ The methods developed assume constant transition probabilities. Further work could usefully seek to relax this assumption, for example, by using external data to inform likely time trends in event rates.

In addition to those studies identified by our review, we are also aware that applied cost-effectiveness studies in oncology are starting to use competing risks and multi-state survival analysis to inform transition rates.³¹⁻³⁵

Recent NICE appraisals

We also reviewed the 10 most recent NICE oncology TAs as of July 2019 to understand current practice with respect to selection, reporting and implementation of modelling approaches and in particular the information used to inform extrapolations of survival data and assess uncertainty around these extrapolations. The review methods are described in Appendix B.

As shown in Table 3, six of the appraisals used a PSM, 3 used a model structure that included a STM component and one used both approaches in response to concerns raised at the first NICE committee meeting relating to the PSM submitted by the manufacturer. Eight of the appraisals referenced TSD 19. STMs were used more frequently in cases where IPD was available for all comparators. TA587 developed a STM, even though IPD was not available for all comparators. In this appraisal STM predictions of PFS and OS were generated for those

comparators for which IPD was available. Hazard ratios describing the effects of the remaining treatments were then applied to these “baseline” event rates.

Table 3: Summary of modelling approaches used in recent NICE oncology appraisals

TA #	Drug appraised	Cancer	Modelling approaches used (includes sensitivity analyses)	IPD available for all comparators
567	Tisagenlecleucel	Diffuse large B-cell lymphoma	Decision tree and PSM	No
573	Daratumumab with bortezomib	Multiple myeloma	PSM	No
577	Brentuximab vedotin	Cutaneous T-cell lymphoma	PSM	Yes
578	Durvalumab	Non-small cell lung cancer	Semi-Markov	Yes
579	Abemaciclib	Breast cancer	PSM	No
580	Enzalutamide	Prostate cancer	Semi-Markov	Yes
581	Nivolumab with ipilimumab	Renal cell carcinoma	PSM	No
584	Atezolizumab	Non-small cell lung cancer	PSM	No
587	Lenalidomide	Multiple myeloma	Hybrid: PSM for 92 weeks then Markov model	No

589	Blinatumomab	Acute lymphoblastic leukaemia	PSM and Semi-Markov model developed	Yes
------------	--------------	----------------------------------	--	-----

PSM = partitioned survival model; STM=state transition model.

Appraisals using PSM

The justifications provided for using PSM included reliance on summary level external data, that the maturity of available data increased the validity of the approach, that there were insufficient clinical data to inform estimation of transition probabilities, and concerns about biases in estimates of PPS when using STM. Three appraisals used the method without justification (TA577, TA579, TA581). Four appraisals acknowledged the independence of clinical endpoints (TA567, TA573, TA584, TA589), though there was limited discussion of the implications of this for extrapolation. Four of the appraisals included the possibility of cure (TA567, TA577, TA581, TA589) by assuming a fixed proportion of individuals achieved cure or that beyond a certain time point all patients were cured.

Generally OS extrapolations were validated against external data and expert clinical opinion. None of the PSMs reviewed presented survival curves for individual clinical events (e.g. progression, pre-progression deaths or post-progression deaths), and mean time spent in individual health states was only discussed in two appraisals (TA567, TA577).

Exploration of uncertainty around survival extrapolations focused on the use of alternative parametric survival models. Some models also explored uncertainty in the duration of treatment effects. The additional structure introduced by models including a cure element allowed for exploration of clinical uncertainties relating to the cure rate, timing of cure, and outcomes following cure.

Appraisals using state transition modelling

The justifications provided for using STM included the need to reflect the link between intermediate prognostic endpoints and OS, the need to reflect the impact of subsequent treatment usage on OS and avoiding logical inconsistencies in models (PFS and OS curves crossing).

All but TA587 reflected time-dependency in event rates and two (TA587, TA589) used multi-state survival analysis to estimate transition probabilities. The others used ad hoc approaches to analysing the IPD. None of the appraisals modelled the potential impact of patient history on subsequent event rates. Concerns were raised in three of the appraisals about the sample size available to estimate the transition probabilities and/or biases in the estimation of PPS (TA578, TA580, TA589).

All but TA589 assessed the internal validity of STM survival predictions, within TA580 concerns were raised about the correspondence between model predictions and longer-term follow-up data. Generally, OS extrapolations were also validated against external data or expert opinion. There was limited discussion of the plausibility of treatment effects on individual transitions and none of the appraisals discussed the mean time spent in individual health states.

The models explored uncertainty at different points in the treatment process including the treatment effect on the intermediate outcome, the impact of the treatment under evaluation on PPS, and the impact of different subsequent therapies on PPS. Appraisals did not routinely examine the sensitivity of the models to both baseline risk of mortality from the progression state and treatment effects on this. Models including a cure element also explored uncertainty in relation to the approach to modelling cure and the point of cure.

Summary of current state of play

PSMs continue to be used extensively to assess the cost-effectiveness of oncology drugs in the NICE TA programme, and broader literature. A growing body of empirical literature now suggests that PSM and STMs can produce quite different survival predictions in the extrapolation period.

Our review of NICE appraisals following the publication of TSD 19 indicates descriptions of PSMs have improved, extrapolations from both model types are generally validated against external empirical data and/or expert opinion, and application of STMs are becoming more sophisticated by allowing for time-dependency and treatment differences in transition probabilities. Across appraisals limited attention was paid to discussing whether incremental life years were accrued in the progression-free or post-progression health states, appraisals using PSM did not use evidence on individual clinical events or STMs to support assessment of extrapolations, and STMs were used to explore uncertainty around PPS to only a limited extent.

Discussion

PSM represents an approach to predicting state membership within cost-effectiveness models that is distinct from commonly used methods such as STMs. In the context of a within-trial analysis or a case in which data have been fully observed, PSM and STM are expected to produce similar results as relationships between endpoints are reflected within the data. When there is a need to extrapolate, the approaches make different assumptions and use different information, and are therefore expected to produce different results. Recent work comparing the predictions generated by PSM and STMs has shown that the approaches produce different predictions of absolute survival and differences in survival between comparators, and that these

differences are attributable to differences in the extrapolation period.^{25,28-30} These differences could have marked implications for decision making informed by cost-effectiveness models.

In TSD 19 we made a series of recommendations to improve use of both PSM and STM. Our review of NICE appraisals following the publication of TSD 19 shows variable adherence to these recommendations. We emphasise here that, regardless of modelling choice, predictions of life years accrued in individual health states should be reviewed for clinical plausibility. This should include reviewing the extent to which incremental life-year gains accrue in individual health states, and in the observed data versus extrapolation periods.³⁶ Survival curves for individual clinical events for each treatment should be used to support an assessment of whether the degree of benefit in the extrapolation period is clinically realistic and where sensitivity analyses should focus. More routine use of STMs to explore key uncertainties in the extrapolation period should be encouraged. For example, if available data on individual clinical events supports an effect of treatment on progression but evidence on post-progression outcomes are highly uncertain due to limited follow-up then it may be important to conduct sensitivity analyses exploring PPS and how this varies by treatment. These sensitivity analyses should be informed by empirical evidence where available but will often require a level of judgement when estimating the impact of a recently licensed treatment on PPS. This should reflect whether, and to what extent, effects of treatment will continue post-progression and how the new treatment modifies post-progression treatment pathways.

Applications of STMs should carefully consider survival model specification. Recent work suggests that inclusion of a covariate describing time taken to progress (or other predictors of prognosis at the time of progression) may improve the accuracy of extrapolations of PPS and therefore OS.^{28-30,32} Use of this approach assumes that the relationship between time to progression and PPS observed within the trial period applies to patients who progress during

the extrapolation period. Further research to support specification of survival regressions in this setting remains of high priority.

Role of the funding source

The study sponsor did not have any role in the study design; in the collection, analysis, and interpretation of data; in the writing of the report; or in the decision to submit the paper for publication.

Figure titles and legends

Figure 1: Determining state membership in a three-state partitioned survival model

OS(t) denotes the OS curve at time t, and PFS(t) the Progression-Free Survival curve at time t.

Figure 2: Specification for a state transition model

$p_{pf,p}$ denotes the probability of disease progression observed prior to death in a model cycle;

$p_{pf,d}$ the probability of death from the progression-free state, and $p_{p,d}$ the probability of death from the progressed state.

References

1. Barton P, Stirling B, Robinson S. Modelling in the economic evaluation of health care: selecting the appropriate approach. *J Health Serv Res Policy*. 2004;9(2):110–118.
2. Brennan A, Chickb S, Davies R. A taxonomy of model structures for economic evaluation of health technologies. *Health Econ*. 2006;15:1295–1310.
3. Sonnenberg F, Beck R. Markov Models in Medical Decision Making: A Practical Guide. *Med Decis Making*. 1993;13(4):322-338.
4. Siebert U, Alagoz O, Bayoumi A, et al. State-transition modeling: a report of the ISPOR-SMDM modeling good research practices task force-3. *Value Health*. 2012;15(6):812-820.
5. Woods B, Sideris E, Palmer S, Latimer N, Soares M. NICE DSU Technical Support Document 19: Partitioned Survival Analysis for Decision Modelling in Health Care: A Critical Review *NICE Decision Support Unit*. 2017.
6. Masucci L, Beca J, Sabharwal M, Hoch JS. Methodological issues in economic evaluations submitted to the pan-Canadian Oncology Drug Review (pCODR). *Pharmacoeconomics Open*. 2017;1(4):255-263.
7. Briggs A, Sculpher M, Claxton K. Decision Modelling for Health Economic Evaluation (1st Edition). Oxford, UK. *Oxford University Press*. 2006:1-76.
8. Davis S, Stevenson M, Tappenden P, Wailoo A. NICE DSU Technical Support Document 15: Cost-Effectiveness Modelling Using Patient-Level Simulation. *NICE Decision Support Unit*. 2014.
9. Guyot P, Ades A, Beasley M, Lueza B, Pignon J-P, Welton N. Extrapolation of Survival Curves from Cancer Trials Using External Information. *Med Decis Making*. 2016;37(4):353-366.
10. 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):743-754.
11. Latimer N. NICE DSU Technical Support Document 14: Survival Analysis For Economic Evaluations Alongside Clinical Trials - Extrapolation With Patient-Level Data *NICE Decision Support Unit*. 2011.
12. Bell Gorrod H, Kearns B, Stevens J, et al. A Review of Survival Analysis Methods Used in NICE Technology Appraisals of Cancer Treatments: Consistency, Limitations, and Areas for Improvement. *Med Decis Making*. 2019;39(8):899-909.
13. Bagust A, Beale S. Survival Analysis and Extrapolation Modeling of Time-to-Event Clinical Trial Data for Economic Evaluation: An Alternative Approach. *Med Decis Making*. 2013;34(3):343-351.
14. Guyot P, Ades A, Ouwens M, Welton N. Enhanced secondary analysis of survival data: reconstructing the data from published Kaplan-Meier survival curves. *BMC Med Res Methodol*. 2012;12(9).
15. Hoyle M, Henley W. Improved curve fits to summary survival data: application to economic evaluation of health technologies. *BMC Med Res Methodol*. 2011;11(139).
16. National Institute of Health and Care Excellence, Guide to the methods of technology appraisal. 2013.
17. 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.
18. Putter H, Fiocco M, Geskus R. Tutorial in biostatistics: Competing risks and multi-state models. *Stat Med*. 2007;26(11):2389–2430.
19. Williams C, Lewsey J, Briggs A, Mackay D. Cost-effectiveness analysis in R using a multi-state modelling survival analysis framework: A tutorial. *Med Decis Making*. 2016;37(4):340-352.
20. Pahuta MA, Werier J, Wai EK, Patchell RA, Coyle D. A technique for approximating transition rates from published survival analyses. *Cost Effectiveness and Resource Allocation*. 2019;17(1).

21. Putter H. Special Issue about Competing Risks and Multi-State Models. *Journal of Statistical Software*. 2011;38(1).
22. Putter H, van der Hage J, de Bock G, Elgalt R, van de Velde C. Estimation and Prediction in a Multi-State Model for Breast Cancer. *Biometrical Journal* 2006;48(3):366-380.
23. Hawkins N, Sculpher M, Epstein D. Cost-Effectiveness Analysis of Treatments for Chronic Disease: Using *R* to Incorporate Time Dependency of Treatment Response. *Med Decis Making*. 2005;25(5):511–519.
24. Beca J, Husereau D, Chan KKW, Hawkins N, Hoch JS. Oncology Modeling for Fun and Profit! Key Steps for Busy Analysts in Health Technology Assessment. *Pharmacoeconomics*. 2018;36(1):7-15.
25. Gao L, Li S-C. Modelled Economic Evaluation of Nivolumab for the Treatment of Second-Line Advanced or Metastatic Squamous Non-Small-Cell Lung Cancer in Australia Using Both Partitioned Survival and Markov Models. *Appl Health Econ Health Policy*. 2019;17(3):371-380.
26. Gibson E, 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.
27. Lee D, Amadi A, Sabater J, et al. Can we accurately predict cost effectiveness without access to overall survival data? The case study of nivolumab in combination with ipilimumab for the treatment of patients with advanced melanoma in England. *Pharmacoeconomics Open*. 2019;3(1):43-54.
28. Pan F, Reifsnider O, Zheng Y, et al. Modeling clinical outcomes in prostate cancer: application and validation of the discrete event simulation approach. *Value Health*. 2018;21(4):416-422.
29. Smare C, Lakhdari K, Doan J, Posnett J, Johal S. Evaluating Partitioned Survival and Markov Decision-Analytic Modeling Approaches for Use in Cost-Effectiveness Analysis: Estimating and Comparing Survival Outcomes. *Pharmacoeconomics*. 2019;38:97-108.
30. Williams C, Lewsey J, Mackay D, Briggs A. 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*. 2016;37(4):427-439.
31. Bensimon AG, Zhou Z-Y, Jenkins M, et al. Cost-effectiveness of pembrolizumab for the adjuvant treatment of resected high-risk stage III melanoma in the United States. *J Med Econ*. 2019;22(10):981-993.
32. Woods BS, Sideris E, Sydes MR, et al. Addition of docetaxel to first-line long-term hormone therapy in prostate cancer (STAMPEDE): modelling to estimate long-term survival, quality-adjusted survival, and cost-effectiveness. *European urology oncology*. 2018;1(6):449-458.
33. Asaria M, Walker S, Palmer S, et al. Using electronic health records to predict costs and outcomes in stable coronary artery disease. *Heart*. 2016;102:1-8.
34. Bongers M, de Ruyscher D, Oberije C, Lambin P, Uyl-de Groot C, Coupe´ V. Multistate Statistical Modeling: A Tool to Build a Lung Cancer Microsimulation Model That Includes Parameter Uncertainty and Patient Heterogeneity. *Med Decis Making*. 2016;36(1):86-100.
35. Castelli C, Combescure C, Foucher Y, Daures J. Cost-effectiveness analysis in colorectal cancer using a semi-Markov model. *Stat Med*. 2007;26(1):5557–5571.
36. Canadian Agency for Drugs and Technologies in Health (CADTH). Guidelines for the Economic Evaluation of Health Technologies: Canada. In:2017.