



This is a repository copy of *A review of recent decision-analytic models used to evaluate the economic value of cancer treatments*.

White Rose Research Online URL for this paper:
<http://eprints.whiterose.ac.uk/169370/>

Version: Published Version

Article:

Bullement, A. orcid.org/0000-0001-7091-0972, Cranmer, H.L. and Shields, G.E. (2019) A review of recent decision-analytic models used to evaluate the economic value of cancer treatments. *Applied Health Economics and Health Policy*, 17 (6). pp. 771-780. ISSN 1175-5652

<https://doi.org/10.1007/s40258-019-00513-3>

Reuse

This article is distributed under the terms of the Creative Commons Attribution-NonCommercial (CC BY-NC) licence. This licence allows you to remix, tweak, and build upon this work non-commercially, and any new works must also acknowledge the authors and be non-commercial. You don't have to license any derivative works on the same terms. 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/>



A Review of Recent Decision-Analytic Models Used to Evaluate the Economic Value of Cancer Treatments

Ash Bullement¹ · Holly L. Cranmer² · Gemma E. Shields³

Published online: 4 September 2019
© The Author(s) 2019

Abstract

Cost-effectiveness analysis provides information on the potential value of new cancer treatments, which is particularly pertinent for decision makers as demand for treatment grows while healthcare budgets remain fixed. A range of decision-analytic modelling approaches can be used to estimate cost effectiveness. This study summarises the key modelling approaches considered in oncology, alongside their advantages and limitations. A review was conducted to identify single technology appraisals (STAs) submitted to the National Institute for Health and Care Excellence (NICE) and published papers reporting full economic evaluations of cancer treatments published within the last 5 years. The review was supplemented with the existing methods literature discussing cancer modelling. In total, 100 NICE STAs and 124 published studies were included. Partitioned-survival analysis ($n = 54$) and discrete-time state transition structures ($n = 41$) were the main structures submitted to NICE. Conversely, the published studies reported greater use of discrete-time state transition models ($n = 102$). Limited justification of model structure was provided by authors, despite an awareness in the existing literature that the model structure should be considered thoroughly and can greatly influence cost-effectiveness results. Justification for the choice of model structure was limited and studies would be improved with a thorough rationale for this choice. The strengths and weaknesses of each approach should be considered by future researchers. Alternative methods (such as multi-state modelling) are likely to be utilised more frequently in the future, and so justification of these more advanced methods is paramount to their acceptability to inform healthcare decision making.

1 Introduction

1.1 Cost-Effectiveness Analysis in Oncology

Cancer imposes a significant economic burden on society, which will continue to increase with the expanding and aging global population. In 2018, it was estimated that

there were 17.0 million new cancer cases and 9.5 million cancer deaths worldwide [1]. By 2040, this is expected to grow to 27.5 million new cancer cases and 16.3 million cancer deaths globally [1]. Innovative cancer treatments have helped to improve survival rates and patients' quality of life; however, these new treatments are often associated with high acquisition costs. As healthcare expenditure continues to rise globally, the need to assess whether the additional costs of new cancer treatments are justified by the improved health outcomes is becoming more prevalent.

Decision modelling plays an important role in the estimation of the value of new cancer treatments. With an ever-increasing demand for new, effective treatments within the constraints of a fixed healthcare budget, cost-effectiveness analysis (CEA) provides decision makers with an objective basis from which decisions may be informed. Through conducting a CEA, the difference in costs and effectiveness (usually presented in the form of quality-adjusted life-years [QALYs]) between two or more interventions is quantified. The decision maker can then consider if additional costs are justified by the improved health outcomes—typically by

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s40258-019-00513-3>) contains supplementary material, which is available to authorized users.

✉ Ash Bullement
abullement@deltahat.co.uk

¹ Delta Hat Limited, 212 Tamworth Road, Nottingham NG10 3GS, UK

² Takeda UK Limited, Building 3, Glory Park, Woodburn Green, Buckinghamshire HP10 0DF, UK

³ Centre for Health Economics, Division of Population Health, Health Services Research and Primary Care, School of Health Sciences, Faculty of Biology, Medicine and Health, University of Manchester, Oxford Road, Manchester M13 9PL, UK

Key Points for Decision Makers

Most models developed for cancer treatments utilise a state transition or partitioned-survival structure.

The choice of model structure is rarely discussed within published studies, though it is a requirement of submissions made to the National Institute for Health and Care Excellence (NICE).

Newer modelling methods (such as multi-state modelling) are expected to be used more frequently in future studies, which aim to address some commonly cited structural limitations.

assessing the cost per QALY gained (or the ‘incremental cost-effectiveness ratio’ [ICER]).

In the UK, national reimbursement decisions regarding cancer treatments are made by the National Institute for Health and Care Excellence (NICE). Companies are invited to submit evidence regarding the safety, efficacy and cost effectiveness of new interventions as part of the technology appraisal (TA) process. Outside of the UK, other health technology assessment (HTA) agencies also make funding decisions regarding the uptake of new cancer drugs, either with or without the formal submission of cost-effectiveness evidence by companies. Where formal evidence submission of cost-effectiveness evidence is not mandated by HTA agencies or payers, published sources of economic evidence (such as peer-reviewed publications) may influence decision making.

When using decision models to generate estimates of the costs and effects of interventions, a range of alternative modelling approaches can be used. These include decision trees, Markov models (discrete or continuous), partitioned-survival analyses (PartSAs; also known as ‘area under the curve’ [AUC] models) and individual sampling models. The choice of model structure can affect the results produced, though multiple models to assess the same decision problem are seldom constructed to explore this [2, 3].

Published guidelines detail the assumptions and interactions underpinning possible modelling approaches and provide direction as to the most appropriate model structure given the disease area and available data. Despite this, the choice of model structure is rarely justified or discussed in published studies [3–5]. NICE Decision Support Unit (DSU) Technical Support Document (TSD) 19 reported the findings of a review of NICE TAs of cancer treatments covering the period May 2013–February 2016 [6]. This review found that the PartSA approach was the most commonly used structure (22/30), and that few appraisals justified the use of this structure beyond case precedence (6/22). The review also

found some misunderstanding between PartSA, Markov and semi-Markov approaches, with ten of the PartSA structures incorrectly labelled as Markov or semi-Markov models.

Several papers have recently been published comparing cost-effectiveness results across different model structures in the field of oncology [7–13]. These indicate that different models may provide different results, thus highlighting the importance of justifying the selected approach and assessing the impact of structural uncertainty (in other words, uncertainty in the functional form of the model constructed).

This study aims to identify and critically evaluate current modelling techniques in oncology, with the aim of highlighting methodological issues and gaps in the literature. Following the identification of these key modelling structures, a summary of the review findings is presented to aid analysts considering different model structures within the field of oncology, and recommendations for future research are made.

2 Methods

A literature review was conducted to identify economic evaluations of cancer treatments, with a focus on the NICE single technology appraisal (STA) process, supplemented with the published literature.

2.1 Search Strategy

Searches were performed in November 2018, capturing relevant publications and STAs published since 2013 up to the date of searching. Database searches were conducted using MEDLINE and EMBASE via Ovid. Search terms included terms specific to economic evaluation and oncology, and results were restricted by article type and language (English). Strategies were pilot tested to ensure all studies already known to the authors were retrieved. Search strategies are provided in Electronic Supplementary Material S.1. STAs were identified through the NICE website by viewing TA guidance by topic (cancer) for the same time period (2013–2018).

2.2 Selection of Studies

Study inclusion criteria are outlined in Table 1. Studies not meeting these criteria were excluded during the two-stage screening process.

Independent screening was undertaken at two stages (firstly of abstracts and titles, then of full papers) for the published literature by a reviewer (GES), with a second and third reviewer each checking a quarter of decisions (AB and HLC). The primary reason for exclusion was recorded at both stages.

Table 1 Inclusion criteria for review

Criterion	Requirement for inclusion
Population	People with cancer (no restriction on the type of cancer)
Intervention	Pharmacological interventions aimed at treating cancer (increasing health and length of life). Interventional studies looking at complications of cancer (e.g. treating anaemia or infections), surgical interventions and precision medicine-focused studies were excluded
Comparator	Comparison with any active intervention, usual care, best supportive care or palliative care
Methods	Studies were required to report the development and use of a decision-analytic model. Multiple technology and highly specialised technology appraisals were excluded
Outcomes	Full economic evaluations (cost-effectiveness or cost-utility studies)
Other	Journal articles published in English language from 2013 up until November 2018. Full-text articles (excluding protocols, case reports, conference proceedings or discussion pieces) were included from the published literature. STAs were included if the necessary documents were available via the NICE website. Publications were excluded if they described the findings of a NICE technology appraisal, or were highlighted within the publication as a country adaptation of a pre-existing published model or NICE STA

NICE National Institute for Health and Care Excellence, STA single technology appraisal

2.3 Data Extraction and Quality Appraisal

A data extraction form was designed prior to screening. This included intervention, cancer type, author-reported model structure, modelled health states (where applicable), and author-reported strengths and limitations specific to the model structure. Specific to the publications, funding source and country setting were extracted. For the NICE STAs, feedback on the model structure from the Evidence Review Groups (ERGs) was extracted. Data extraction was completed by one reviewer (GES: published papers; AB: STAs) with a quarter of the data extraction checked by a second reviewer (HLC). Review findings are presented via narrative synthesis.

2.4 Supplementary Summary of Modelling Structures

To provide a more detailed background of the findings of the review, short summaries of the alternative modelling approaches identified were developed based on methods literature already known to the authors. The strengths and weaknesses of each approach were identified from a targeted review of published materials, as well as practical considerations that the authors deemed of relevance for inclusion.

3 Results

3.1 Identified Studies (Single Technology Appraisals and Published Papers)

An overview of the identification of relevant studies is presented as a Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) diagram in Fig. 1.

After excluding duplicates and ineligible publications, a total of 100 NICE STAs and 124 published studies were identified as relevant.

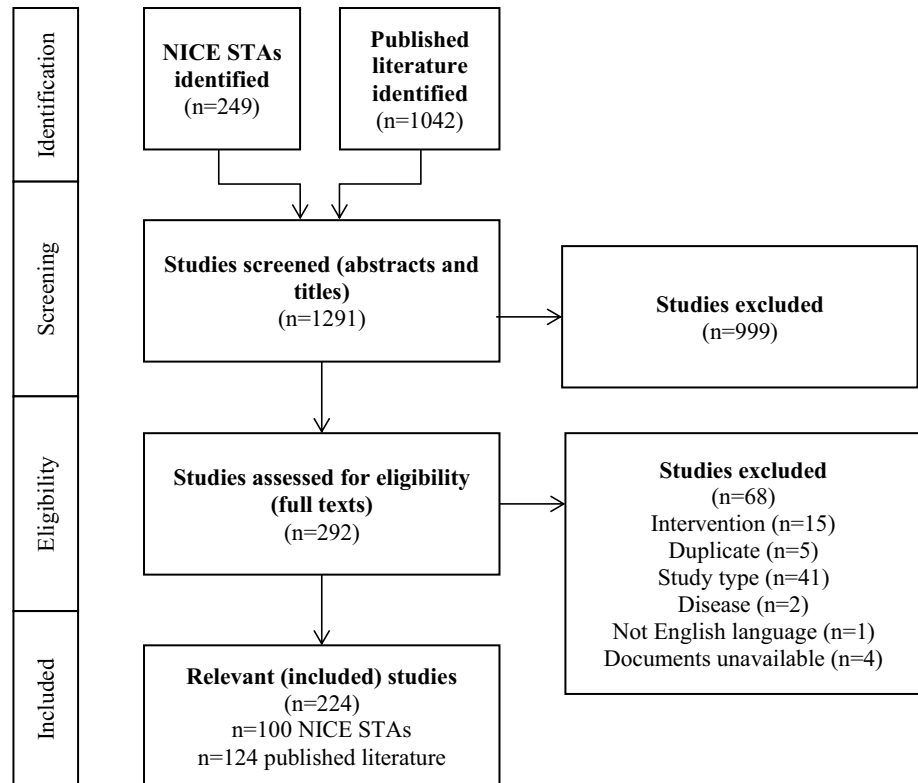
3.2 Overview of Identified Studies

Table 2 provides a summary of the identified studies. Across both the literature and the NICE STAs, there appears to be an increasing number of oncology cost-effectiveness models published each year. Note that one identified STA and four papers from the published literature reported on multiple economic models within one publication.

Figure 2 presents the model structures employed across the NICE STAs and the literature. PartSAs were demonstrated to be the most popular structure across the NICE STAs ($n = 54$), followed by discrete-time state transition structures ($n = 41$) and decision tree or other cohort-based combination structures ($n = 6$). State transition models most commonly applied the Markov assumption ($n = 19$). However, 16 STAs relaxed this assumption via a semi-Markov approach. Five STAs did not provide any further detail on their state transition approach (i.e. Markov or semi-Markov). The remaining models considered simulation approaches, which also fall under the state transition bracket and were described as patient-level simulation ($n = 1$) and time series individual simulation ($n = 1$). Of the four combination structures, two considered a decision tree populating a state transition model and two considered a decision tree populating a PartSA.

In comparison, discrete-time state transition models were demonstrated to be the most popular structure in the published literature ($n = 102$), followed by a PartSA ($n = 15$) and either a decision tree or other cohort-based combination structures ($n = 10$). Five of the published papers did not describe their model structure. State transition models

Fig. 1 Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) diagram for inclusion of relevant studies. *NICE* National Institute for Health and Care Excellence, *STA* single technology appraisal



were most often reported as applying a Markovian (memoryless) assumption ($n = 88$), with only four relaxing this assumption and considering a semi-Markov approach. Models reporting funding from pharmaceutical companies were more frequently PartSA structures (13/40; 33%) than were those with other funding sources. Patient-level simulations were described as discrete event simulation ($n = 4$), patient-level simulation ($n = 2$), micro-simulation ($n = 1$), time series individual simulation ($n = 1$), simulation ($n = 1$) and a macro-simulation Markov ($n = 1$) models. Combination structures relate to those which consider more than one type of structure within the model; in terms of the identified published papers, five papers considered a decision tree feeding into a state transition Markov model.

When reviewing each of the NICE STAs and published papers, it was evident that there was some confusion between state transition and PartSA structures. This was also previously reported in the review detailed in NICE DSU TSD 19 where authors found that ten of the PartSA models had been erroneously labelled as Markov models [6]. Whilst there is insufficient information provided in most studies to conclusively identify model misspecifications, there were some clear examples where the model was labelled as a Markov or semi-Markov state transition model but further description indicated these were in fact PartSAs. For example, health state occupancy calculations were informed by partitioning the overall survival (OS) curve using intermediate outcomes

such as progression-free survival (PFS) or event-free survival (EFS)—occupancy in the progressed disease state was then calculated by the difference between the OS and intermediate event curves, indicating a PartSA structure. Other examples included reported parameters focused on the coefficients relevant to parametric curves fit to the OS and intermediate outcomes data rather than health state-specific transitions, as well as the application of hazard ratios to baseline OS and intermediate outcomes curves with no description on how the competing events reflected by the intermediate outcome (e.g. PFS) were separated, as required in a state transition model. Consequently, we believe that many of the reported state transition models were actually implemented as a PartSA. The inability to conclusively identify model structures from the literature highlights both the confusion between state transition and PartSA structures as well as the lack of transparency in reporting.

3.3 Description of Most Common Modelling Approaches

Our review identified two modelling approaches as the most commonly used in oncology in recent years: the PartSA and the discrete-time state transition approaches. Detailed descriptions of each of these approaches can be found in the literature, as well as the relative advantages and limitations

Table 2 Summary of identified studies

Characteristic	NICE STAs [n (%)]	Published literature [n (%)]
Year published		
2013	3 (3)	19 (15)
2014	9 (9)	17 (14)
2015	7 (7)	22 (18)
2016	25 (25)	26 (21)
2017	28 (28)	29 (23)
2018 (up to 31 October)	28 (28)	11 (9)
Country		
North America	–	57 (46)
Europe (excluding UK)	–	27 (22)
Asia	–	25 (20)
UK	100 (100)	9 (7)
South America	–	4 (3)
Australasia	–	2 (2)
Cancer type		
Blood and bone marrow cancers	30 (30)	28 (23)
Breast cancer	8 (8)	29 (23)
Lung cancer	15 (15)	15 (12)
Skin cancer	12 (12)	4 (3)
Colorectal cancer	2 (2)	10 (8)
Ovarian cancer	3 (3)	8 (6)
Prostate cancer	6 (6)	5 (4)
Others ^a	24 (24)	25 (20)

NICE National Institute for Health and Care Excellence, STA single technology appraisal

^aOther cancers represents those that had less than ten total publications in total grouped together (published literature and STAs combined)

of each method [6, 14, 15]. A summary of these methods is given in Sects. 3.3.1 and 3.3.2.

3.3.1 Partitioned-Survival Analysis

PartSA models are characterised by health states with state membership determined by a series of (usually) *independently* modelled, non-mutually exclusive survival curves; N represents the number of health states in the model and to which $N-1$ survival curves are required. The PartSA structure is most commonly applied as a three-state model, which makes use of two modelled survival curves for OS and PFS. The area underneath the PFS curve represents the proportion of patients who are yet to progress, and the area between the OS and PFS curves represents the proportion of patients who have progressed and are still alive. The remaining proportion of patients (i.e. patients who are neither progressed nor yet to progress) represents those who have died.

In this three-state model, the assumption of independent survival curves necessitates that OS and PFS are

independent of each other, i.e. progression is not considered to be a prognostic factor of death within the fitted OS model. The dependency between OS and PFS is captured in the within-trial period. However, this is not reflected in the extrapolation beyond this point if independence is assumed. Therefore, the more immature data, the less dependence captured between OS and PFS in the long-term extrapolations.

3.3.2 State Transition Model

State transition models are defined by distinct events or health states that individuals experience and transition between; the speed at which transitions occur are called transition probabilities or rates. These health states should be mutually exclusive and exhaustive. Unlike the PartSA approach, state transition models capture the dependency between events; the extent to which this dependence is captured depends on the specific state transition model applied. In the example of a three-state transition model (pre-progression, progressed disease and death), there are three transitions which are estimated: (1) pre-progression to progressed disease; (2) pre-progression to death; and (3) progressed disease to death. Therefore, the dependence between progressed disease and death is captured in the extrapolations through the evolving proportion of patients in the progressed disease health state.

Most commonly in oncology, discrete time periods have been considered employing either a Markov or semi-Markov approach. The discrete time periods are implemented as model cycles—fixed time periods over which model transitions are calculated (e.g. daily, weekly, monthly). The semi-Markov approach relaxes the ‘memoryless’ feature of the Markovian assumption by allowing transitions to depend on time spent within an intermediate health state; for example, the transition from progressed disease to death may depend on how long a patient has spent in the progressed disease health state.

3.4 Justification and Criticism of Modelling Approaches

Within the NICE STA process, companies are required to provide a detailed explanation for the selected modelling approach. The role of the independent ERG is to critically appraise the companies’ rationale as part of the appraisal process, and ultimately determine if the model is appropriate to inform decision-making or not.

The most common strength of the submitted model structure noted by the company was the similarity of the chosen structure to previous appraisals conducted by NICE and/or other published literature ($n = 75$). The second most common strength noted was the ability to use data directly from the pivotal trial(s) to inform health state occupancy within

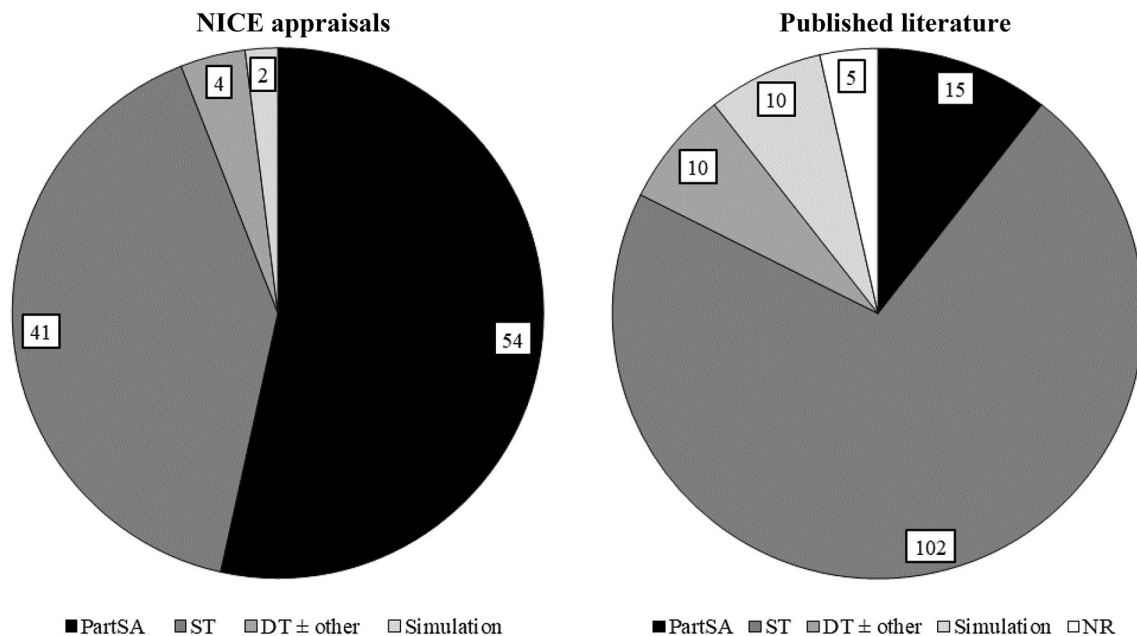


Fig. 2 Comparison of model structures adopted in published studies versus the NICE STA process. The total number of model structures is greater than the total number of identified papers or appraisals as some publications reported multiple economic models. *DT ± other* decision tree ± other combination model structure, *NICE* National

Institute for Health and Care Excellence, *NR* not reported or documents unavailable, *PartSA* partitioned-survival analysis, *ST* state transition model (Markovian or semi-Markovian), *STA* single technology appraisal

the submitted model ($n = 20$)—this was primarily raised as a strength of the PartSA structure as the endpoint of many trials of cancer treatments is either OS or PFS. Other strengths raised included the ability to reflect time dependency ($n = 6$), applicability of the structure to a treatment pathway with multiple treatment lines ($n = 6$), being aligned with the irreversible nature of transitions ($n = 5$), providing a straightforward and/or transparent approach ($n = 5$), avoiding specific issues with the extrapolation of OS ($n = 4$), allowing for incorporation of external data ($n = 4$), avoiding the need for tunnel states to be included ($n = 2$), exhibiting intuitive interpretation ($n = 2$), and permitting calibration of model transitions ($n = 2$).

The ERGs rarely commented on the merits of the companies' submitted model structures. We found no explicit instances where the ERG commented on the appropriateness of the company's model structure beyond a general comment that the model was either aligned with convention, appropriate to inform decision-making, or was considered transparent and simple.

In only seven of the 104 STAs, the company noted at least one limitation of their chosen modelling approach (i.e. choice of PartSA, state transition or another model structure). The most common limitation highlighted by the submitting company was that the PartSA structure was considered 'rigid', as individual transitions are not modelled. Despite reporting limitations, none of the NICE STAs

explored the impact of structural uncertainty through an alternative model structure. One appraisal submitted with two different models (TA284; one semi-Markov state transition model and one PartSA). However, data from different populations informed these models and they were viewed in isolation [16].

Criticisms raised by ERGs in relation to the model structure used by the company included the lack of calibration between outcome measures (e.g. through independently fitted survival curves) ($n = 5$), an 'over-simplification' of the final health state (such as 'progressed disease') ($n = 4$), seemingly 'counterintuitive' results (e.g. an improvement in PFS leading to a higher ICER) ($n = 4$) and structural assumptions that were considered 'inappropriate' (such as a constant hazard of a given event occurring) ($n = 3$). It was noted that relatively few ERG reports specifically flagged issues relating to the model structure, but we suspect some of the criticisms may have either been raised during the appraisal committee meetings or within other sections of the appraisal documentation (e.g. in relation to the incorporation of clinical efficacy data).

Within the published literature, there was limited description provided around the choice, justification and design of model structure. A minority of published studies (22%; 27/124) reported any strengths or limitations associated with the chosen model structure, and, where reported, the information provided was very brief. Two papers from the

published literature reported four different Markov state transition models in each publication and compared the results across these, emphasising the need to assess structural uncertainty [17, 18]. A further two papers from the published literature reported on multiple model structures; however, these reflected different stages in the disease pathway and did not provide an assessment of structural uncertainty.

4 Discussion

This review provides an update of the literature search published in NICE DSU TSD 19, in terms of NICE STAs considering cancer treatments published since 2013, and supplements this with published cost-effectiveness analyses identified from the literature. Our motivation for conducting this review was twofold: (1) to identify what modelling techniques are most commonly applied in the field of oncology; and (2) to highlight the methodological issues and gaps evident across the available studies.

Our research has identified a lack of transparency in model specification, justification in terms of choice of model structure and an exploration of the impact of structural uncertainty on the cost-effectiveness model results. As these findings impede our ability to address our original research question regarding which modelling techniques are used and their methodological issues, our discussion focuses on how modellers could improve on the historical modelling and reporting practice.

4.1 Transparency in Model Specification

All identified papers reported the type of model structure that had supposedly been implemented within the analysis. However, it was not always possible to corroborate the stated model structure with the structure truly implemented, and it was evident that some published models were incorrectly labelled. To address transparency in model specification, some journals have advocated the publication of the economic model file and tabulated model parameters as supplementary material. In providing these materials, readers may objectively identify the model structure employed, yet an improved reporting standard of model structures is still needed.

4.2 Justification for Model Structure

We found that the published literature contained limited rationale defending the assumptions underpinning the choice of model structure, which is aligned with the findings of previous studies [3–5]. The NICE STAs provided more detailed discussion regarding model structure; though a requirement

of the submission template specifies justification of model choice. Nevertheless, companies most often justified their choice of structure with case precedence and ability to incorporate data from the trial or literature directly. It is our opinion that these two reasons alone do not reflect good modelling practice, and this is particularly concerning as in many cases the ERG's report also noted precedence as a key reason why the model structure was appropriate. The NICE DSU TSD19 provides detailed recommendations concerning the justification of PartSA models specifically, which, based on the findings of our review, is currently under-used [6].

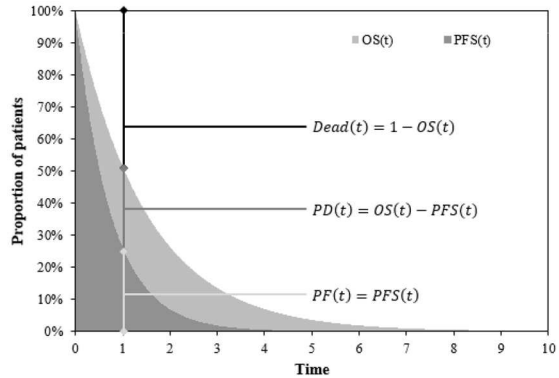
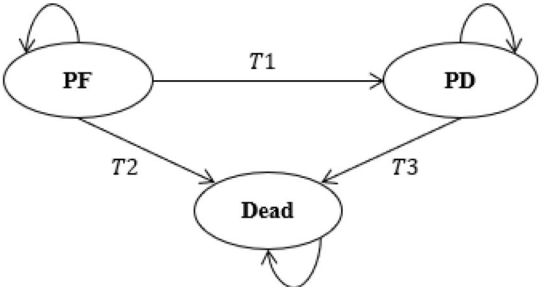
4.3 Exploration of Structural Uncertainty

None of the identified studies utilised multiple model structures (based on the same data) to assess structural uncertainty. This finding is despite the recommendation from NICE DSU TSD 19, which states 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” [6]. There are a number of reasons why multiple models may not be developed in practice, including budget and time constraints, knowledge and expertise of software packages, and other practical constraints. However, data granularity is a pertinent point in cancer modelling as most published clinical trials report only a selection of outcomes (e.g. just OS and PFS), making it difficult to use these data outside of a PartSA approach. Ideally, survival curves for individual clinical events (e.g. time to progression) should be reported in clinical trial publications in addition to ‘standard’ endpoints, in order to inform a broader range of model structures.

Data maturity is also an important consideration, particularly in HTAs where decisions are often made based on early clinical trial data cuts. Immature data may encourage modellers to use a simplistic PartSA approach in favour of a more ‘data hungry’ structure. More recent advances in real-world data collection may now provide modellers with the ability to supplement maturing trial data with other data sources in order to inform a more complex model structure.

Presenting multiple model structures may introduce further uncertainty in relation to which structure offers the most robust estimate of cost effectiveness. The ‘best’ estimate of cost effectiveness should be derived from the model structure that most closely aligns with the key features of the disease and technology [6]. Furthermore, if there are differences in the cost-effectiveness estimates from the model structures, these results should be accompanied by both an explanation of the mechanism driving the differences and a justification for these differences which is supported by the data and the disease pathway. Increased transparency associated with model structures, as described earlier, will allow decision makers to more robustly select the ‘best’ estimate

Table 3 Model structures reference table

Structure	Summary	Diagram	Strengths	Weaknesses
PartSA	Fit survival curves to relevant outcomes (e.g. OS and PFS) to predict proportion of patients residing in each health state over time		<ul style="list-style-type: none"> Intuitive application of OS and PFS curves to inform health state occupancy Can use KM curves without access to individual patient data Can reflect complex hazard functions through the use of advanced extrapolation methods 	<ul style="list-style-type: none"> Commonly assumes independence of outcomes (e.g. OS and PFS) Average time an individual is expected to stay in each health state cannot be calculated for all states Can only be applied to processes in which patients move forward through a series of progressive health states
DT ST	Calculate probabilities of moving from one health state to another over time (may be constant or time-varying)	 $PF(t) = PF(t - 1) \times (1 - T1 - T2)$ $PD(t) = PF(t - 1) \times T1 + PD(t - 1) \times (1 - T3)$ $Dead(t) = PF(t - 1) \times T2 + PD(t - 1) \times T3$	<ul style="list-style-type: none"> Can capture dependency between outcomes (e.g. progression and death) Treatment effect can be specified for individual transitions Does not require the estimation of one over-arching OS curve Possible to more easily incorporate time-dependency into calculations Can consider multidirectional transitions 	<ul style="list-style-type: none"> Can be difficult to interpret as commonly used trial endpoints (such as OS and PFS) may not directly inform the model Requires calibration of component transitions to be carefully considered so as to avoid misleading probabilistic analysis results May not be possible to implement for some decision problems due to lack of available data from published sources

DT ST discrete-time state transition, KM Kaplan-Meier, OS overall survival, PartSA partitioned-survival analysis, PD progressed disease, PF progression-free, PFS progression-free survival, *t* time, *T1* transition probability 1 [PF to PD], *T2* transition probability 2 [PF to dead], *T3* transition probability 3 [PD to dead]

of cost effectiveness. However, it is important to note that decisions by HTA agencies are seldom based on a sole cost-effectiveness estimate, and therefore providing multiple model structures offers another mechanism to assess decision uncertainty.

4.4 Recommendations and Conclusions

Guidelines from NICE DSU TSD 19 document provide a useful, under-utilised reference as to what information should be reported and discussed in relation to each modelling approach [6]. However, we recommended the development of more consistent and appropriate terminology in terms of model structure and methods to reduce the current misreporting of model structures. To aid with this, we have provided a comparison of the state transition and PartSA model structures in Table 3. Furthermore, we advocate increased use of the reporting guidance provided in NICE DSU TSD 19 for both HTA submissions and published literature regarding the description of how health state occupancy is calculated, the assumptions underpinning this and how relative efficacy estimates are incorporated (including how these may change over time, e.g. to reflect treatment waning effects). As a minimum, we recommend that journals request inclusion of a model parameters table as supplementary material (as is standard in submissions made to NICE).

Further research is required to understand which model structure should be used and when. This is important to deepen our understanding of the bias introduced by the assumptions underpinning different model structures. Key contexts to consider are different stages of disease, tumour types, and interventions with different mechanisms of action and likely prognoses (e.g. gene therapies and immunotherapies). This research may involve building multiple model structures and comparing the results—presentation of results should include tabulations showing the states in which life-year and QALY differences between interventions accrue and a justification of why these differences should be considered plausible. Research may also involve re-visiting established models when later data cuts become available to assess how well the earlier model predicted long-term outcomes, or performing a simulation study wherein potential causes of bias may be enabled and disabled as needed in order to inform optimal model structure selection. Additional research is also required to further advance modelling methods; for example, the need to develop methodology to allow the extraction of transition probabilities from published summary data only (as without this, the use of the PartSA approach is expected to be chosen by default) and appropriate statistical modelling to address competing risks and/or sequential events (e.g. multi-state modelling).

This review has some limitations that should be highlighted. First, the literature search was restricted to the

English language and studies published within the last 5 years. These restrictions were selected based on a combination of pragmatism, relevance to current modelling practice and generalisability of the findings across studies. Nevertheless, the restricted approach may have led to the exclusion of some relevant studies. Secondly, we extracted the modelling approach used based on author specification. As documented in previous published reviews, we found some of the structures to be incorrectly reported, such as those reported as Markov or semi-Markov when they were a PartSA structure [6]. Searches were kept simple to produce an overview of the published literature, rather than a comprehensive systematic review. In addition, the grey literature was excluded. A published review identified a link between published studies and more favourable results [19]; thus, if certain model types are more likely to produce less favourable results, these may be published less often. Finally, whilst we did exclude duplicate publications, country adaptations were included. Therefore, the results of this review may over-represent the evidence base capturing the same model published multiple times.

Author Contributions AB drafted the initial manuscript, based on an outline developed by AB, HLC and GES. GES led the review of previous published studies, and the retrieval of relevant studies (including contacting NICE). AB led the review of previous NICE appraisals. HLC led the consistency checking of results from both reviews, and performed the quantitative analysis presented within the paper. All authors reviewed each version of the manuscript.

Compliance with Ethical Standards

Funding The authors received no specific funding for this work.

Conflict of interest Ash Bullement, Holly L. Cranmer and Gemma E. Shields declare that they have no conflict of interest.

Research data policy Data sharing is not applicable to this article as no datasets were generated or analysed during the current study.

Open Access This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

1. American Cancer Society. Global cancer facts & figures. 4th ed. Atlanta: American Cancer Society; 2018.
2. Drummond MF, Barbieri M, Wong JB. Analytic choices in economic models of treatments for rheumatoid arthritis: what makes a difference? *Med Decis Mak.* 2005;25:520–33. <https://doi.org/10.1177/0272989X05280561>.

3. Brennan A, Chick SE, Davies R. A taxonomy of model structures for economic evaluation of health technologies. *Health Econ.* 2006;15:1295–310. <https://doi.org/10.1002/hec.1148>.
4. Roberts M, Russell LB, Paltiel AD, Chambers M, McEwan P, Krahn M, ISPOR-SMDM Modeling Good Research Practices Task Force. Conceptualizing a model: a report of the ISPOR-SMDM modeling good research practices task force-2. *Med Decis Mak.* 2012;32:678–89. <https://doi.org/10.1177/0272989X12454941>.
5. Husereau D, Drummond M, Petrou S, Carswell C, Moher D, Greenberg D, et al. ISPOR Health Economic Evaluation Publication Guidelines-CHEERS Good Reporting Practices Task Force. Consolidated Health Economic Evaluation Reporting Standards (CHEERS)—explanation and elaboration: a report of the ISPOR Health Economic Evaluation Publication Guidelines Good Reporting Practices Task Force. *Value Health.* 2013;16:231–50. <https://doi.org/10.1016/j.jval.2013.02.002>.
6. 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. 2017. Available from <http://nicedsu.org.uk/wp-content/uploads/2017/06/Partitioned-Survival-Analysis-final-report.pdf>.
7. Williams C, Lewsey JD, Mackay DF, Briggs AH. Estimation of survival probabilities for use in cost-effectiveness analyses: a comparison of a multi-state modeling survival analysis approach with partitioned survival and markov decision-analytic modeling. *Med Decis Making.* 2017;37:427–39. <https://doi.org/10.1177/0272989X16670617>.
8. Minacori R, Bonastre J, Lueza B, Marguet S, Levy P. How to model survival in cost-effectiveness analysis? Differences between Markov and partitioned survival analysis models. *Value Health.* 2015;18:A704. <https://doi.org/10.1016/j.jval.2015.09.2639>.
9. Gibson E, Koblbauer I, Begum N, Dranitsaris G, Liew D, McEwan P, et al. Modelling the survival outcomes of immuno-oncology drugs in economic evaluations: a systematic approach to data analysis and extrapolation. *Pharmacoeconomics.* 2017;35:1257–70. <https://doi.org/10.1007/s40273-017-0558-5>.
10. Degeling K, Franken MD, May AM, van Oijen MGH, Koopman M, Punt CJA, 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;57:60–7. <https://doi.org/10.1016/j.canep.2018.09.008>.
11. Gibson EJ, Begum N, Koblbauer I, Dranitsaris G, Liew D, McEwan P, 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;22:531–44. <https://doi.org/10.1080/13696998.2019.1569446>.
12. 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 partition survival and Markov models. *Appl Health Econ Health Policy.* 2019;17:371–80. <https://doi.org/10.1007/s40258-018-0452-0>.
13. Lee D, Amadi A, Sabater J, Ellis J, Johnson H, Kotapati S, et al. Can we accurately predict cost effectiveness without access to overall survival data? The case study of nivolumab in combination with ipilimumab for the treatment of patients with advanced melanoma in England. *Pharmacoecon Open.* 2019;3:43–54. <https://doi.org/10.1007/s41669-018-0080-5>.
14. Siebert U, Alagoz O, Bayoumi AM, Jahn B, Owens DK, Cohen DJ, et al. State-transition modeling. *Med Decis Making.* 2012;32:690–700. <https://doi.org/10.1177/0272989X12455463>.
15. Sonnenberg FA, Beck JR. Markov models in medical decision making. *Med Decis Mak.* 1993;13:322–38. <https://doi.org/10.1177/0272989X9301300409>.
16. National Institute for Health and Care Excellence. Bevacizumab in combination with paclitaxel and carboplatin for first-line treatment of advanced ovarian cancer. Technology appraisal guidance [TA284]. London: NICE; 2013. Available at: <https://www.nice.org.uk/guidance/ta284/history>. Accessed 28 Aug 2019.
17. Le QA. Structural uncertainty of Markov models for advanced breast cancer: A simulation study of lapatinib. *Med Decis Making.* 2016;36:629–40. <https://doi.org/10.1177/0272989X15622643>.
18. Le QA, Bae YH, Kang JH. Cost-effectiveness analysis of trastuzumab emtansine (T-DM1) in human epidermal growth factor receptor 2 (HER2): positive advanced breast cancer. *Breast Cancer Res Treat.* 2016;159:565–73. <https://doi.org/10.1007/s10549-016-3958-x>.
19. Bell CM, Urbach DR, Ray JG, Bayoumi A, Rosen AB, Greenberg D, et al. Bias in published cost effectiveness studies: systematic review. *BMJ.* 2006;332:699–703. <https://doi.org/10.1136/bmj.38737.607558.80>.