ORIGINAL RESEARCH ARTICLE

Development and Validation of the TRansparent Uncertainty ASsessmenT (TRUST) Tool for Assessing Uncertainties in Health Economic Decision Models

Sabine E. Grimm[1] · Xavier Pouwels[1] · Bram L. T. Ramaekers[1] · Ben Wijnen[1] · Saskia Knies[2] · Janneke Grutters[3] · Manuela A. Joore[1]

Abstract

Background An increasing number of technologies are obtaining marketing authorisation based on sparse evidence, which causes growing uncertainty and risk within health technology reimbursement decision making. To ensure that uncertainty is considered and addressed within health technology assessment (HTA) recommendations, uncertainties need to be identified, included in health economic models, and reported.

Objective Our objective was to develop the TRansparent Uncertainty ASsessmenT (TRUST) tool for systematically identifying, assessing, and reporting uncertainties in decision models, with the aim of making uncertainties and their impact on cost effectiveness more explicit and transparent.

Methods TRUST was developed by drawing on the uncertainty and risk assessment literature. To develop and validate this tool, we conducted HTA stakeholder discussion meetings and interviews and applied it in six real-world HTA case studies in the Netherlands and the UK.

Results The TRUST tool enables the identification and categorisation of uncertainty according to its source (transparency issues, methodology issues, and issues with evidence: imprecision, bias and indirectness, and unavailability) in each model aspect. The source of uncertainty determines the appropriate analysis. The impact of uncertainties on cost effectiveness is also assessed. Stakeholders found using the tool to be feasible and of value for transparent uncertainty assessment. TRUST can be used during model development and/or model review.

Conclusion The TRUST tool enables systematic identification, assessment, and reporting of uncertainties in health economic models and may contribute to more informed and transparent decision making in the face of uncertainty.

✉ Sabine E. Grimm
Sabine.grimm@mumc.nl

[1] Maastricht University Medical Centre, Department of Clinical Epidemiology and Medical Technology Assessment, School for Public Health and Primary Care (CAPHRI), P. Debyelaan 25, PO Box 5800, 6202 AZ Maastricht, The Netherlands

[2] Zorginstituut Nederland, Willem Dudokhof 1, 1112 ZA Diemen, The Netherlands

[3] Department for Health Evidence, Radboud University Medical Centre, Post 133, PO Box 9101, 6500 HB Nijmegen, The Netherlands

1 Introduction

Uncertainty is rife in health technology assessment (HTA), which leads to increased complexity in decision making. Drug approval processes are typically based on evidence that exhibits several limitations [1–3]. These may include, among others, a lack of data on a certain relevant health outcome (e.g. long-term treatment effectiveness), limited generalisability of evidence to the decision context, small sample sizes, and data immaturity. The result is that health economic models used to inform such decisions (further referred to as 'models') suffer from uncertainty in, for instance, their inputs, extrapolations, and/or assumptions [4, 5]. This uncertainty about the (cost) effectiveness of new health technologies means HTA decision making is cognitively challenging

△ Adis

Key Points for Decision Makers

In health economic decision making, uncertainty information is currently not reported systematically, which leaves decision makers with the cognitively challenging task of translating pieces of qualitative and quantitative uncertainty information into an overall assessment of uncertainty and risk.

When considering uncertainty in health economic models, undue emphasis is placed on imprecision issues, which may lead to misrepresentation of uncertainty and ultimately of risk.

The TRansparent Uncertainty ASsessmenT (TRUST) tool enables the systematic identification, assessment, and reporting of uncertainties in health economic models, which may contribute to more informed and transparent decision making in the face of uncertainty.

[2]. This situation is worsening: Recent pressures to grant market access earlier in the drug development process mean decisions are having to be made based on increasingly immature evidence [6–9] or on single-arm studies [10, 11]. If these uncertainties are not fully appreciated, decisions may be based on a biased view of expected incremental net benefit and the associated uncertainty. This in turn would lead to an increased risk of making a suboptimal decision, which can be described as opportunity losses, or potential losses in population health that could be averted with gathering additional evidence [12].

Unfortunately, uncertainties are often not fully characterised in models [5]. We think this happens for two reasons: (1) undue emphasis on imprecision issues and (2) difficulty precisely expressing all uncertainty. The undue emphasis means that uncertainty about the quality and the generalisability of evidence to the setting in question is not always considered [13]. This is an omission that has long been recognised in the assessment of quality of evidence [14] and the risk assessment literature [15, 16]. The GRADE (Grading of Recommendations Assessment, Development and Evaluation) working group, for instance, highlighted the lack of focus on the quality of clinical evidence and proposed a system to acknowledge and even rate it [14, 17]. To the best of our knowledge, no tool for systematically identifying uncertainty about the quality of evidence used in models is currently available. Regarding (2), uncertainty is most completely characterised with a full explicit probability distribution [4, 5, 18, 19]. In theory, it may be possible to represent all uncertainty in this way, for example using methods to express expert opinion [20, 21] and/or structural uncertainty [22–24]. In practice, tight HTA timelines and resource constraints may

be prohibitive. Other barriers include that people may not be able to express their beliefs in the form of probabilities [25] or that the distribution of the additional structural uncertainty parameter may be difficult to specify [13]. Therefore, most models do not incorporate all uncertainty in a quantitative manner and thus do not fully express the risk associated with a decision as a numerical value [13, 26].

In the worst case, the lack of systematic identification and assessment of all uncertainties present in models may result in a state of uncertainty ignorance, leaving decision makers unaware of the risk associated with a decision. But even where uncertainty information is present, it is usually scattered throughout reports, and decision makers are left with the cognitively challenging task of identifying all uncertainties and translating pieces of qualitative and quantitative uncertainty information into an overall assessment of uncertainty and risk [15]. A unified, comprehensive framework for identifying uncertainties in models, assessing their impact, and reporting the results is therefore needed.

We developed a tool for systematically identifying, assessing and reporting uncertainty in health economic submissions. The paper is structured as follows: we first describe the development and validation of the TRUST (TRansparent Uncertainty ASsessmenT) tool. In Sect. 3, we present findings from interviews and case studies and the final tool, as well as its place in the reimbursement decision-making process. We conclude with a discussion and conclusion.

2 Development of TRUST

2.1 The Theoretical Framework

We developed a first version of the tool based on the literature on uncertainty and risk frameworks in HTA and risk assessment disciplines [14, 16, 18, 27–29]. Relevant literature was identified in a review of the literature on uncertainty frameworks. We used a pragmatic approach to this review and started with a citation search of a paper known to us [16]. We then checked the references of a recently published interdisciplinary literature review of uncertainty frameworks [18] for relevant uncertainty literature. Concepts surrounding uncertainty from the identified literature were checked for applicability and relevance in the HTA decision-making context by a group of researchers that drew upon experience in HTA decision making in the Netherlands and the UK. MJ, JG, BR, SG, and XP have been active members of a UK National Institute for Health and Care Excellence (NICE) evidence review group (ERG) and have built models that informed reimbursement decision making. SK has been a senior policy advisor of the Dutch Health Care Institute (ZIN) since 2012 and MJ a member of the scientific advisory committee of ZIN since 2013.

2.1.1 Model Aspects

To systematically identify and assess uncertainty in models, it was necessary to establish the locations where uncertainties could occur. Walker et al. [27] developed a "conceptual basis for the systematic treatment of uncertainty in model-based decision support activities such as policy analysis, integrated assessment and risk assessment." The authors identified the following locations: context (including natural, technological, economic, social, and political representation), model (structure and technical), and inputs (driving forces, system data), parameters, and model outcomes. Standard HTA dossier templates from the UK and Netherlands were used to inform more detailed descriptions of model aspects.

2.1.2 Sources of Uncertainty

Different uncertainty frameworks provide different taxonomies of uncertainty and use different terminology to describe uncertainty. Walker et al. [27] considered that uncertainty must be viewed as a three-dimensional concept: its nature (whether it is epistemic or aleatory), its level (from deterministic knowledge to ignorance), and its locations (what is affected in the model-based assessment). Epistemic uncertainty is that which could be reduced with further knowledge; aleatory uncertainty is caused by inherent variability [18]. In our paper, much like in the related health policy decision-making literature [28], we focus on epistemic uncertainty, i.e. that which can be addressed in some form. The 'levels' in the Walker et al. [27] framework range from deterministic knowledge via statistical uncertainty (precision), scenario uncertainty (where the mechanism leading to different possible outcomes is not well understood), and recognised ignorance (scientific evidence for developing scenarios is weak) to total ignorance (unknown unknowns). These different levels were defined with the intention to call for different actions to address them. This list of levels is a pragmatic summary of the seven 'sources' of epistemic uncertainty previously identified [16]. The Walker et al. [27] framework has been adapted for the NUSAP (Numeral, Unit, Spread, Assessment and Pedigree) approach to uncertainty assessment [29], which added methodological unreliability as a further type of uncertainty.

The GRADE working group outlined a broader framework for assessing quality of evidence [17], establishing five limitations that can lead to a downgrading of the quality: study limitations resulting in bias, inconsistent results (across studies), indirectness of evidence, imprecision, and publication bias [14, 17]. Van der Bles et al. [18] distinguished between 'levels' of direct and indirect uncertainty, where direct uncertainty pertains to imprecision in estimates that can be communicated in quantitative terms, and indirect uncertainty relates to the quality of the underlying knowledge, which will typically be communicated as a list of caveats (GRADE was named as an example tool that facilitates the assessment of indirect uncertainty). The International Society for Pharmacoeconomics and Outcomes Research–Society for Medical Decision Making (ISPOR–SMDM) taskforce report [28] makes a distinction between methodological uncertainty, parameter uncertainty, stochastic uncertainty, heterogeneity, and structural uncertainty.

Our final selection of possible sources of uncertainty was informed by these taxonomies and based on the consideration that, as different approaches are used in practice to manage uncertainty, it should be possible to match sources of uncertainty relevant to decision makers with an approach to dealing with uncertainty [27].

2.2 Expert Feedback

The TRUST tool was based on the theoretical framework just described and further developed based on several rounds of feedback from HTA experts. After the theoretical framework was established, initial feedback was obtained by means of a discussion meeting with advisors of ZIN. The revised TRUST tool was presented at academic conferences (EUHEA 2018, Maastricht and IHEA 2019, Basel) and at further discussion meetings at ZIN, Maastricht University Medical Center, and Radboud University Medical Center. TRUST was presented during an HTA committee meeting for tomosynthesis for breast cancer population screening in the Netherlands. We conducted individual in-depth semi-structured interviews with different international stakeholders ($N = 11$) to finalise the content of the tool, validate it, and obtain insights into barriers to its use, user-friendliness, and feasibility. These stakeholders were chosen to represent the following stakeholder groups: HTA policy advisors ($n = 2$), independent academic advisors for HTAs ($n = 4$), an appraisal committee member ($n = 1$), and industry stakeholders ($n = 4$, one working in a pharmaceutical company and three for HTA consultancies); convenience sampling was used to establish contact. The interviews were recorded, transcribed, and analysed by themes. The interview guide can be found in Appendix 1 in the Electronic Supplementary Material (ESM). Interviews were used to inform the selection of model aspects and sources of uncertainty to include in TRUST, terminology, and to highlight potential feasibility and usability issues.

2.3 Case Studies

The tool was tested for its validity and whether it 'worked' in practice by means of application in five retrospective HTA case studies on expensive medicines and one ongoing HTA, the latter being on tomosynthesis for breast cancer population

screening in the Netherlands. Four of the retrospective case studies were for ZIN in the Netherlands, and these were performed by two researchers independently, who then compared notes to identify challenges in the use of TRUST and potential inter-rater variability. The fifth retrospective case study was a NICE appraisal (pembrolizumab for treating relapsed or refractory classical Hodgkin lymphoma); we present this here for illustration purposes. This application of TRUST was based on the information available in the article, ERG report, and company submission [11]. The case studies were selected based on convenience sampling: the ZIN referred us to the four retrospective Dutch case studies, and both the Dutch prospective study and the NICE study were projects in which members of our research team were involved.

3 Results

3.1 Findings from Stakeholder Interviews and Case Studies

Insights from stakeholder interviews were an important part of the development of this tool. The most important findings included the following (see Box 1 for more detail):

1. Keep it concise and simple
2. Rating uncertainty with a score may create a false sense of validity and be time consuming

3. Definitions and descriptions need to be clear
4. Time was the main barrier to the use of TRUST

The selected case studies represented a range of potential uncertainty issues encountered in health economic models (Table 1). They included appraisals of one screening technology (tomosynthesis for breast cancer population screening in the Netherlands) and five drugs (pembrolizumab for treating relapsed or refractory classical Hodgkin lymphoma in the England and Wales national health service, and lumacaftor/ivacaftor vs. best supportive care for treatment of cystic fibrosis in patients aged ≥ 12 years homozygous for the F508del mutation in the *CFTR* gene, nivolumab for treatment of non-small-cell lung cancer, rifaximin vs. lactulose alone for the treatment of recurrent episodes of manifest hepatic encephalopathy in patients aged ≥ 18 years, and eculizumab for the treatment of atypical haemolytic uremic syndrome; the latter four were from the Netherlands).

3.2 The TRUST Tool

The TRUST tool integrates the outlined theoretical frameworks on model aspects and sources of uncertainty in one taxonomy (Fig. 1). The rows present model aspects, and columns relate to sources of uncertainty (Appendix 2 in the ESM presents the full TRUST tool and updated versions can be found here: https://osf.io/8eay7/?view_only=90222b3e31 834e469c24a92d587e247d).

Box 1 Interview and case study findings

1. Keep it concise and simple

  Although there were conflicting viewpoints on this (some respondents wished to add more response options or additional detail), the general feedback was that usability and interpretation of findings would improve with a version that was as concise as possible. The TRUST tool was simplified, for example by having dichotomous questions in the identification of sources of uncertainty in each model aspect, rather than providing three- or four-level response options. A summary was developed that only shows the main issues at one glance (Fig. 1 and Appendix 2 in the ESM).

2. Rating uncertainty with a score may create a false sense of validity and be time consuming

  The first version of TRUST included a scoring system that was inspired by GRADE. This was intended to help assess, based on the different sources of uncertainty, the severity of the overall uncertainty associated with one model aspect. However, this caused multiple issues: there was significant inter-rater variability due to the subjective nature of assigning a score to the severity of uncertainty. Respondents considered that they would spend a long time trying to assign the score. This was confirmed in the case study applications. Rating uncertainty was also deemed potentially misleading: users might mistake the resulting score for a quantitative assessment or weight. As a result, a scoring system of uncertainty was not included in TRUST.

3. Definitions and descriptions need to be clear

  Respondents voiced difficulty with understanding the exact meaning of certain combinations of model aspects and sources of uncertainty. In response, we developed a key to the TRUST tool, with definitions and examples for each cell (Appendix 2 in the ESM).

4. Time is the main barrier to the use of TRUST

  The main barrier to its use was the perceived time it would take to complete TRUST. Opinions differed on this: some interviewees thought it was easy to complete while developing or reviewing a model; others stated it may present an additional burden for analysts and reviewers. Our case studies indicated that, after some of the issues with definitions and levels were resolved, TRUST could be completed within approximately 1 h when the dossier and model were already known. Application in retrospective case studies showed that completing TRUST could be difficult and time consuming when relying on information presented in pharmacoeconomic dossiers where uncertainty information was not easily accessible.

ESM electronic supplementary material, *GRADE* Grading of Recommendations Assessment, Development and Evaluation; *TRUST* TRansparent Uncertainty ASsessmenT tool

Table 1 Matrix of features observed in 6 case studies

Features	Tomosynthesis	Pembrolizumab	Lumacaftor / ivacaftor	Nivolumab	Rifaximine	Eculizumab
Features of case study						
Drug		X	X	X	X	X
Other technology	X					
ICER above threshold	X	X	X	X		
Single arm studies		X				
Indirect treatment comparison		X				
Long-term extrapolation beyond available data		X	X	X	X	X
Clinical evidence from setting different to decision problem	X	X	X	X	X	X
Issues in clinical evidence due to intermediate outcomes			X			
Issues in TRUST model aspects observed in case study						
Issues with context / scope	X	X				X
Issues with structural uncertainty		X			X	
Issues with selection of evidence / review		X	X		X	X
Issues with effectiveness			X	X	X	
Issues with relative effectiveness	X		X	X	X	
Issues with adverse events		X			X	X
Issues with HRQoL	X	X	X		X	X
Issues with cost & resource use	X	X	X	X		X
Issues with model implementation	X				X	X

Fig. 1 The TRUST tool part 1: identifying uncertainty

3.2.1 Model Aspects

The selected model aspects included the 'context/scope', the 'model structure', 'selection of evidence' to inform the model, 'model inputs', its 'implementation', and 'outcomes'. Our choice of model aspects was informed by Walker et al. [27]. Any deviations from this relate to the terminology used (to adapt to the target audience). With the 'scope/context' of the appraisal, we considered the definition of PICOTP (Patients, Intervention, Comparator, Outcomes, Time horizon, Perspective) of the decision problem. The aspect 'model structure' is about the selection of health states and/or events and how they relate to each other. 'Selection of evidence' refers to the review(s) informing the inputs of the assessment. We further subdivided the 'model inputs' section, distinguishing between transition probabilities/time-to-event estimates/accuracy estimates; relative effectiveness estimates; adverse events; utilities; and resource use and costs. By 'model implementation' we refer to its technical implementation and the calculations made. 'Model outcomes' are the reported results of the model.

3.2.2 Sources of Uncertainty

Summarising our findings, we consider that actionable sources of uncertainty include 'transparency issues', 'methodological issues', 'imprecision', 'bias and indirectness', and 'unavailability'.

Transparency issues were included based on our experience and our interviews. We found that transparency issues were fairly common in models, and their descriptions and the addition of this item was deemed valuable in interviews. Ideally, such transparency issues can be addressed early on, with clarification requests. Based on the ISPOR–SMDM taskforce report and the NUSAP framework [28, 29], we considered uncertainty in methods as an additional source of uncertainty to be included in TRUST: by this, we refer to cases in which modelling methods are chosen that deviate from best practices or the reference case. This uncertainty could be addressed by means of clarification requests and scenario analysis.

The items of imprecision, bias and indirectness, and unavailability are issues with the evidence itself and shed light on its quality, credibility, and relevance. The item imprecision was non-controversial and was covered in all taxonomies, albeit with different terminology: parameter uncertainty in the ISPOR–SMDM taskforce report, direct uncertainty in van der Bles et al. [18], statistical uncertainty (precision) in Walker et al. [27] and NUSAP [29], and imprecision in GRADE [14]. In TRUST, bias and indirectness covers multiple aspects related to issues with the evidence. The GRADE items of 'study limitations', 'indirectness', and 'publication bias' are all included here [14],

as is indirect uncertainty per van der Bles [18] and structural uncertainty as per the ISPOR–SMDM taskforce report [28]. This source of evidence also includes scenario uncertainty per Walker et al. [27], as potential biases may be explored through scenarios when parameterisation is not an option. This is best addressed using scenario analysis or elicitation of expert opinion.

We disregarded GRADE's inconsistent results, as we considered those to be issues with imprecision or bias and indirectness, which could respectively be parameterised or explored in scenarios [13]. Finally, we added unavailability to account for recognised ignorance per Walker et al. [27] and considered that such issues could be addressed with (extreme) scenario analysis, for example, informed by additional expert opinion.

3.2.3 Combination of Model Aspects and Sources of Uncertainty

In TRUST, combinations between aspects and sources of uncertainty are depicted in cells (Fig. 1, and Appendix 2 in the ESM). In these cells, we ask the following questions for each of the model aspects:

1. Is there a lack of clarity in presentation, description, or justification (transparency issue)?
2. Is there a violation of best research practices, existing guidelines, or the reference case (methodology issue)?
3. Is there an issue with particularly wide confidence intervals, small sample sizes, or immaturity of data (imprecision)?
4. Is there risk of bias or indirectness (bias and indirectness)?
5. Is there a lack of data or insight (unavailability)?

These are dichotomous questions, with 'yes' or 'no' answers provided as response options. 'Yes' turns the cell red, and 'no' turns it green to enable a quick graphical overview of the main uncertainties in the assessment. A 'not applicable' option is also available. Furthermore, for questions 2–5, 'intransparent' can be selected when a transparency issue was reported when answering the first question. A 'remarks' column allows the analyst or reviewer to provide detail on the identified uncertainties (Appendix 2 in the ESM).

Some of the combinations are more intuitive than others, and some may be downright impossible. Our experience from case studies and feedback from interviewees resulted in the following conclusions: imprecision issues are unlikely to arise within the model aspects of context/scope, model structure, and selection of evidence. Unavailability was deemed non-intuitive for the model aspect model structure. For the model aspect implementation, the relevant sources

of uncertainty were transparency issues and methods issues only, and, for model outcomes, we only considered transparency issues, as the others would be an aggregation of uncertainties identified in the previous model aspects. The impossible (or at least unlikely) combinations are greyed out in TRUST. For all other possible combinations, we provide explanations and examples in Appendix 2 in the ESM.

3.3 Potential Impact of Uncertainties on Cost Effectiveness

The next step in TRUST is the assessment of the impact on cost effectiveness (Fig. 2). In this section of TRUST, the analyst can indicate whether the described uncertainty is appropriately reflected in the probabilistic sensitivity analysis (PSA) (and hence expected value of perfect information). The next question asks whether the uncertainty is explored in scenario analysis.

Based on PSA and/or scenario analysis, we ask the final question: whether the identified uncertainty has an impact on cost effectiveness. Where an uncertain aspect is parameterised and reflected in the PSA, these results can be used to identify the most impactful uncertainties. Where neither a PSA nor scenario analyses are available, a subjective judgement must be made. The response options to this question include 'likely high impact', 'likely low impact', 'likely no impact', and 'unknown'. The responses 'likely high impact'

and 'unknown' turn these cells red. An application of the TRUST tool is shown in Box 2 and Table 2.

3.4 Proposed Use of TRUST in the Decision-Making Process

Based on the discussions with stakeholders, we think of the process of dealing with uncertainty in HTA as comprising the following three phases, with specific courses of action and roles pertaining to each (Fig. 3): phase I, identify; phase II, explore impact; and phase III, manage.

In phase I, we suggest first identifying issues with transparency and methods (phase Ia). These can best be addressed by writing or requesting clarifications, redoing any analyses that were not methodologically sound, and/or performing scenario analysis on the methodological uncertainties. In phase Ib, uncertainties caused by imprecision, bias and indirectness, and/or unavailability are identified. These can, respectively, be addressed by PSA/deterministic sensitivity analysis and plausible and/or extreme scenario analysis. In addition, it is possible that issues with bias and indirectness, unavailability, and methodological issues are parameterised and therefore reflected in the PSA.

Phase II explores the impact of uncertainty on, first, cost effectiveness (IIa) and then risk (IIb). An overview of high-impact uncertainties can aid the design of potential research schemes, which can be assessed using established methods

Fig. 2 The TRUST tool part 2: assessing the impact of uncertainties on cost effectiveness

Box 2 Application of TRUST in a case study

NICE appraisal titled 'Pembrolizumab for treating relapsed or refractory classical Hodgkin lymphoma' (2018)

Pembrolizumab was compared with standard of care in patients who did and did not receive prior autologous SCT.

This NICE appraisal suffered from a lot of uncertainty (Table 2).

Most impactful uncertainty locations: model structure, transition probability and relative effectiveness estimates.

Model structure: the main uncertainty was that the time at which patients would receive allogeneic SCT was modelled as a fixed time point (bias). Not included in PSA or scenario analysis (alternative time point submitted later upon request).

Transition probability estimates: distributions to model OS were selected without appropriate justification (methods). Not included in the PSA but explored in scenario analysis.

The company deemed their own trial to be immature for modelling OS (imprecision), and the estimation of post-allogeneic SCT OS was based on a study that included 13 patients (imprecision). Both included in PSA.

Relative treatment effectiveness: using a naïve comparison for two single-arm studies (methods and bias). Not in PSA, but in scenario with matched-adjusted indirect comparison.

Comparator data had a mixed population of patients who did and did not receive prior autologous SCT (bias). Not explored in PSA or scenarios (scenario with alternative data for one population submitted later upon request).

Unknown impact on cost effectiveness: context/scope and selection of evidence.

Context/scope: comparators were omitted, best supportive care (unavailability of data) and nivolumab due to direction by NICE (unavailability and methods). No PSA or scenarios.

Selection of evidence: a list of health economic publications on this topic was not provided (transparency). No PSA or scenarios.

Conclusions:

1. Example of the increasing number of applications for reimbursement based on single-arm evidence and immature survival data.
2. Most impactful uncertainties were issues with bias and indirectness, methods and imprecision and were not always included in the PSA.
3. Reaching a reimbursement decision under such circumstances is potentially associated with a significant risk.

NICE National Institute for Care and Excellence, *OS* overall survival, *PSA* probabilistic sensitivity analysis, *SCT* stem cell transplant

Table 2 Illustration of TRUST in pembrolizumab for Hodgkin lymphoma case study

	Sources of uncertainty					Impact on cost effectiveness		
	Transparency	Methods	Imprecision	Bias	Unavailability	Uncertainty not reflected in PSA?	Uncertainty not explored in scenario analysis?	High impact on cost effectiveness?
Context / scope		Exclusion of comparator from scope	/////	May not represent UK clinical practice	No data on potential comparator led to its exclusion	/////	Yes	Unknown
Model structure			/////	AlloSCT assumed at one fixed time point. One relevant health state excluded	/////	Yes	Yes	Likely high
Selection of evidence	Health economic literature not provided	Intransparent	/////	Intransparent	Incomplete results from health economic review	/////	Yes	Unknown
Inputs — Effectiveness		Selection of distributions for OS lacked justification	Immature OS, small sample size for post-alloSCT	Only survey data on proportion of responders		Yes	Yes	Likely high
Relative effectiveness		Naïve unanchored indirect comparison, mixed comparator arm	Few events for OS of comparator	Single-arm study, comparator data mixed for two populations		Yes		Likely high
Adverse events								Likely no impact
Utilities		Use of one time-point utility data, despite repeated measurements		Different sources used for utility and progression decrement			Yes	Likely low
Resource use & costs				Mix of chemo-regimens based on assumption				Likely low
Implementation			/////	/////	/////	/////		
Outcomes		/////	/////	/////	/////	/////		

AlloSCT - allogeneic stem cell transplant, OS - overall survival, UK - United Kongdom

---

[26]. When a financial scheme is an option, phase II is repeated with the scheme incorporated in the modelling. Phases I and II are within the responsibility of the analyst/reviewer.

Lastly, the decision maker will appraise the risk in phase III and determine which, if any, risk-management options such as Managed Entry Agreement (MEA) schemes, are most suitable and arrive at a recommendation. The TRUST tool is positioned in phases I and IIa.
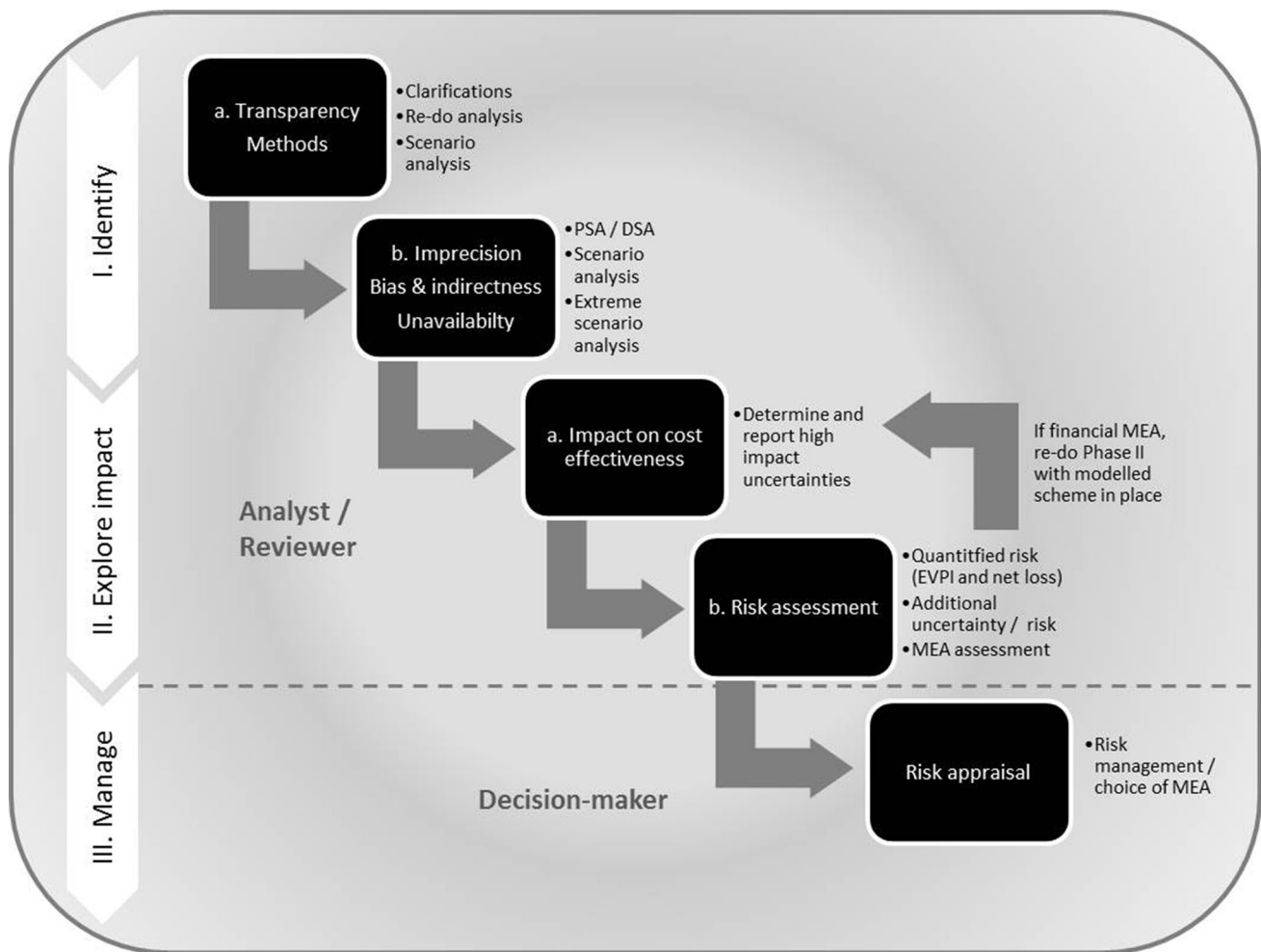
Fig. 3 The process of uncertainty and risk assessment. *DSA* deterministic sensitivity analysis, *EVPI* Expected Value of Perfect Information, *MEA* Managed Entry Agreement scheme, *PSA* probabilistic sensitivity analysis

4 Discussion

We developed the TRUST tool for systematically identifying, assessing, and reporting uncertainty in health economic models. To the best of our knowledge, this is the first tool to enable the transparent identification, assessment, and reporting of uncertainties in different model aspects in one place and the first to highlight the different sources of uncertainties. This is important because our experience and the literature indicate shortcomings in how uncertainty information is currently presented in health economic submissions [13]. Furthermore, we hope that TRUST will, in the future, prevent uncertainties due to bias, indirectness, and unavailability from being overlooked, an issue that may cause a misrepresentation of risk. Implementation of TRUST in six case studies showed that some of the most impactful uncertainties were issues not with imprecision but with bias and indirectness. A re-thinking of how we consider uncertainties in models appears warranted, especially given that increased

leniency in evidential requirements and handling of evidence has been observed, at least in the UK setting [30]. Knowledge about sources of uncertainty is also important because it helps recognise that different sources of uncertainty may call for different types of analyses. We hope this tool will contribute to the systematic identification, assessment, and reporting of uncertainties in HTA and therefore to more informed and transparent decision making in the face of uncertainty.

TRUST borrows from a range of uncertainty frameworks in different disciplines, which we view in two categories: those within and those outside the field of health economics. Compared with existing uncertainty frameworks outside the field of health economics [16, 18, 27, 29], TRUST has been adapted to match the needs of HTA analysts and decision makers, for instance with a focus on epistemic uncertainty and it being aligned with typical structures of health economic reporting. TRUST adds to the frameworks inside the field of health economics [5, 28] by considering different

sources of uncertainty and requiring an assessment for each model aspect in one place. Where other uncertainty tools have already been completed for a certain decision problem, for instance, a GRADE assessment is available for the studies informing relative treatment effectiveness, results of this can be used to fill in TRUST, if consideration is given to whether the decision problem settings are truly equivalent and the source is trustworthy. TRUST also complements existing tools on model validation [31, 32]. Assessing uncertainty is to be seen as a separate albeit related exercise to validating a model or model aspects. A model that is rigorously validated and even considered valid will still suffer from uncertainty. We therefore wish to caution against assuming uncertainty is resolved when validation has taken place.

A strength of this work is that TRUST has its foundation in the uncertainty framework literature. The content of this tool was further validated through feedback from experts with different backgrounds relevant to the HTA decision-making process during its development. Experts thought this tool was of potential value in uncertainty assessment but thought the proof would be in its use in actual decision making. Application in six real-world HTA case studies served as further validation: all uncertain model aspects could be registered without difficulty, and inter-rater variability was reduced after the tool's amendment. A caveat is that these six case studies may not represent all the unique features that may be observed in reimbursement applications; however, they did cover, amongst other features, single-arm studies, the use of indirect treatment comparisons, clinical evidence from different settings, structural uncertainty, long-term extrapolation beyond the available follow-up, issues related to the use of intermediate outcomes, and how these related to health states (Table 1).

We acknowledge that the part of TRUST concerning the impact of uncertainties on cost effectiveness may elicit subjective answers, especially when uncertainty is not sufficiently explored through PSA or scenario analysis. However, we wish to highlight that, where these considerations are not made explicit, they will be made implicitly, and potentially in a non-systematic and opaque fashion. We consider an explicit subjective judgement to be superior to no judgement at all or to a subconscious judgement that may implicitly drive decision making. We recognise that representing the impact of uncertainty can be challenging, both methodologically and in terms of time constraints. We argue for parameterising uncertainty as much as possible, and, if this is infeasible, we propose representation of uncertainty in terms of ranges of the incremental cost-effectiveness ratio or net benefit statistics. In addition, results from plausible and extreme scenario analysis and qualitative statements should be reported [13, 18]. A limitation of this work is the transferability of TRUST to settings other than the UK or the Netherlands. However, the tool can be tweaked in

terms of the model aspects covered and response options if interested individuals wish to apply it outside these settings and believe that model aspects do not cover all aspects considered. We believe the sources of uncertainty to be fairly universal and grounded in the literature so would not advise tweaking those without good justification.

A further limitation relates to the added time it may take to use TRUST. Interviewees suggested that, once in the process of reviewing evidence submitted by companies, or indeed developing an evidence submission, it might only take an hour to complete. However, we also found that completion could be more difficult and therefore time consuming if the evidence was not easily accessible. If it was not possible to complete TRUST within the review process, we consider this to be a potentially important finding that should be communicated, as it may mean that the evidence and the model could not be fully assessed. Furthermore, if explicit consideration of uncertainty is desired in a world that relies on increasingly sparse evidence, processes may have to be adjusted to allow for these additional analyses.

Knowledge of uncertainty is crucial for risk management [15, 18]. In the process of HTA, communication of uncertainty is therefore key, but is also a challenge. The TRUST tool can help identify and report uncertainty, but effective communication to decision makers may require effective summaries and visualisation of uncertainty and risk. Communicating uncertainty does not necessarily undermine trust [18] but can indeed be appreciated by the target audience [33]. It can and should, in our opinion, be viewed as a duty of science [34], as Høeg expressed, "That is what we meant by science. That both question and answer are tied up with uncertainty, and that they are painful. But that there is no way around them. And that you hide nothing; instead, everything is brought out into the open" [35]. Further research should therefore be performed on the communication of results from TRUST to committees and decision makers, considering whether it changes perceptions and/or behaviours [18, 36]. Here, both outcomes would be desirable: we wish to effect change in the perception of uncertainty and risk in an appraisal; ideally, this would translate into better risk management. Further research should also focus on the link between the TRUST tool outcomes and risk-management options.

5 Conclusion

The TRUST tool enables the systematic identification, assessment, and reporting of uncertainties in health economic decision models and may contribute to more informed and transparent decision making in the face of uncertainty. Further research is needed on uncertainty and risk communication and the link to risk-management options.

Compliance with Ethical Standards

References

1. Lipska I, Hoekman J, McAuslane N, Leufkens HGM, Hovels AM. Does conditional approval for new oncology drugs in Europe lead to differences in health technology assessment decisions? Clin Pharmacol Ther. 2015;98(5):489–91.

2. Grutters JPC, van Asselt MBA, Chalkidou K, Joore MA. Healthy decisions: towards uncertainty tolerance in healthcare policy. Pharmacoeconomics. 2015;33(1):1–4.

3. Makady A, van Veelen A, de Boer A, Hillege H, Klungel OH, Goettsch W. Implementing managed entry agreements in practice: the Dutch reality check. Health Policy. 2019;123(3):267–74.

4. Briggs AH. Handling uncertainty in cost-effectiveness models. Pharmacoeconomics. 2000;17(5):479–500.

5. Claxton K. Exploring uncertainty in cost-effectiveness analysis. Pharmacoeconomics. 2008;26(9):781–98.

6. Davis C, Naci H, Gurpinar E, Poplavska E, Pinto A, Aggarwal A. Availability of evidence of benefits on overall survival and quality of life of cancer drugs approved by European Medicines Agency: retrospective cohort study of drug approvals 2009–13. BMJ. 2017;359:j4530.

7. Dickson R, Boland A, Duarte R, Kotas E, Woolacott N, Hodgson R, et al. EMA and NICE appraisal processes for cancer drugs: current status and uncertainties. Appl Health Econ Hea. 2018;16(4):429–32.

8. Gyawali B, Hey SP, Kesselheim AS. Assessment of the clinical benefit of cancer drugs receiving accelerated approval. JAMA Intern Med. 2019;179(7):906–13.

9. Sabry-Grant C, Malottki K, Diamantopoulos A. The cancer drugs fund in practice and under the new framework. Pharmacoeconomics. 2019;37(7):953–62.

10. Grimm SE, Armstrong N, Ramaekers BLT, Pouwels X, Lang S, Petersohn S, et al. Nivolumab for treating metastatic or unresectable urothelial cancer: an evidence review group perspective of a NICE single technology appraisal. Pharmacoeconomics. 2019;37(5):655–67.

11. Grimm SE, Fayter D, Ramaekers BLT, Petersohn S, Riemsma R, Armstrong N, et al. Pembrolizumab for treating relapsed or refractory classical hodgkin lymphoma: an evidence review group perspective of a NICE single technology appraisal. Pharmacoeconomics. 2019;37(10):1195–207.

12. Howard R. Information value theory. IEEE Trans Syst Sci Cybern. 1966;SSC2(1):22–6.

13. Bilcke J, Beutels P, Brisson M, Jit M. Accounting for methodological, structural, and parameter uncertainty in decision-analytic models: a practical guide. Med Decis Mak. 2011;31(4):675–92.

14. Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. BMJ. 2008;336(7650):924–6.

15. Klinke A, Renn O. A new approach to risk evaluation and management: risk-based, precaution-based, and discourse-based strategies. Risk Anal. 2002;22(6):1071–94.

16. van Asselt MBA, Rotmans J. Uncertainty in integrated assessment modelling: from positivism to pluralism. Clim Change. 2002;54:75–105.

17. Balshem H, Helfand M, Schunemann HJ, Oxman AD, Kunz R, Brozek J, et al. GRADE guidelines: 3. Rating the quality of evidence. J Clin Epidemiol. 2011;64(4):401–6.

18. van der Bles AM, van der Linden S, Freeman ALJ, Mitchell J, Galvao AB, Zaval L, et al. Communicating uncertainty about facts, numbers and science. R Soc Open Sci. 2019;6(5):181870.

19. Briggs A, Sculpher M, Claxton K. Decision modelling for health economic evaluation. Oxford: Oxford University Press; 2006.

20. Iglesias CP, Thompson A, Rogowski WH, Payne K. Reporting guidelines for the use of expert judgement in model-based economic evaluations. Pharmacoeconomics. 2016;34(11):1161–72.

21. Soares MO, Sharples L, Morton A, Claxton K, Bojke L. Experiences of structured elicitation for model-based cost-effectiveness analyses. Value Health. 2018;21(6):715–23.

22. Bojke L, Claxton K, Sculpher M, Palmer S. Characterizing structural uncertainty in decision analytic models: a review and application of methods. Value Health. 2009;12(5):739–49.

23. Strong M, Oakley JE. When is a model good enough? deriving the expected value of model improvement via specifying internal model discrepancies. SIAM/ASA J Uncertain Quantif. 2014;2(1):106–25.

24. Strong M, Oakley JE, Chilcott J. Managing structural uncertainty in health economic decision models: a discrepancy approach. J R Stat Soc C-Appl. 2012;61:25–45.

25. O'Hagan A, Oakley JE. Probability is perfect, but we can't elicit it perfectly. Reliab Eng Syst Safe. 2004;85(1–3):239–48.

26. Grimm SE, Strong M, Brennan A, Wailoo AJ. The HTA risk analysis chart: visualising the need for and potential value of managed entry agreements in health technology assessment. Pharmacoeconomics. 2017;35(12):1287–96.

27. Walker WE, Harremoes P, Rotmans J, Van der Sluijs JP, Van Asselt MBA, Janssen P, et al. Defining uncertainty: a conceptual basis for uncertainty management in model-based decision support. Integr Assess. 2003;4(1):5–17.

28. Briggs AH, Weinstein MC, Fenwick EAL, Karnon J, Sculpher MJ, Paltiel AD, et al. Model parameter estimation and uncertainty analysis: a report of the ispor-smdm modeling good

research practices task force Working Group-6. Med Decis Mak. 2012;32(5):722–32.

29. Bouwknegt M, Havelaar A. Uncertainty assessment using the NUSAP approach: a case study on the EFoNAO tool. EFSA supporting publication. 2015;EN-663.

30. Charlton V. NICE and Fair? Health technology assessment policy under the UK's National Institute for Health and Care Excellence, 1999–2018. Health Care Anal. 2019. https://doi.org/10.1007/s10728-019-00381-x.

31. Eddy DM, Hollingworth W, Caro JJ, Tsevat J, McDonald KM, Wong JB, et al. Model transparency and validation: a report of the ISPOR-SMDM modeling good research practices task force-7. Med Decis Mak. 2012;32(5):733–43.

32. Vemer P, Ramos IC, van Voorn GAK, Al MJ, Feenstra TL. AdViSHE: a validation-assessment tool of health-economic models for decision makers and model users. Pharmacoeconomics. 2016;34(4):349–61.

33. Versluis E, van Asselt MBA, Kim J. The multilevel regulation of complex policy problems: uncertainty and the swine flu pandemic. Eur Policy Anal. 2019;5(1):80–98.

34. Saltelli A, Funtowicz S. Evidence-based policy at the end of the Cartesian dream: the case of mathematical modelling. In: Guimaraes Pereira A, Funtowicz S, editors. Science, philosophy and sustainability: the end of the cartesian dream. London: Routledge; 2015.

35. Høeg P. Borderliners. Toronto: Delta Publishing; 1993.

36. Fischhoff B, Brewer N, Downs J. Communicating risks and benefits: an evidence-based user's guide. FDA: FDA; 2011.