CONSIDERING AND COMMUNICATING UNCERTAINTY IN HEALTH TECHNOLOGY ASSESSMENT

HTAi Global Policy Forum 2021 Background Paper

February 2021

Introduction

The purpose of this background paper is to inform the discussion at the HTAi Global Policy Forum (GPF) meeting which, for the first time, will be held virtually across the world in February 2021. The topic chosen for the meeting is "Considering and Communicating Uncertainty in Health Technology Assessment (HTA)" a theme felt to be especially timely during the COVID-19 pandemic. The topic, and overall outline, was selected by HTAi GPF member representatives in early 2020 and further refined through virtual topic scoping and breakout meetings held during June and July of 2020. The meeting's main aim is to discuss, at a strategic and policy level, the impact of uncertainty on deliberations and outputs in HTA and examine how multiple types of uncertainty are best handled and then communicated to multiple stakeholders. It will consider whether global approaches and/or considerations for the treatment of uncertainty can be developed, acknowledging the contextual importance of attitude to uncertainty and risk and its possible variation across settings. The intention is that the focus of the GPF discussions remain policy-oriented, rather than at a detailed operational or methodological level.

To support the aims of the GPF meeting, this paper presents an overview of key methods and considerations related to uncertainty available in the published literature. This is supplemented by the concerns identified by HTA users, producers and other stakeholders as well as those identified by GPF members. These concerns were elicited during 20 expert informant interviews conducted by the GPF Scientific Secretary and Chair, with experts selected to represent a variety of stakeholder perspectives and "insider" knowledge (see the Acknowledgements for further details). In addition to this, a survey of the current not-for-profit members of the GPF was conducted to determine what explicit methods are in place for considering and communicating uncertainty in their respective organizations. A total of 16 responses from 14 not for profit member organizations were received. Review and further input from the HTAi GPF Organizing Committee, the wider HTAi GPF membership, and members of the HTAi Board was also received during the development of this background paper.

It is important to highlight that management and communication of uncertainty as described in this paper is primarily through the lens of the high-income country. This is reflective of the majority of the GPF membership as well as the published literature. However, the same conditions and concerns are also present for low- and middle-income countries. Indeed, the HTAi Asia and Latin America Policy Fora have had wide ranging topics that have implicitly and explicitly included conversations around uncertainty in those regions (https://htai.org/policy-forum/). In addition, the focus of this paper and the GPF discussions will be primarily from the perspective of the HTA community (i.e. those concerned with using or producing HTAs).
We acknowledge that other perspectives of uncertainty (such as that of regulators) will differ from those of the HTA community due to different remits, functions and scopes. We also note that there are areas of significant innovation and change in these other arenas (particularly in the regulatory field), which is seeing collaborative efforts to streamline activities and accelerate regulatory processes. Current examples include Project Orbis between the US, Canada, Singapore, Switzerland, Brazil and Australia and MIT's NEW Drug Development ParadIGmS (NEWDIGS). Project Orbis aims for concurrent submission, HTA review and regulation for oncology products. NEWDIGS is a multi-stakeholder collaborative focused on enhancing the capacity of the global biomedical innovation system to deliver better therapeutics faster. While the work in this arena is of great significance to the HTA community and one that has consequences, particularly for technology manufacturers who deal with both perspectives in parallel, it is beyond the scope of this HTAi GPF.

Prior Policy Fora Topics Relevant to Uncertainty

The January 2020 GPF discussed deliberative processes in HTA(1) and at this meeting, GPF members agreed on three core principles of deliberative processes--transparency, inclusivity and impartiality. These are closely linked with uncertainty, as it is during the deliberative process that the uncertainty in the available data (hereafter referred to as "input uncertainty") is debated and decisions on how to manage the uncertainty (and what level of uncertainty can be tolerated) are made. Being transparent about what uncertainty exists and how this impacts the decision, communicating the impact of uncertainty to all relevant stakeholders, and an impartial approach to managing uncertainty (i.e., consistency and predictability in how the uncertainty is

handled) are also relevant principles for the current topic. There have been other GPF topics that are closely linked to the current area of focus, further demonstrating how embedded the concept of uncertainty is in all facets of HTA. For example, in 2019, the GPF considered "Real World Evidence (RWE) in the Context of HTA", (2) noting that a key potential use of RWE is to generate evidence that can reduce uncertainty after initial technology adoption, but that RWE itself can be considered unreliable if parameters for its use are not appropriately outlined. In 2016, the GPF considered the changing paradigm for HTA (3) and touched on many aspects related to addressing uncertainty such as early scientific dialogue. Prior to this, the GPF discussed managed entry agreements (MEAs) in 2010 (4) and coverage with evidence development (CED) in 2007, (5) two vehicles for handling uncertainty through linking price to value and collection of additional evidence.

The intention is to not repeat and return to any of these topics in detail during the 2021 GPF discussions but for GPF members to draw on these resources as required. Table 1 lists some questions and topics that will be most relevant for the 2021 GPF.

Table 1 Selected key questions relevant to the 2021 HTAi GPF

Domain	Questions
Future of Input Uncertainty	How can the HTA community better prepare for increasing input uncertainty from the below, without sacrificing timely access: ▪ Accelerated regulatory approvals or other abbreviated regulatory processes? ▪ Highly specialized treatments (e.g. gene therapies)? What is the opportunity cost of resolving increasing uncertainty; How can HTA and regulators be better aligned in considering uncertainty? Who should influence who (or what can we learn from each other?) Are there particular types/areas of uncertainty that external stakeholders (e.g., patients and clinicians) can help to resolve?
Managing Uncertainty	What are the potential supporting actions to improve the consistency and predictability of management of uncertainty for stakeholders? What is the role of stakeholders in facilitating the understanding and management of uncertainty during committee deliberation? How is uncertainty conveyed to and considered by deliberative committees or other bodies; can this be standardized across jurisdictions? Are there conditions where additional uncertainty is universally acceptable (either by disease or by type of technology), and should standard criteria for their determination apply?
Communicating Uncertainty	What are the key considerations for communicating uncertainty (and tolerability of uncertainty) to different stakeholder groups, including patients, clinicians and payers? What are the current innovations in this space? What are the resource implications for communicating uncertainty? Are there lessons from the COVID-19 pandemic that can be applied?

Background

In the Cambridge English Dictionary, uncertainty is defined as: "not knowing what to do or believe, or not able to decide about something; not known or fixed, not completely certain"(6). This broad definition demonstrates the variety of uses and meanings of the word(7). While there is no single, standalone definition of "uncertainty" in the HTA Glossary(8), the concept is clearly articulated and inherent in many definitions and terms related to HTA methods and processes. By the very nature of what HTA is and what it seeks to do, uncertainty will always exist at some level (9). Therefore, considering uncertainty is a fundamental and inherent component of HTA. However, the process for how these concepts are dealt with is contextual and will vary across jurisdictions and by stakeholder perspective. The types and level of uncertainty, how uncertainty is considered and managed when arriving at recommendations or decisions, and how the various uncertainties are conveyed to multiple stakeholders are all critically important to consider; if any one of these elements is not considered carefully then trust in HTA findings will surely be reduced.

Conceptualizing Uncertainty

The complexities of considering and communicating uncertainty become more manageable by conceptualizing them using an "Input-Throughput-Output" (ITO) model (as utilised to great effect in the 2020 Global Policy Forum on Deliberative Processes(1)). The ITO model is often used to illustrate information processes and complex pathways of care, and it has similarities to other general descriptions of HTA systems and processes for well-informed policy making in health care(10). As with the previous HTAi GPF on deliberative processes, the ITO model provides a useful framework for considering different types of uncertainty and the roles that each play in HTA activities. This background paper will provide an overview of each of the domains; importantly, however, the discussions at the virtual meeting will focus primarily on the throughput and output domains, as input during the scoping process indicated that these domains would potentially benefit from development of a set of key considerations and/or recommendations. These may include but are not limited to: development of white paper(s), position statements, and other actions such as additions to the HTA Glossary on the definition of uncertainty in HTA.

Firstly, "input" can generally be considered the collection of material (evidence, information, and perspectives) that informs HTA activities. This sets the stage for consideration of "input uncertainties" by deliberative bodies in HTA and come primarily in the form of:

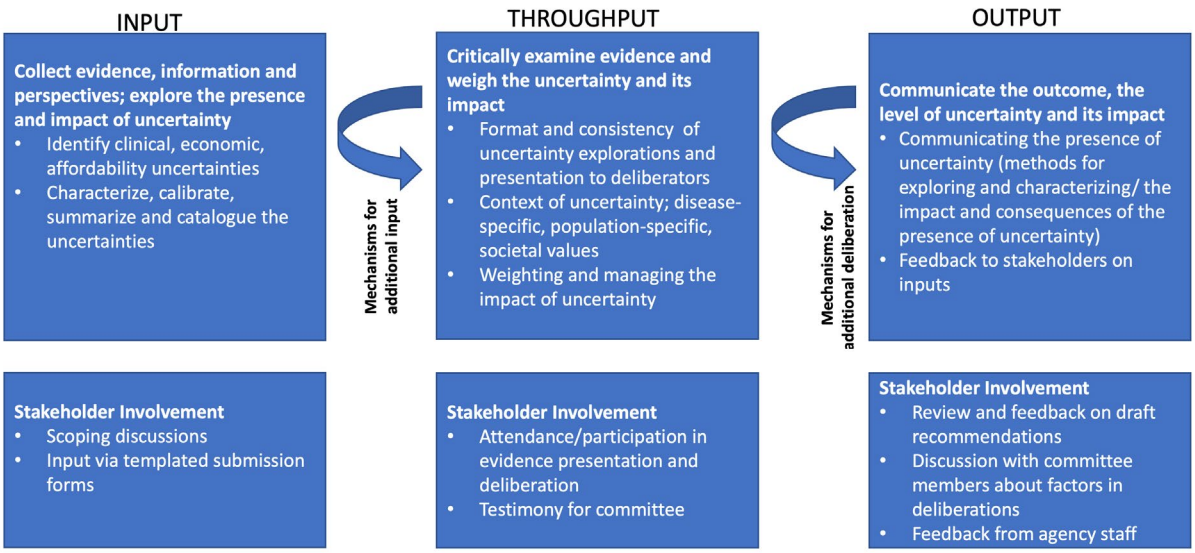
- clinical uncertainty from concerns regarding trial design, population, and generalisability/heterogeneity
- economic model "structural" uncertainty from model design and operation
- economic model "parameter" uncertainty from data that is used in the economic model (including temporal uncertainty from the extrapolation of shorter-term data over long-term time horizons)
- affordability uncertainty from estimates used to calculate budget impact of a technology in a healthcare system, often driven by uncertainty in clinical evidence and economic extrapolations

The "throughput" stage describes how the various input uncertainties are handled; in other words, the weighting of the facts, data, values and reasons that will lead to a collective judgement for a key HTA process (e.g., topic selection, scoping consultation, adoption decision or recommendation). This is the stage where the interplay between the clinical, economic and affordability uncertainties are considered, and consensus is sought, or votes are taken. Handling uncertainty during deliberations can vary according to perspective and societal values. Of course, the risk tolerance of each individual involved in the decision-making process will also vary based on their own interpretation of the key inputs, and this can result in further deliberative uncertainty. The effects of uncertainty considered at the throughput stage can have varying levels of influence on the resulting recommendations.

Finally, "output" refers to the way in which the level of uncertainty and its impact on the recommendations is communicated and any learning is consolidated. The importance of clearly and understandably conveying the types of uncertainty described above to decision-makers, patients, the media, and the general public should not be underestimated. Stakeholder-friendly methods of conveying the input and throughput

uncertainties are important to ensure that the resulting HTA recommendations, decisions, and specific actions taken to reduce uncertainty are understood and fundamentally trusted—this is not to say that all stakeholders will agree with the decision, but should be able to agree that any uncertainties were handled in a fair, impartial and trustworthy manner. As indicated in Figure 1, stakeholder involvement crosses all aspects of the ITO model.

Figure 1 – Representation of the Input-Throughput-Output Model in the context of uncertainty in HTA



Input Uncertainty

Input uncertainty can be expressed in terms of variation in the data collected for an HTA, elements of the evidence base for which data are not available, limitations on who the findings apply to and over what timeframe, and other examples. As such it falls into the broad categories of clinical, economic and affordability uncertainty. However, there can be further uncertainties around other aspects of HTA (as HTA is more than just a consideration of the clinical and economic evidence (11)). Examples include uncertainty in how and when a technology should be implemented in practice (implementation uncertainty), whether the best price is being offered for the technology (i.e. pricing/relational uncertainty often related to uncertainty around development and production costs, competitive market development, turnover and other factors such as patent duration), and uncertainty around societal values for when circumstances exist and conditions are met for greater uncertainty to be acceptable (value uncertainty) are all additional areas of uncertainty that appraisal committees often must consider. The section below focusses more on traditional types of input uncertainty but these additional concepts and "meta" uncertainties, as well as the relative weight given to each during deliberations, are often at play.

First described by psychologists Luft and Ingham in 1955 using the Johari window analysis technique(12), and in accordance with terminology popularised by Donald Rumsfeld (former US Secretary of Defence), there are "known knowns, known unknowns and unknown unknowns". Of these areas of uncertainty, it is the known unknowns that are most relevant at the point at which an HTA is conducted, given the expectation that resolving uncertainty that is known to exist would make the evidence considered by HTA bodies more complete; this is otherwise known as epistemic uncertainty(7). This form of uncertainty is particularly challenging for devices, diagnostics, technologies for ultra-rare conditions, highly specialised/precision technologies, and other potentially innovative technologies that have come to market on an accelerated regulatory pathway.

The subsections below describe each major type of input uncertainty (clinical, economic model and affordability) followed by descriptions of some of the most common methods seen to characterize the input uncertainty that is present.

Clinical Uncertainty

Uncertainties regarding the effects of treatments are inevitable(13); when a technology is first tried in humans, the effects can be anticipated but cannot be known. The evidence generated may have wide confidence intervals (i.e., data variation) or an apparent benefit related to some outcomes or subgroups of patients but not others. Variability from person to person may be challenging to understand. Added to this, clinical trials cannot provide a complete picture of how a technology will work in practice. In further framing clinical uncertainty, the Population, Intervention, Comparator, Outcomes, Timing, Setting (PICOTS) framework is a useful way to consider how uncertainties may be introduced into the clinical evidence base(14). Those familiar with the use of the PICOTS framework in HTA will recognize some of the more common uncertainties that HTA bodies must grapple with. For example, the trial population may not be reflective of the target population in practice (for example a trial in relatively healthy volunteers will likely have different effects in a frail older population with comorbidities). The intervention itself will be delivered consistently within a clinical trial setting but may have lower adherence (and therefore reduced effectiveness) in practice. The comparator in the study may not reflect standard of care in the country in which the HTA is taking place, and therefore the relative effectiveness may be uncertain. The outcomes in the study setting may not reflect what is important to clinicians and patients or may be surrogate to the actual outcome of interest in the HTA, therefore the actual benefit of a technology may be uncertain. The timing or duration of the study may be too short to be certain about the long-term benefits (and adverse effects) associated with the technology. The setting of the study may not reflect the setting of delivery of the technology in the real world (e.g., academic or highly specialized centers vs. community practice), and therefore the effects in a real-world setting may be uncertain.

In a recent study by Vreman et al.(15) that looked at the greatest areas of concern for HTA agencies and regulatory bodies, the greatest concern was uncertainty around the long-term effectiveness of a technology,

a finding that did not differ by type of agency or geography. This was echoed by the responses to the Not-for-Profit GPF member survey, but in addition it was noted that uncertainty in the structure and parameters used in the economic model was of great concern.

The study design can also create an extra level of uncertainty, with small and/or single arm studies creating a divide in terms of how acceptable/comfortable people are with them – particularly for estimating clinical efficacy. It should be noted that there is typically greater comfort with other estimates coming from non-clinical trial data (such as quality of life or utility data). There are also an increasing number of studies where participants in the placebo arm of RCTs are permitted to switch to the intervention arm, thereby essentially reducing a component of the RCT to a single-arm study(16). Clinicians (and therefore many technology appraisal committee members) have been trained to consider statistical significance as a critical demonstration of benefit or harm and it is challenging to go against an ingrained belief system (Expert informant); study design issues such as these may limit or preclude the use of significance testing. On the other hand, statistically significant results may be observed for an interim or surrogate outcome measure for which the clinical significance of the findings is unknown.

Finally, in an era of increasing accelerated regulatory pathways, it is becoming increasingly common for much of the input data from early clinical trials to be provided as data in confidence to regulators and HTA agencies. This brings additional challenges as data are not published and peer-reviewed, and are not identifiable through literature searching. This leads to an increased reliance on trusting the technology manufacturer to provide nearly all of the data on which the HTA is based. In the response to the Not-for-Profit GPF member survey, there was an even split between agencies that do and do not accept data in confidence from technology manufacturers.

Characterizing Clinical Uncertainty

There are a number of methods and approaches available to attempt to characterize uncertainty in the clinical parameters. The simplest way in which to do this is with a qualitative (text) summary. Here the level of uncertainty can be listed and synthesized narratively for the decision maker to read and understand; however, this does not provide an intuitive and easily digested summary of the uncertainty, nor does it attempt to score or quantify the impact of the uncertainty. A variety of technical approaches have also been used, the most common of which are described below.

QUALITY MEASURES

In addition to qualitatively describing the uncertainty, assessing and providing a "score" of the uncertainty present in the clinical inputs is a technique commonly seen by many HTA agencies. Some popular examples of such measures include GRADE (Grading of Recommendations, Assessment, Development and Evaluation), the Cochrane Risk of Bias tool, and the Effective Health Care Methods Guide (used by AHRQ in the US), as well as others. These tools do not reduce any uncertainty but rather provide a means to evaluate the likely impact of clinical uncertainty. The GRADE system is a transparent framework for developing and presenting summaries of evidence and serves to provide a systematic approach for making clinical practice recommendations(17). It is the most widely adopted tool for grading the quality of evidence for making recommendations, with over 100 organizations worldwide having officially endorsed the use of GRADE(18). Reviewers assign one of four levels to categorize the strength of evidence (also known as certainty or quality of evidence): very low, low, moderate, and high. The Cochrane risk of bias tool assigns ratings of low, high, and unclear risk of bias in 6 specific domains for a given study, the results of which can be compared across all studies in a sample (19). The Effective Health Care approach is conceptually very similar to GRADE and the strength of evidence across studies (for each outcome) receives a high, moderate, low or insufficient rating (20) based on domains such as consistency across studies, directness of the outcome of interest, precision of the findings, and others. Typically randomized controlled trials are considered superior to observational studies as sources of evidence (given a greater a number of selection and other attendant biases possible in the latter), but the certainty for a given study may be affected by a host of factors.

Criticisms of quality measures such as GRADE and others generally focus on the notion that they are

essentially subjective and cannot be implemented consistently. This criticism is noted particularly for devices, diagnostics and innovative technologies. It does however provide a reproducible and transparent framework for grading the level of uncertainty in clinical inputs, and a summary that is generally easily understood by committee members. Additional research to address the limitations of the GRADE tool, such as exploring the inherent biases and inter-rater variances/reliability can also be of value.

In the response to the survey, most respondents stated that they used a specified checklist (such as GRADE) or adapted versions of GRADE (for example at the Center for Healthcare Quality and Control, [CHQC] in Russia where they have created a numerical score for GRADE categories and levels of evidence that they then use to inform adoption decisions). The Institute for Clinical and Economic Review (ICER) was the only organization to note in the survey that they have created their own Evidence Rating Matrix™, in which uncertainty is represented on a distinct axis (i.e., low, moderate, or high certainty) in addition to the magnitude of potential net health benefit. See Annex 1 for copies of the CHQC and ICER checklists.

SURROGATE ENDPOINTS

As mentioned, accelerated regulatory pathways and a desire for more rapid HTAs (to facilitate faster treatment access for patients) is leading to an increased reliance on evidence from surrogate endpoints. In reviews of the effects of using surrogate endpoints, it was found that, on average, these measures overestimate treatment effects(21). The strength of the evidence for the surrogate can be evaluated systematically, and commonly three levels are ascribed: Level 1 is clinical trial evidence of treatment effects on the surrogate corresponding to effects on the patient-related outcome; Level 2 is evidence from epidemiological or observational studies that demonstrates a consistent relationship between the surrogate and the patient-related outcome; and Level 3 is biological plausibility from pathophysiological studies or from understanding of the disease process(22). Validation methods (calculating the correlation of the effects on the surrogate and clinical endpoint, or the R^2 value) and validation values (that is, accepting only surrogates with a correlation value above a certain level, such as an R^2 of greater than 0.49 as employed by IQWiG) can also be used. Challenges arise particularly where a technology is developed with a novel mechanism of action and evidence around surrogacy of outcomes and effect does not yet exist at the time of the HTA assessment.

In a recent study by Grigore et al.(23), a review of HTA agency policies found that 40% of agencies studied had methodological guidelines that made specific reference to consideration of surrogate outcomes. In the results of this review, the HTA agencies with specific guidance noted the lack of methodological consensus around the level of evidence necessary, all agencies highlighted a preference for randomized trial data to support the association in the treatment effect between the surrogate and final endpoint, including the use of meta-regression analysis methods.

Economic Model Uncertainty

Uncertainty around economic evaluation in HTA can be broadly split into structural and parameter uncertainty. These types of input uncertainty are described in more detail within the sub-sections below. There can also be additional uncertainty introduced around the heterogeneity and stochastic variance of the economic model (24). Stochastic (first-order) uncertainty relates to the fact that individuals facing the same probabilities will experience the effects of a disease or an intervention differently due to random variation. This type of uncertainty is informed by confidence intervals and ranges of treatment effects and is typically primarily of concern for rare diseases and small patient populations. Heterogeneity describes the variability between the responses to an intervention that can be explained by the differences in the demographic and/or clinical characteristics of patients (for example age-specific results for the impact of an intervention on mortality). This type of uncertainty is informed by subgroups and stratification of a patient population.

As indicated, there is a wide range of input uncertainties that can be present within a HTA and must be considered. Most HTA agencies that consider cost-effectiveness as a factor are seeking a primary point estimate of the incremental cost-effectiveness ratio to compare to a threshold (or threshold range) and inform a recommendation. This poses challenges given that clinical, parameter, and structural uncertainty may all affect the derivation of that single point estimate. Sometimes the uncertainty in the clinical evidence

base may be too great to fully consider cost-effectiveness, for example in the assessment of a treatment for a very rare condition with poorly-understood outcomes. Typically, however, a point estimate for the cost-effectiveness is sought, and the major uncertainties summarized and presented. As with clinical input uncertainty, this can be done qualitatively or using technical approaches. Some of the most common approaches presented to HTA agencies are detailed below.

PARAMETER UNCERTAINTY

All economic models have parameters that must be estimated, and economic models can only be as reliable as the parameters (inputs) that they utilize (25). A key area of parameter uncertainty (also known as a second-order uncertainty) specifically relates to the fact that the probabilities and other estimates assigned to an economic model are uncertain because they are observed (for example within a clinical trial) and then estimated. The sample size of the observed dataset is therefore a key consideration in determining parameter uncertainty (smaller trials typically result in wider confidence intervals around a point estimate). Parameter uncertainty can therefore link directly with the areas of clinical uncertainty described above, and any problems with a study's internal or external validity and generalizability to a real-world setting are fed through in model parameter uncertainty.

Other parameter estimates may be subject to significant uncertainty, albeit for different reasons. Cost and utility estimates may be uncertain because they are derived from external sources and may not align with the target population for modeling, for example. Parameter uncertainty also arises when there are conflicting estimates from multiple studies, or when there are no available data for a required value and expert opinion must be utilized. In these instances, additional parameter uncertainty can arise if parameters are not chosen in an evidence-based way (for example if estimates are "cherry-picked" from the clinical evidence base or from asking a few key opinion leaders) or unrealistic assumptions are used(26). As per the Not-for-Profit GPF member survey response, the most common methods for characterizing parameter uncertainty are deterministic and probabilistic sensitivity analyses.

DETERMINISTIC SENSITIVITY ANALYSES

Deterministic sensitivity analysis (DSA) considers the impact of individual economic model parameters on the cost-effectiveness ratio. One or more parameter inputs can be changed manually to evaluate what effect the change in the parameter(s) has on the result. The range that the parameters are varied across is usually pre-specified (often representing the upper and lower limits of the 95% confidence interval or some other measure of variance) (27). Univariate sensitivity analysis refers to the modification of a single parameter at a time, and two-way sensitivity analysis involves modification of two parameters simultaneously; less commonly, multivariate sensitivity analysis involves the modification of several parameters at the same time (however, usually no more than five). The results of these sensitivity analyses can be presented as stack bar charts or as a "tornado diagram", in which those parameters appearing with the greatest impact on model results appear at the top, with subsequent parameters with lower sensitivity presented below; the resulting figure resembles the funnel cloud of a tornado (28).

The main difficulties with conducting deterministic sensitivity analyses arise when the ranges of the parameters are highly uncertain (for example a study may be small and have very wide confidence intervals); this may result in estimates that are not clinically plausible or even relevant, and including these values in a DSA can lead to a skewed perception of the impact of the uncertainty.

A variant of deterministic sensitivity analyses known as "threshold analyses" are also increasingly common and are used to assess the 'tipping-point' of an input parameter. For example, at what value of parameter X does the output change to the point that the recommendation based on the result would be altered? An increasing number of HTA agencies use these types of analyses, particularly where price negotiation is within the remit of the agency. In this situation, threshold analyses are commonly used to determine the price at which a cost-effectiveness threshold is reached, which can then be used for price and/or discount negotiations.

PROBABILISTIC SENSITIVITY ANALYSES

Probabilistic sensitivity analyses (PSA) are a form of sensitivity analysis in which all parameter inputs are varied at once. In probabilistic sensitivity analyses, rather than individual parameter estimates or points on a range, the parameters are sampled from a representative distribution, either an observed one from patient-level data or a distributional form that fits the data well (29). The model is then run over multiple iterations (typically 1,000 or more), each producing a unique cost-effectiveness estimate that can be compared to existing decision-making thresholds.

Cost-effectiveness acceptability curves have become a common way to present the results of a PSA(30), but other approaches such as incremental/net health benefit curves, rankograms and scatter plots also exist(31). The main concerns when conducting PSA are whether the number of simulations that are performed is sufficient (with little explicit guidance on this provided) (32). In addition, even the most robust PSA cannot adequately address structural uncertainty; a greater number of iterations will not mitigate uncertainty introduced if the model does not realistically portray disease trajectory or typical clinical practice, for example. While PSA is often described as an acceptable form of sensitivity analysis by HTA agencies (33), our expert informant interviews suggested that HTA agencies are only typically using deterministic sensitivity analyses routinely and that PSA are relatively under-utilized. The recent NICE methods guide review consultation has recognized this under-utilization and has identified PSA as an area requiring a "major change" in its methods update (34).

CALIBRATION OF EXTRAPOLATION

In the clinical trial setting, the true long-term effects of a technology are rarely ever observed, as clinical trials tend to be no more than two years in duration. Many technologies assessed by HTA agencies are chronic therapies intended for lifetime use. To account for this, the observed effects are typically extrapolated to provide a best-case estimate of what is likely to be the longer-term outcome of use of a technology (35). The accuracy of any extrapolation depends on the reliability of the modelling, with various options available to provide and calculate a "best fit" for the data(36), often using available epidemiologic or other long-term observational studies for comparison purposes. Where extrapolation becomes particularly challenging is when there are relatively few observed events (for example, few deaths in a short trial when modelling survival), or where participants have switched from a placebo to intervention arm (as is increasingly common in oncology studies) (16). In addition, there is increased extrapolation of the effects of disease modifying therapies, where the expected treatment effect from what may be a single intervention timepoint are extrapolated far beyond what is observed. When model calibration has been used to derive parameters, the "uncertainty around the calibrated values should be reported and reflected in deterministic or probabilistic sensitivity analyses or both"(25).

STRUCTURAL (MODEL) UNCERTAINTY

Structural uncertainty is uncertainty about the functional form of an economic model(37). If the structure of the model does not reflect what is happening in real life (for example if all relevant health states are not included) then the results of the economic model may not be reliable (even if all of the inputs are correct). Examples may include not enough health states (which would result in a lack of accuracy in estimates) or too many health states (which may not reflect reality and could result in a number of assumptions being required to inform each health state rather than evidence). The time horizon of the economic model is another aspect of model structure that can commonly generate additional uncertainty (with longer term time horizons also necessitating longer-term extrapolation of treatment effects and costs), and thus also increasing extrapolation uncertainty as described above.

Structural uncertainty is often not explored in depth, although it may have just as much impact, if not more, than parameter uncertainty(25). Recent approaches for characterizing structural uncertainty have sought to parameterize the structural uncertainties into the model. Adding parameters or varying elements of the structure can be undertaken and these are commonly presented as scenario analyses whereby different parameters or model assumptions are varied to represent different scenarios that may be possible within a

particular healthcare system or setting.

In a recent paper by Afzali et al. (38), five approaches to characterizing structural uncertainty were identified, including scenario analyses; model selection; model averaging; parametrization and discrepancy (see Table 2 for details of these approaches). Where complete rebuilds of the model are considered necessary then this is of course challenging. In these cases, the most appropriate course of action may be simply a qualitative summary, making the presence of uncertainty and possible impact on the findings as explicit as possible(25). In accordance with the Not-for-Profit GPF member survey, guidelines on addressing structural uncertainty are not typically provided by HTA agencies, though scenario analysis is the most common approach to characterizing uncertainty of this type. The Patient-Centered Outcomes Research Institute (PCORI) has stated that the principles for assessing structural uncertainty are undergoing review now.

Table 2 - details of approaches to characterizing structural uncertainty

Approach	Description
Scenario Analyses	Simplest approach in which structural uncertainty is handled, by presenting the predictions of alternative models based on plausible changes in structural assumptions
Model selection	The underlying concept is to select the best model (among a set of alternative options) on the basis of how well the model's outputs match the observed data for one or more model outputs
Model averaging	The predictive performance of alternative model structures is weighted against the data used to populate the model according to some measure of model adequacy
Parameterization	Some sources of structural uncertainty are expressed as uncertain parameters in the model (e.g., assumptions regarding patient and/or clinician behavior) and included in probabilistic sensitivity analyses
Discrepancy	The discrepancy between the model evaluated at its true inputs, and the true costs and true health effects, is calculated by decomposing the model into a series of sub-functions and considering the error at each sub-function

Affordability Uncertainty

Affordability is an increasingly important component for consideration by HTA processes, particularly in low- and middle-income countries (39), but increasingly also in high-income settings. Core aspects for evaluating the potential budget impact of a new therapy are the size of the target population and costs of the technology in practice (40). Uncertainty can arise in estimating the true population size (are all patients with the condition known, are there factors that would prevent or encourage patients to seek access to the new technology, will greater disease awareness increase the population size?). The costs of a new technology in practice can be challenging to estimate if the delivery is not clear (for example not knowing what healthcare setting and support will be needed), or if the place in the treatment pathway is unclear. Costs of the treatment can also include downstream cost offsets which can be difficult to quantify at the point that an HTA is being conducted, and can also lead to increased uncertainty. Typically, utilization uncertainty is greater than unit cost uncertainty. Finally, as the inputs to the budget impact analysis are often generated by the simulation model used to inform the cost effectiveness analysis, all of the possible clinical and economic uncertainties already described can also be present in these estimates as well.

Typically, guidance around calculating costs and population sizes are provided by HTA agencies, although some agencies have no specified methods. For example, in their survey response ICER stated that they let the end user decide what the level of uptake will be, and offer multiple price points to ascertain potential budget impact in a given setting. Specific examples that were provided in the survey response included submissions to the Pharmaceutical Benefits Advisory Committee (PBAC) in Australia, which are expected to include a refined spreadsheet template with details of cost and uptake. The CQHC in Russia stated that they provide guidance on what kind of costs should be considered and recommend sensitivity analysis for the size of the population.

Summary Approaches to Handling Economic Model Uncertainty

VALUE OF INFORMATION ANALYSES

Value of Information (VoI) analysis is a technical approach that provides a methodological framework that explicitly considers the level of input uncertainty, parameter uncertainty, and structural uncertainty in an HTA (41). VoI focuses on the likelihood of making a "wrong" decision if the technology is adopted, as such it can be used to understand what the cost of resolving residual uncertainty is. The Expected Value of Information (EVI) is the value of additional research and determines the extent to which further information will reduce the uncertainty(42). The intent is to allow a comparison of the potential benefits of further research with the costs of further investigation which provides an assessment of the value of investing limited healthcare resources in research or provision of the health technology (43).

Further steps and analysis methods can be used to determine the expected value of sample information (EVSI) and the expected value of (partial) perfect information (EV[P]PI). These pre-posterior forms of analysis aim to estimate the increased utility that a decision maker would have with access to an additional sample of information or the price that one would be willing to pay in order to gain access to perfect information. Essentially, these approaches are attempts to quantify the trade-off between making a potentially incorrect decision and generating more evidence (44). While there is much still to be done in terms of education, a recent ISPOR taskforce on VoI provides recommendations for good practice when planning, undertaking or reviewing the results of various VoI analyses(45). However, these methods are generally seen by many as academic exercises and to date have rarely been seen in practice by HTA agencies. It should be noted however that the use may increase in the future; NICE, for example, have highlighted that there is a "case to introduce wider use of the expected value of perfect information (or other equivalent concepts, such as payer uncertainty burden) as an additional tool to help committees understand the likelihood and consequences of decision error for parameter uncertainty" in their recent consultation review of the Methods Guide.

TOOLS FOR CATALOGUING MODEL UNCERTAINTY

There are a range of other summary approaches to characterizing uncertainty that have been reported in the literature; however, when comparing these with the approaches favored by HTA agencies, there is little overlap. Two more recent developments of note are the TRansparent Uncertainty ASsessmenT (TRUST) tool (46) and the TRUST4RD adaptation for rare disease (47).

The TRUST tool is a general tool that was developed to systematically identify, assess and report uncertainties in decision (economic) models with the aim to make uncertainties and their impact on cost effectiveness more explicit and transparent. In the validation of the TRUST tool (via HTA stakeholder interviews and application to six case studies) the authors state that stakeholders found it to be feasible and of value for transparent uncertainty assessment, but with the main barrier to use a lack of time to complete the necessary fields. Table 3 is a reproduction of the summarized TRUST approach, with identification of the sources of uncertainty conducted first followed by an assessment of the likely impact of the uncertainties on the cost effectiveness analysis.

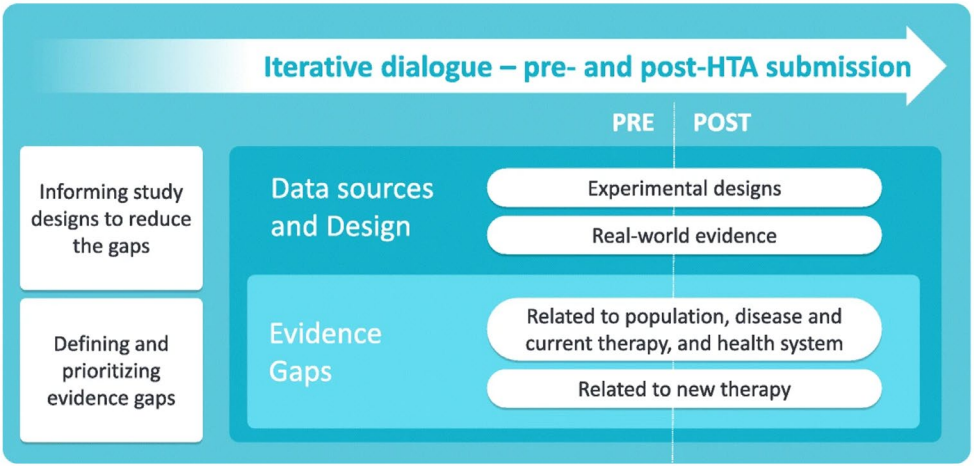
Table 3 – TRUST Table taken from Grimm SE, Pouwels X, Ramaekers BLT, Wijnen B, Knies S, Grutters J, et al. Development and Validation of the TRansparent Uncertainty ASsessmenT (TRUST) Tool for Assessing Uncertainties in Health Economic Decision Models. PharmacoEconomics. 2020;38(2):205-16.

			SOURCES OF UNCERTAINTY					IMPACT ON COST EFFECTIVENESS		
			Transparency	Methods	Imprecision	Bias & Indirectness	Unavailability	Uncertainty not reflected in PSA?	Uncertainty not explored in scenario analysis?	High impact on cost effectiveness?
ASPECTS		Context/scope								
		Model structure								
		Selection of evidence								
	INPUTS	Effectiveness								
		Relative Effectiveness								
		Adverse events								
		Utilities								
		Resource use & costs								
		Implementation								
		Outcomes								

Grey cells are unlikely combinations

TRUST4RD has been developed with Orphan Medicinal Products (OMP) in mind, as these technologies are associated with higher levels of input uncertainty (due to evidence from small or non-controlled trials, surrogate or immature outcome measures, and abbreviated follow-up, among other concerns) (48). As depicted in Figure 1, the TRUST4RD approach, developed through multi-stakeholder dialogue, aims to identify uncertainties of most concern for decision-makers by developing an iterative and informed dialogue so that potential approaches to uncertainty resolution can be discussed. The intended result is that future evidence generation can be more directed and will be more likely to demonstrate the value of a technology with less uncertainty than would have otherwise been presented.

Figure 1 - TRUST4RD components, from Annemans L, Makady A. TRUST4RD: tool for reducing uncertainties in the evidence generation for specialised treatments for rare diseases. Orphanet journal of rare diseases. 2020;15(1):127.



No reports of HTA agencies utilizing the TRUST tool or TRUST4RD approach have been received to date. However, more HTA agencies are engaging in early dialogue with stakeholders to better understand the probable levels of input uncertainty and to guide evidence generation plans accordingly. As highlighted by expert testimony, stakeholder input can be particularly useful in identifying where uncertainty will likely be a factor in advance of the HTA process itself. Where there are early scoping discussions stakeholder inputs can be sought at a stage that is often the most useful to shape the evidence generation plans and technology submissions.

WHO IS RESPONSIBLE FOR UNCERTAINTY?

In considering the many ways in which uncertainty may present in the evidence base for HTA, and the many ways in which it can be summarized and then evaluated, the notion of who is responsible for uncertainty is raised. Clearly the technology manufacturer must shoulder much of the burden of proof, as they are the owners of the technology and therefore drivers of the evidence base. Indeed, in the circumstances where traditional trial evidence is challenging to generate, there may still be alternative avenues of evidence generation (for example greater detail on burden of disease data, comparator usage) that may be possible to increase the committee comfort with a decision problem.

There are, however, considerations of responsibility for other stakeholders. For example, HTA agencies have a responsibility to ensure efficient use of taxpayer money in conducting HTAs that appropriately inform resource allocation, while understanding and taking societal (and possibly political) values into account. This means that they are responsible for helping to understand and mitigate the effects of uncertainty, as well as ensuring that the appraisal committees are sufficiently equipped to make the best decisions. As summarized above, there have been a range of technical and methodological advances in characterizing uncertainty for decision-making. One could argue in fact that the methods that are available to attempt to mitigate uncertainty are far in advance of what is typically undertaken and presented to appraisal committees. The advances have come with complaints; methods may not be perceived as intuitive and their complexities may be challenging for all members of an appraisal committee to understand. The notion that the technical adaptations for addressing uncertainty are a "black box" remains a continued criticism. The implication is that each committee member must develop technical abilities advanced enough to fully understand the methods and the subsequent results in order to integrate them into their decision making. This can become onerous and demanding for deliberators. Similarly, academic groups who often provide independent input or review of manufacturer submissions also have a responsibility to aid the HTA agencies in the understanding of the key drivers of uncertainty and their impact. Finally, other stakeholders such as patients and clinicians often benefit from adoption of technology by a health system. As such they are also responsible for helping to provide additional context and clarification of any input uncertainties as much as possible.

THE FUTURE OF INPUT UNCERTAINTY

Finally, as noted several times during the expert interviews, input uncertainty is "nothing new" for the field of medical devices and diagnostics. In this research field, evidence generation for these technologies often relies on single-arm trials (where placebo controls are not possible), small sample sizes, rapid technology evolution, and "learning curve" uncertainty (where the clinician becomes more proficient with experience) (49). Many feel that the uncertainties that device manufacturers and appraisers have been grappling with for years are now just becoming reality for those producing and appraising pharmaceutical technologies. However, as noted, with the advent of precision medicines (often with companion diagnostics), gene and stem cell therapies, and regenerative medicines, the line between drug and device is blurring and input uncertainty is increasing. Artificial intelligence and digital technologies are also bringing new complexities and uncertainties that must be considered. All of this is potentially compounded by new accelerated licensing pathways that are bringing technologies to licensure more quickly and with arguably less evidence (which, even when available is often presented as data in confidence) (21). Nearly 75% of respondents to the Not-for-Profit GPF member survey felt that input uncertainty is increasing in line with the issues outlined above.

Such a high degree of input uncertainty is leading to a need for more adaptable and nuanced tools or pathways for managing the uncertainty. Increasingly, there are examples where the uncertainty is genuinely unresolvable in a meaningful timeframe and HTA agencies are starting to give consideration to where it might be appropriate to manage such technologies with abbreviated or alternative HTA processes (Expert informant). Half of the responses to the Not-for-Profit GPF member survey suggested that HTA agencies are putting approaches in place to prepare for increasing uncertainty by moving to lifecycle approaches, with iterative appraisal processes, rapid reviews, and greater acceptance of real world and qualitative evidence. A specific example is that of multigene and multi-purpose investigative technologies as considered by the Medical Services Advisory Committee in Australia, where a concept is being introduced whereby "exemplars"

can form the basis of facilitating reviews of additional genes, purposes and/or medical conditions with less evidentiary burden. However, there is the risk that the additional evidence generated may indeed increase uncertainty if the results were not as expected (for example if the technology is poorly implemented or used in a cohort of patients with challenging comorbidities) and the conclusions may not be straightforward. Additionally, such approaches can be perceived as burdensome (particularly by technology manufacturers), potentially requiring perpetual evidence generation.

Throughput Uncertainty

The "throughput" stage describes how the various input uncertainties are handled; in other words, the weighting of the facts, values and reasons that will lead to a collective judgement. It is possible that new information can be presented by stakeholders (such as manufacturers, patients and clinicians); however, it is at this stage where the interplay between the clinical, economic and affordability uncertainties are considered, and consensus is sought, or votes are taken. Here the presentation of the uncertainty to an appraisal committee is critical, with a view toward facilitating consistent and transparent management of the key unknowns. As highlighted above, there can be multiple and varied levels and types of uncertainty present in any one HTA and a deliberative committee must quickly understand this and decide how the uncertainty impacts their deliberation. For example, NICE explicitly note in the current Methods Guide that they "will be more cautious about recommending a technology when they are less certain about the ICERs presented in the cost-effectiveness analysis" (50).

While uncertainty is inevitable in HTA, the notion that some uncertainties are unresolvable is key to consider. At what point is the level of uncertainty great enough that a decision must be deferred, or the technology cannot be recommended at all? The opportunity cost of resolving uncertainty (e.g., with a resource intensive evidence generation request) must be carefully considered, including the situations where the uncertainty is highly unlikely to be resolved in a meaningful timeframe. The question of whether uncertainty is unresolvable is becoming more common, and HTA agencies are starting to consider where it might be appropriate to manage such technologies with abbreviated or alternative HTA processes, which may even include specialized managed entry agreements and/or pricing arrangements to allow patient access while acknowledging some uncertainties may be unresolvable (Expert informant). Such approaches should not, however, come at the price of reducing quality evidence generation whenever this is truly possible.

In contemplating uncertainty in HTA, it is also important to also consider the notions of risk (and appetite for risk) as well as confidence. The level of uncertainty that is acceptable for an individual decision maker (including patients) is largely dependent on their appetite for risk. If a decision maker is more risk averse then they will require more certainty to make a positive recommendation, particularly where the consequences of a "wrong" decision are far-reaching (for example if an expensive technology does not realise the value it is expected to and healthcare resources are diverted from a more cost-effective existing standard of care). For a risk-taking decision maker (e.g., in populations with terminal prognoses), a higher level of uncertainty can be tolerated; greater risk tolerance is commonly observed in situations in which a technology is considered to be innovative, or has "plausible promise"(51) in an area with significant unmet need. These risk appetites may vary by individual when compared with the broader society in which the HTA agency operates, and this can create conflict and tension, where the societal values are unknown/unclear or become politicized. Confidence is a separate but critically important concept, as this is arguably the opposite of uncertainty in the context of decision making. Reducing uncertainty increases the level of confidence any decision maker will have in the evidence and their subsequent recommendations, irrespective of their appetite for risk.

It is the essentially the risks associated with uncertainty (i.e., making the wrong decision) that is what matters most to patients, clinicians, payers, healthcare, and even political systems, although whether the decision is wrong will certainly vary by stakeholder perspective. It is of particular concern if the perception that the decision is always wrong in one direction (i.e. always reimburse technologies that are not cost-effective). This however must be weighed with the risk of making no decision, whereby patients have no access to potentially beneficial technologies, or by the risk of generating more evidence to reduce the uncertainty (and

using scarce resources in doing so). Considering uncertainty in a realistic and pragmatic manner - but in a way that does not jeopardise high-quality evidence generation where it is possible - is a critical consideration.

Understanding Uncertainty in the Throughput Stage

The above techniques and approaches describe some ways in which the level of input uncertainty can be described and summarized to an appraisal committee. Qualitative, graphical, and otherwise technical ways of summarizing the uncertainty (including Bayesian analyses whereby prior expectations and information is applied to the analysis in a variety of ways) can all aid a committee in understanding the uncertainty. This is an area under current development; for example, the Australian PBAC is progressing work around consistency of terminology used in relation to its assessment of evidence, which includes describing the presence and impact of uncertainty.

(Non-manufacturer) Stakeholder Input

Stakeholders (particularly patients and clinicians) can be invaluable in helping to understand the level and impact of input uncertainty during the decision-making process. Input in the form of expert testimony and/or written submissions can help an appraisal committee understand the severity of a disease, the experience of comparator technologies, and the place of a new technology in a treatment pathway. Where sensitivity analyses have been presented, stakeholder input can be used to help the committee understand the most plausible and clinically relevant range of parameters (52). Inviting stakeholder input in this way has been criticized, as it is subject to bias due to conflicts of interest and is at best anecdotal in nature (53). Those who responded to the not-for-profit GPF member survey stated that most HTA agencies invite written submissions and/or invite stakeholders to attend committee meetings and/or accept public commentary. Other techniques for eliciting non-manufacturer stakeholder input include: discrete choice experiments; best-worst scaling; willingness to pay experiments; analytical hierarchy processes; focus groups; interviews; ethnographic fieldwork and deliberative inclusive methods (i.e. Delphi) (54). The Swedish Council on HTA (SBU) and Swedish Dental and Pharmaceutical Benefits Agency (TLV) in Sweden have a particular emphasis on unmet needs and human dignity for the prioritization of resources and so have set a precedence for using quality of life, patient willingness to pay, and other trade-off studies as supporting evidence for reimbursement decisions.

Other methods to gather and summarize stakeholder views on input uncertainty may be beneficial to reduce the impact of conflicts of interest and to establish a more holistic perspective (55). Examples of these methods are being trialed – for example the National Health Care Institute, Zorginstituut Nederland (ZIN) in the Netherlands have recently used an online Delphi model and public debate through social media to elicit a wider range of stakeholder views on the introduction of the Da Vinci Robot(56). The HTAi Patient and Citizen Interest Group (PCIG) are also developing a template for an international Summary of Information for Patients (SIP) (57). The SiP was originally a manufacturer-generated document that forms an essential part of the submission to the Scottish Medicines Consortium. The intent of the SiP is to provide a lay language summary of the technology being appraised to facilitate patient input into the HTA process. Development of a standardized international template will allow other HTA agencies to use a similar approach, facilitate the adoption by industry, and simplify the input for patient groups. Tools such as this can be used to help patient groups target where feedback and input is likely to be particularly useful for an appraisal committee to consider. It is feasible that such a tool could also be used by clinicians and other external stakeholders to understand the evidence base, where key uncertainties lie, and where input is likely to be most valuable.

Impact of Uncertainty on HTA Deliberations

Once the magnitude of the input uncertainty has been understood, decision makers must grapple with the uncertainty during the deliberative process and the impact of the uncertainty must be weighed by the appraisal committee. In addition to appetite for risk, each individual committee member must consider the available evidence and the uncertainty that is present and come to conclusions about how confident they are in the strength of the evidence and estimates presented. It must be highlighted however that, due to the very nature of HTA, the uncertainty can never be completely removed or mitigated. In the Not-for-Profit GPF member survey, the majority of responses stated that no explicit guidance was provided to committee

members on how uncertainty estimates should be handled during deliberations. Kaiser Permanente in the US noted that they do provide ratings of the quality of evidence and how this effects uncertainty regarding the estimate of effect, others (such as PBAC in Australia) stated that the guidance was implied rather than explicit; the primary aim is to ensure that there is consistency in how uncertainty is presented and considered, and that where there are apparent inconsistencies that these differences were defensible.

Following introduction of the summary of the evidence base (including the descriptions and measures of uncertainty previously described), the committee must then deliberate. It is at this point in the process that individuals can highlight areas of greatest concern alongside all other relevant matters. As highlighted by the responses to the Not-for-Profit GPF member survey, individual considerations on uncertainty are not recorded. Almost half of the respondents to the survey stated that consensus between members is sought with a collective committee view on the uncertainty necessary. This will then inform the recommendation of the committee and may include measures to reduce the uncertainty that is present, but may have a downside if variation in the perceived role that uncertainty plays is not documented.

Managing Throughput Uncertainty

Very generally speaking, the two main ways that input and decision uncertainty can be mitigated are through financial means (e.g., tying reimbursement to proof that the technology works as advertised), or through generation of additional evidence required to reduce uncertainty. The most common vehicles to achieve these goals are financial or outcome-based agreements (managed entry agreements [MEAs] and performance-based agreements), or what is known as "coverage with evidence development", in which a technology is listed and reimbursed for a temporary period while additional evidence is generated, typically triggering an updated appraisal at a later date. Different health systems and payer/provider frameworks have different capacities to allow access to technologies while attempting to reduce uncertainty. The willingness or ability of a payer or health system to enter into arrangements is therefore a large influential factor on how uncertainty can be managed (58).

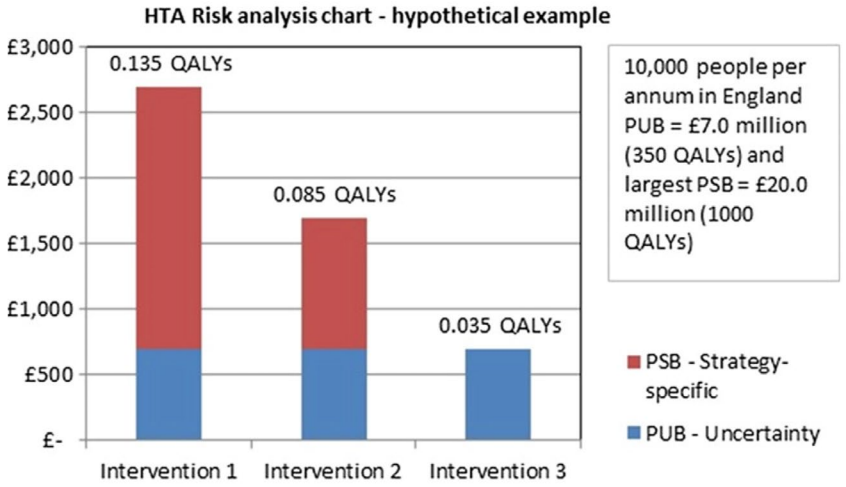
Managed entry and performance-based agreements, while typically financially-based (e.g., a lower price that is raised when performance meets a certain standard, or a higher price that is reduced when performance fails to meet that standard), may have variants. For example dose-capping schemes may be employed, (such as that for ranibizumab for macular degeneration agreed to by Novartis and NICE) where cost-effectiveness conclusions were based on a certain number of doses, after which the costs of treatment were borne solely by the technology manufacturer. A recent review by Vreman et al. investigated the feasibility of managed entry agreements for innovative therapies in different settings and combinations, allowing decision makers the flexibility to construct an agreement most suited to their preferences, to the technology and to their healthcare setting (59).

Concerns that have been raised around these managed entry and performance-based agreements are that the requirements for such schemes are generally labour and resource intensive – with the manufacturer voluntarily agreeing to the additional burden, but perhaps an undue burden also placed on health care systems, payers, patients, and clinicians. However, with adequate infrastructure the burden can be reduced over time. Additionally though, to be meaningful the schemes must be based on very clear requirements for the evidence generation (that is, the committee must be very clear on what additional uncertainty exists and what would be sufficient to change a decision). However, these schemes can only generate evidence for as long as they are in place and may only do little to reduce uncertainties around extrapolation of effects. Manufacturers also often feel disincentivised by such schemes if the price of the technology after additional evidence generation never goes up (it will more likely stay the same or be reduced and so prices can be higher at the outset to account for this). Conversely, HTA agencies and payers can experience a sense of frustration when the schemes are challenging to implement, and the expected effects are not experienced but the price remains the same. These types of situations can reduce trust between HTA agencies and technology manufacturers.

Some methods have been developed that attempt to quantify and visualize the need for approaches such as

MEAs; an example of this is the "HTA risk analysis chart" (60) as shown in Figure 2. This approach is set in the context of a decision problem with at least two technologies compared and where a technology is considered optimal if it provides the greatest "payoff" (as determined by the technology with the greatest net benefit). There are two main types of risk – the first is the risk that the decision is incorrect due to uncertainty in the evidence base. This risk is then visualized as the payer uncertainty burden (PUB) – this is equivalent to the EVPI and based on the results of the probabilistic sensitivity analyses. The second risk is the strategy-specific risk burden associated with choosing a specific non-optimal strategy and is termed payer strategy burden (PSB). For the strategy that is expected to be optimal, the PSB will be zero and for each "non-optimal" strategy the value of the PSB becomes positive. A graphical representation of the resulting HTA risk analysis chart (based on a hypothetical example) is below in Figure 2. Here the PUBs are the same height for each intervention and intervention 3 is the optimal technology (with zero PSB). If there is a large PSB then price-based managed entry agreements may be of greater value; if there is a large PUB then further evidence generation would be of greater value.

Figure 2 - Example of the HTA Risk Analysis Chart, from 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



Health technology assessment risk analysis chart illustrated in a hypothetical example. *PSB* payer strategy burden, *PUB* payer uncertainty burden, *QALY* quality-adjusted life-year

Other approaches use EVSI modelling to calculate what the risk of making conditional approvals might be (61). However, there is no evidence that such technical approaches have been widely used by HTA agencies as an aid to managing uncertainty.

Some HTA agencies can also defer recommendations, or recommend a technology is only used in the context of clinical trials, a form of coverage with evidence development known as "only in research" typically for higher risk interventions or "only with research" for lower risk interventions (62) . Both of these approaches are more risk averse, as the health system does not pay for the technology while the evidence is being generated. The primary concerns are that a broader set of patients may be denied timely access to a potentially beneficial technology, which is particularly challenging for end-of-life conditions or conditions that have no other viable treatment options. The use of only in research recommendations are a less common approach, (63) and some HTA agencies do not have this as a recommendation option at all.

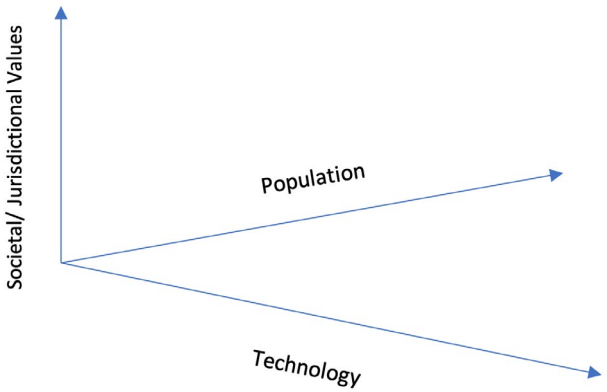
As mentioned in the Background section, however, any recommendation for managing the uncertainty must be pragmatic. For example, recommending significant additional evidence generation for a technology given for an ultra-rare disease or for a very targeted population is unlikely to be successful in its aims given the small number of patients throughout the world to include in additional clinical trials (although it is noted that the budget impact for ultra-rare diseases may be lower and therefore a higher tolerance of uncertainty may be acceptable). Even outside of these conditions, it is widely acknowledged that conditional reimbursement

approaches are a 'heavy lift' for all stakeholders and adopting such approaches in every situation is unsustainable. In the situations where uncertainty is unresolvable in a timely manner, then the level of uncertainty must be weighed against the risk to the HTA body as well as the payer and healthcare system which will have to implement any recommendation. There must be a genuine likelihood that the generation of new evidence will lead to a change in recommendation and the recommendations must be themselves ethical (64). Without these assurances, the trust between manufacturer and HTA agency will be reduced. Or to put it another way, as stated at the HTAi Asia Policy Forum on Managed Entry Agreements in 2018 (65), an MEA should not be used in lieu of a negative recommendation.

Context of Throughput Uncertainty

There are some contexts, however, where an HTA agency may have a higher threshold or tolerability for greater uncertainty (66). Examples include a condition that severely affects children (67), a rare or ultra-rare disease, (meaning that larger evidence base is more challenging and other treatments are less likely) (68), technologies that are given at the end of life as last-resort treatments, and those that may be potentially regenerative or curative (69). Many of these situations are subject to what is known as the "rule of rescue", describing situations involving an identifiable individual in immediate peril and the injunction to rescue that individual no matter the cost(70). While it is difficult to quantify exactly what additional weight society places on children, rare diseases, and end-of-life treatments, appraisal committees often try to consider some aspect of the societal perspective when weighting the impacts of uncertainty in these contexts (71). Circumstances where additional uncertainty might be tolerated can be categorized by two key domains: population (that is, rare/ultra-rare indications, children) and technology (for example Advanced Therapy Medicinal Products, ATMPs (72)), weighted against societal values and jurisdictional context as depicted in Figure 3. In the circumstances where HTA agencies do not routinely employ a societal perspective or some other "special-case" consideration, HTA agency recommendations may not completely reflect societal values. This can lead to politicization or judicialization of the HTA process, where HTA recommendations can be overturned within the legal setting (something that is becoming increasingly common in the Latin America region(73)).

Figure 3 - Contexts of Throughput Uncertainty



In the recent update to the NICE methods guide, it is explicitly stated that if these 'modifiers' (i.e. factors of that influence or change the considerations, decisions or recommendations) are applied then they should be applied cautiously and supported by "evidence of societal value or a robust and reasoned justification"(74). Contextual issues were also highlighted recently by a one-year review undertaken by ICER in the US, in collaboration with NICE and CADTH. In this review, ICER have set out criteria for eligibility for technologies to be deemed "single and short-term transformative therapies," and when given for ultra-rare diseases, specific methods can be applied. These include making cure proportion modelling as a standard reference case, presenting optimistic and conservative benefit scenarios, conducting threshold analyses for durability of treatment effect and discussion specifically around the uncertainties and controversies surrounding the technology. Annex 2 provides additional details from selected HTA agencies about their published methods for considering uncertainty.

Griffiths et al. (75), reviewed all HTA appraisals from 2000-2014 conducted by the National Institute for Health and Care Excellence (NICE, England), Scottish Medicines Consortium (SMC, Scotland), Pharmaceutical Benefits Advisory Committee (PBAC, Australia), and Canadian Agency for Drugs and Technologies in Health (CADTH, Canada). All instances in which the cost-effectiveness result was higher than the accepted threshold were analysed. The results showed that NICE recommended the highest proportion of submissions, with results higher than the threshold (34% accepted without restrictions; 20% with restrictions), followed by PBAC (16% accepted without restrictions; 4% with restrictions), SMC (11% accepted without restrictions; 14% accepted with restrictions), and CADTH (0% recommended without restrictions; 26% with restrictions). Most of the submissions analysed were classified as cancer and autoimmune diseases. Reasons for accepting submissions reporting cost-effectiveness findings above the threshold included substantial clinical benefit over the standard of care and addressing an unmet therapeutic need. In response to the survey of Not-for-Profit GPF members, most agencies stated that there were no explicit guidelines for when there should be special consideration of uncertainty, and that the situations are implied rather than explicit. It was also highlighted that having specified considerations (such as orphan or end-of-life conditions) may be unhelpful as they can generate artificial boundaries, and more case-by-case contextual consideration may in fact be more appropriate.

Consistency and Predictability

As described, there are many types of input uncertainty that need to be summarized and presented for consideration. Each individual committee member must understand and then weight the impact of uncertainty on their decision-making. However, it is very difficult to codify how uncertainty has been weighted by individual committee members. While a formulaic approach may be desired by technology manufacturers, this is difficult as it rarely fully captures all the elements of human nature that are dynamically at play and some HTA committees do not accept such approaches. Multi-criteria decision analysis (MCDA) (76) is a possible tool to address this; however, some feel this itself adds an additional layer of subjectivity (or "fake objectivity") in its implementation(77). Fundamentally, there are times when attitudes to risk, tolerance of uncertainty and confidence in what is being presented are difficult to capture neatly, and their variation may be simply a function of the informed but nonetheless human judgment that is at the heart of HTA. Consistency in how uncertainty is handled and considered within and between committees at a single organization through to how it is handled across jurisdictions is clearly an area of great concern for technology manufacturers, but difficult to synthesize given the different societal values at play. Nevertheless, it is critically important for technology manufacturers to see consistency in how uncertainty is handled generally in similar situations (e.g., within the same disease or across similar technologies). It is the predictability of how uncertainties are prioritized, valued and managed that becomes the key concern for manufacturers and other stakeholders.

In a review by Allen et al., (66) that compared the recommendations on common HTAs between Australia, Canada, England and Scotland found that between 2009 and 2013 there were 26 technologies that had been assessed by all 4 agencies for the same indications. Of these, a detailed review was conducted on 7 of the most divergent case studies. The authors identified that the HTA activities varied because of different mandates and unique political, social and population needs, but that the differences in recommendations could primarily be attributed to uncertainties surrounding cost-effectiveness, clinical benefit, safety, trial design and submission timing, with comparator choice also a key reason for differences in recommendations.

Stakeholder Input

As mentioned, the primary role of stakeholders such as patients and clinicians is to try and clarify any input uncertainty (both before and after technical approaches have been conducted) by providing lived experiences and contributions to the real world experience (78). Often, however, particular input uncertainties become more relevant during the committee deliberations (for example during deliberation on manufacturer and/or academic submissions and reports). In these circumstances, the use of stakeholders to reduce uncertainty can become inconsistent and can vary according to how prepared stakeholders are in advance of any deliberations. In a recent survey by the European Patients Forum, the majority of the 18 decision makers

that responded stated that patient input had "a high impact on increasing the accuracy in measuring the needs and preferences of patients and led to a better understanding of the impact of technologies in a real-life context"(79). The opportunity cost of stakeholder involvement in helping to consider uncertainty must also be considered; patient organizations take time and invest effort in producing responses to submissions and recommendations. There also needs to be a realistic expectation of what uncertainty can be reduced by stakeholders and it must not be a tokenistic or academic exercise (80). Stakeholder involvement must remain relevant and meaningful to maintain trusting relationships between stakeholders and HTA agencies.

Case Studies

In order to more easily conceptualize how HTA agencies across the world consider uncertainty in the same evidence base, three case studies are described (see Annex 3 for further details). Case Study 1 (CAR T therapy) describes the case of a novel, single-administration immunological treatment, considered to represent a "breakthrough" in the treatment of lymphomas and leukemias affecting children and older adults. The clinical evidence was primarily comprised of short, single-arm, small studies and the uncertainty in the evidence base was acknowledged by all HTA agencies. Most agencies reviewed in this case study did eventually recommend the CAR-T in question for the cancers of interest, along with the implementation of financial mechanisms such as performance-based payment schemes.

Case study 2 (treatments for Spinal Muscular Atrophy, SMA) is similar to that of one of the CAR T therapies, in that the condition affects young children and one of the treatments considered (Zolgensma™) is a "one-shot" intervention. The list price for Zolgensma was the highest in the world at the time of introduction ($2.125 million USD). The other treatment considered (Spinraza®) was also an innovative intervention but was delivered in multiple doses with an initial treatment and maintenance doses that could ultimately cost $5 million USD per patient over time. The clinical evidence base for both treatments has been considered as relatively robust (in the context of treatments for an ultra-rare condition) with randomized controlled trials, open label dose escalation and extension studies. Concerns that were highlighted, however included that the control and intervention groups were imbalanced in terms of important prognostic factors, the study populations were not fully reflective of the real world population and the studies are all relatively short and lengthy extrapolation of treatment effects was required to estimate the full benefit of treatment. While appraisals of Zolgensma are awaited in some jurisdictions, Spinraza has been recommended for use by most jurisdictions, but typically with strict patient selection and with the implementation of various risk mitigation strategies in the form of stopping rules and data collection requirements. Zolgensma has also been provided in a number of jurisdictions with schemes such as the ATU Framework in France whereby exceptional use of drugs that do not have a marketing authorisation that treat serious or rare diseases with no other appropriate alternative treatments can be granted.

Case study 3 highlights the use of transvaginal implants for the surgical repair of pelvic organ prolapse and stress urinary incontinence. This older technology was granted marketing authorisation without a well-developed clinical evidence base. Following safety reports and numerous lawsuits filed by patients who have experienced significant adverse effects, there are now a variety of recommendations made by regulatory and HTA agencies throughout the world ranging from not recommended for any use to recommended for use in certain patient subgroups, and in some countries (for example in South America) there are no restrictions on the recommendations for use. The uncertainty in the evidence base has demonstrated the importance of patient education and counselling and the need for surgeon credentialling. Studies with longer follow up are still underway.

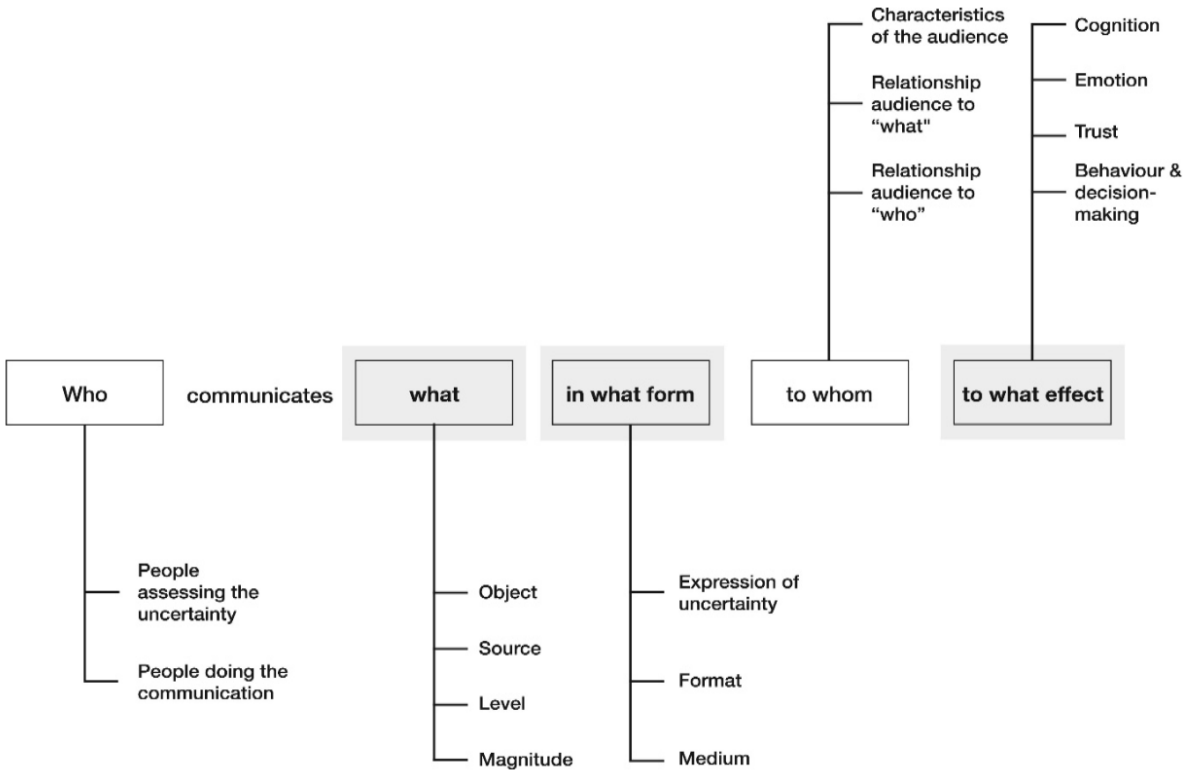
Output Uncertainty

Output uncertainty refers to the way in which the level of uncertainty and its impact on the recommendations is communicated and any learning is consolidated. In many fields the communication of uncertainty (what uncertainty is present and the impact that uncertainty had on the decision) is "[shied] away from"(7). Most often, the complexity of the situation may make the understanding challenging and prohibits effective communication (81). Additionally, where there is a lack of consistency in how uncertainty is considered this makes communicating uncertainty in a coherent and logical manner all the more challenging. It is also important to acknowledge the inevitability of the presence of uncertainty; "new scientific knowledge usually includes greater uncertainty and researchers often don't know how much of the picture it shows"(82). The challenge in communication is recognizing the uncertainty that exists while ensuring there is public trust in what is certain about the data.

A Framework for Communicating Epistemic Uncertainty

In a recent paper by van der Bles et al (7), a possible framework for how to communicate epistemic uncertainty (that is, the known unknowns) is presented. This was based on Lasswell's venerable model of communication (83) and essentially addresses who communicates what, in what form, to whom, and to what effect while acknowledging the relevant contexts (see Figure 4). This could be a useful framework for HTA agencies to consider using while developing outputs and communications on the presence and impact of uncertainty.

Figure 4 - Framework for Communicating Epistemic Uncertainty



Broadly speaking, in the above framework and within the context of communicating uncertainty in HTA, the following applies:

- 'Who' are the technology appraisal committee who assess the uncertainty and the HTA agency who communicates the impact of the uncertainty on the committee deliberations and decisions or recommendations
- 'What' is being communicated is the technology as the article of uncertainty (in terms of facts, numbers or models); the source (or reasons) for the uncertainty; the level of the uncertainty (whether direct or

indirect); and the magnitude of the uncertainty (e.g., 3 well-conducted trials with small variations in treatment effect vs. 1 small single-arm study with only short-term data). The context of the uncertainty (as described previously) such as whether the disease affects children or is a technology for a rare condition, is also important to communicate

- 'Form' is how uncertainty is expressed (in full, or briefly); the format of the communication (numbers, visualizations, and/or verbal statements); the medium of the communication (online, broadcast, verbal, print)
- 'Whom' is about the characteristics of the stakeholder audience (levels of numeracy and literacy); knowledge of the field; the relationship of the topic to the stakeholders (how emotionally invested they are); the relationship of the HTA agency to the stakeholder (how much the stakeholders trust the HTA agency and perceive them to be credible)
- 'What effect' is considering the effect that the communication will have on the stakeholders cognitively and emotionally, the trust that stakeholders place in the communication, and the impact on behaviour and decision-making.

Communicating to Stakeholders

Some additional perspectives and recent examples relevant to HTA outputs are considered below. As previously noted, there is an increasing amount of data that is presented in confidence (if this is accepted by the HTA agency). Describing the presence and impact of the uncertainty if it is to be kept confidential is clearly challenging, however there are attempts to rectify this. For example, beginning at the end of 2020, and unless strict criteria are met, PBAC in Australia is committed to publishing all clinical data relied on for the decision-making, regardless of how obtained.

PATIENTS

When communicating the impact of uncertainty to personally invested stakeholders such as patients, it is important to acknowledge the differences between population and individual uncertainty. Decisions and recommendations are typically made on population level data with societal values and perspectives, but they affect individual patients. Understanding (and accepting) uncertainty that exists at the population level is challenging when decisions that have been made may negatively impact the individual patient's life. This is particularly true where a patient has access to no other viable treatment options (particularly so for ultra-rare, orphan conditions) or if they have personal positive experiences with the technology (for example in the clinical trial setting). This can be compounded by an inherent sense of unfairness if a technology is approved for one patient population but not another.

There are examples demonstrating attempts to improve communication of uncertainty to the public and to individual patients. Graphical displays such as traffic light systems are becoming more common as a way to quickly display level of confidence in the evidence or finding (or to conversely indicate how much input uncertainty is present). Typically input uncertainty can be rated as red, amber or green (high, moderate, low) which can give a high-level overview of how uncertain the evidence base is. Icon arrays are also often used to assist stakeholders to understand probabilities (for example probability of diagnosing cancer based on different sensitivities and specificities associated with diagnostic tests). However, as with GRADE, similar criticisms of graphical representations of uncertainty apply as the ratings are subjective, can be inconsistently applied and also do not necessarily provide detail on what the impact of the uncertainty might be. The survey of Not-for-Profit GPF members reflected this discomfort, as half of the agencies that responded do not make any specific attempts to explain the uncertainties to patients and the public. It was acknowledged as an area for improvement within the survey by some. For those Not-for-Profit GPF members who responded to this question positively in the survey, most provide lay language summaries for patients and the public that integrate uncertainty concerns.

Many HTA agencies around the world now also provide lay language summaries alongside other publicly available summaries of the appraisal committee deliberations (such as the Public Summary Document by PBAC in Australia). One organization that has invested heavily in developing lay summaries is the

Patient-Centered Outcomes Research Institute (PCORI) in the United States. PCORI have established the Patient-Centered Outcomes Research Translation Center (84) to make sure that PCORI research findings are comprehensible and useful to patients, clinicians and others in making healthcare decisions. This is particularly important given the move towards more shared decision making in the US and worldwide. At the PCORI Translation Center, health communication and literacy experts prepare a summary of the findings for the general public using standard and consistent formats, which are then reviewed and tested by a Technical Expert Panel. Coupled with a commitment to open access journal articles, this is one of the biggest investments in communicating uncertainty to patients recognized to date.

Providing constructive feedback on how stakeholder input did (or did not) help resolve uncertainty is also a key component of the process; this ensures that contributions are appropriately acknowledged and continually improved. The provision of feedback should also ideally take the form of a two-way dialogue rather than a one-way monologue, although the former is clearly more time-consuming and resource intensive. CADTH is undertaking a consultation (February 2021) to ascertain how patients can be better informed about the uncertainties in the evidence base and how they can best contribute to a recommendation. In addition to this, CADTH also have an Implementation Support and Knowledge Mobilization Team that embedded in the CADTH jurisdictions that provides services and knowledge tools to CADTH customers and stakeholders about the reports and recommendations. The Team also provide training and capacity building workshops to the policy counterparts in Canada, alongside working with policy makers, funders and clinicians to identify priorities to help CADTH define some of their activities.

TECHNOLOGY MANUFACTURERS

The impact and consideration of uncertainty is critical for a technology manufacturer to understand. In the context of pricing negotiations, or where there are recommendations for additional evidence generation, the manufacturer needs to be very clear on what the uncertainty was, what it meant and what is now required of them (58). Ensuring that the uncertainty that impacted the decision is clearly articulated will facilitate the generation of evidence that will hopefully aid in the reduction of uncertainty. Current mechanisms for facilitating this are detailed summary reports and minutes of appraisal committee meetings. In the case of developing patient access schemes or negotiating prices, dialogue between manufacturer and HTA agency and/or payer can also be undertaken.

Where there is ineffective communication of uncertainty and its impact, this can result in a labor and resource-intensive appeal against the recommendation by the technology manufacturer (in jurisdictions where this mechanism is available). However, as noted previously, much of the uncertainty that HTA agencies must contend with comes from the clinical evidence base and so is introduced by decisions on what is measured, in whom and for how long. Early scientific dialogue and advice can go some way to mitigate this, as a mechanism to address what the likely considerations of the uncertainty will be at an HTA committee level, which can then inform the clinical trial plans accordingly.

HEALTH SYSTEM STAKEHOLDERS

Other health system stakeholders (such as clinicians, policy makers and payers) also need to understand the level of input uncertainty and its impact on the recommendation. These are the stakeholders who typically must implement the recommendation made by an appraisal committee and as such a sound understanding is essential. As mentioned previously, this is particularly true where mechanisms such as managed entry or outcome-based agreements have been employed to manage uncertainty. Where there is ambiguity about what something means, or a stakeholder does not know why something has been recommended then this can also create general feelings of doubt or unease, resulting in reduced trust and confidence in the HTA process. In response to the survey of Not-for-Profit GPF members, respondents stated that uncertainty is explained to other health system stakeholders primarily through publicly available summary documents, and that no other special vehicles are utilized; it appears that there is room for further innovation.

Impact of the COVID-19 Pandemic

It is impossible to develop a background paper for a HTAi Global Policy Forum without mention of the impact of the current COVID-19 pandemic. In regard to the current topic, the impact of the pandemic is potentially far reaching, with many of the true effects still yet to be realized.

For example, regulators have responded to the COVID-19 pandemic with the increasing use of emergency regulatory pathways (to assess new or existing technologies in the context of treating COVID-19), with the FDA approving remdesivir as the first licensed treatment for COVID-19 in October 2020 after an initial emergency authorization in May (85). The ramifications of this are still only just being understood, but it is likely that the levels of input uncertainty around these rapidly assessed technologies will increase. For example, remdesivir was priced based on the assumption that the final clinical trial results would show a clear mortality benefit, in line with an early HTA extrapolation, however this was adjusted as more data became available (86). In this case, the mortality benefit was not demonstrated, and because no formal HTA appraisal was done, there is no mechanism to modify the price. Indeed, the role of HTA agencies in assessing additional treatments or vaccines is unclear, and it is possible that the greater levels of uncertainty accepted for COVID-19 interventions may become more acceptable when considering interventions for other conditions.

As already mentioned, accelerated regulatory pathways are already in full force aside from COVID-19. Some argue that these processes are simply removing redundancies and streamlining; but by definition this means that HTA agencies will receive less data and with shorter follow-up (cited as an area of great concern for HTA agencies)(87). In addition, the increased flexibility and agility being shown by regulators is creating a greater expectation for HTA agencies to do the same (Expert informant). The greatest risk to HTA is of course that in the rush to approve new technologies that HTA will be seen as superfluous entirely.

Other major impacts of the COVID-19 pandemic include changes to the throughput format and processes of HTAs. Committee meetings that were face-to-face are now virtual. This will impact the capacity and ability for stakeholders to be involved and truly be heard. The committee deliberations will be likely affected, but the dynamics of these changes are yet to be fully understood. This is similarly true for the opportunities for information sharing, which are now virtual or through online media. Relationship building and development of trust can be more challenging in this way, but conversely it may open more opportunities for those who are typically unable to attend public meetings, potentially leading to a more inclusive approach. Values may also change, and it is possible that technologies that are innovative and stimulate the economy may be valued higher than before the pandemic.

The COVID-19 pandemic has also certainly increased the public's understanding of medical research and a greater appreciation of the risk associated with it(88). The effect of the increase in understanding of medical research on the public's perception of HTA is, as yet, largely untested. However, alongside the increased appreciation and understanding of medical research, the pandemic has also resulted in what the World Health Organization (WHO) has termed an "infodemic" of misinformation. A recent study by Evanega et al (89) identified over 1.1 million news articles that disseminated, amplified or reported on misinformation related to the pandemic. The WHO has also released "Communicating and Managing Uncertainty in the COVID-19 Pandemic: A Quick Guide" (ref) as an aid for those discussing and presenting information on the pandemic.

Finally, the economic impacts of the COVID-19 pandemic are also being felt across the world, and the effects of suspending national economies to stop the spread of the disease will have consequences for many years to come. One of these is likely to be on the financial sustainability of health care systems, with possible reductions in overall future spending. The reality for HTA agencies in this "new normal" could mean that HTA is seen as more crucial than ever in prioritising how to spend even more scarce healthcare resources, or it could mean it is a step that becomes bypassed altogether. HTA's ability to evolve and improve in considering and communicating the inherent uncertainty in medical research should not be overlooked. While there remain a number of policy questions for discussion at the HTAi GPF, it is clear that there is a wealth of experience within the HTA community that seeks to ensure patients have timely access to right technologies while balancing the use of precious healthcare resources. The challenge will be to maintain this approach in an increasingly complex landscape for evidence and decision-making.

Acknowledgements

The Policy Forum Chair and Scientific Secretary would like to thank the following expert informants for speaking at length with them about Considering and Communicating Uncertainty in HTA. The information and insights they provided were an important contribution to the background paper and helped to stimulate thinking about the meeting program. Karen Facey (University of Edinburgh, UK); Ann Single (HTAi PCIG Chair); Mark McIntyre (Boston Scientific); Mohit Jain (Biomarin, UK); Andrew Wilson (Pharmaceutical Benefits Advisory Committee (PBAC), Australia); Robyn Ward (Medical Services Advisory Committee, MSAC, Australia); Andrew Mitchell (Department of Health, Australia); Wija Oortwijn (Radboud University Medical Centre, Netherlands); Wim Goettsch (ZIN, Netherlands); Meindert Boysen and Helen Knight (National Institute for Health and Care Excellence (NICE), UK); Allan Wailoo (University of Sheffield, UK); Jean Slutsky (PCORI, USA); Laura Sampietro-Colom (Hospital Clinic of Barcelona, Spain); Andres Pichon-Riviere (IECS, Argentina); Ken Bond (Institute for Health Economics (IHE), Canada); Tina Wang and Neil McAuslane (Centre for Innovation in Regulatory Science, CIRS, UK), Nicole Mittman (Canadian Agency for Drugs and Technology in Health, CADTH, Canada), Nuriya Musina (Center for Healthcare Quality and Control, Russia) and Stefan Sauerland (Institute for Quality and Efficiency in Healthcare, IQWiG, Germany).

Annex 1 – Examples of Evidence Rating Mechanisms

Scales from the Agency for Healthcare Quality and Control, Russia

The scales below were provided by the Russian Agency for Healthcare Quality and Control. Clinical trials of pharmaceuticals are assessed using these scales.

ASSESSMENT SCALE OF LEVELS OF EVIDENCE

Trial characteristics	Level of evidence of results	Assessment scale (score)
Systematic reviews and meta-analysis	I	10
Blind randomized clinical trials	II	9
Open randomized clinical trials	II	8
Network meta-analysis (incl. indirect and mixed comparisons)	III	7
Cohort studies	IV	6
Case-control study	V	5
Case report and case series	VI	4
Experts views	VII	3

ASSESSMENT SCALE OF LEVELS OF STRENGTH OF EVIDENCE

Level of strength	Level of evidence of results	Assessment scale (grade)
A	Evidence is convincing: strong evidence for the proposed statement is provided	3
B	Relative strength of evidence: there is enough evidence to recommend a proposed drug for inclusion (exclusion) in (from) a reimbursement list	2
C	Insufficient evidence: there is not enough evidence to recommend a drug, but it could be recommended under other circumstances	1

INTEGRATED QUANTITATIVE QUALITY ASSESSMENT

Trial characteristics	Level of evidence	Level of strength	Overall score
Systematic review of randomized clinical trials and meta-analysis with a low or moderate risk of bias	I	A	Is a product of an assessment (score) of level of evidence and assessment (score) and level of strength
Randomized clinical trials with a low or moderate risk of bias	II	A	
Network meta-analysis (incl. indirect and mixed comparisons) with a low or moderate risk of bias	III	A	
Cohort studies with a low or moderate risk of bias	IV	B	
Case-control study with a low or moderate risk of bias	V	B	
Case report and case series	VI	C	
Experts views	VII	C	
Studies of any design with a high risk of bias (of low methodological quality)	I II III IV V VI VII	C C C C C C C	

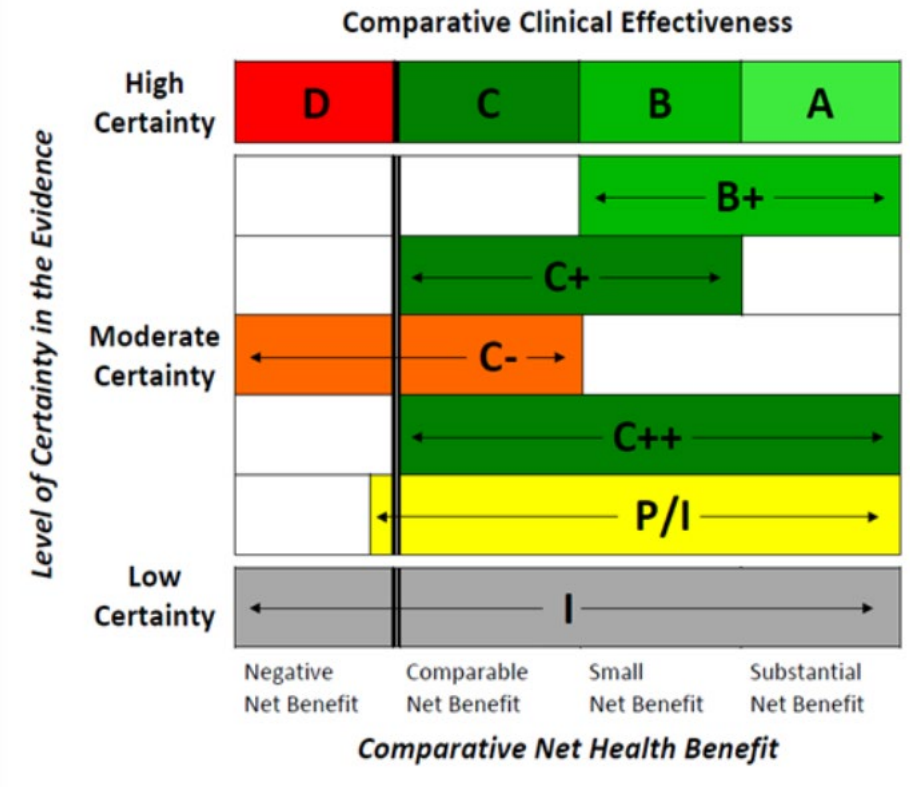
Overall integrated quantitative quality assessment of clinical trials of drugs:

Score is over 18 – enough to recommend a drug for inclusion into a reimbursement list

Score is over 12 – enough to recommend an orphan drug for inclusion into a reimbursement list

ICER – EVIDENCE RATING MATRIX ™

The ICER Integrated Evidence Rating ™ combines a rating for comparative clinical effectiveness and a rating for comparative value. The clinical effectiveness rating arises from a joint judgement of the level of confidence provided by the body of evidence and magnitude of the net health benefit – the overall balance between benefits and harms. This method for rating the clinical effectiveness is modelled on the "Evidence-Based Medicine (EBM) matrix" developed by a multi-stakeholder group convened by America's Health Insurance Plans.



- A = "Superior" – High certainty of a substantial (moderate-large) net health benefit
- B = "Incremental" – High certainty of a small net health benefit
- C = "Comparable" – High certainty of a comparable net health benefit
- B+ = "Incremental or Better" – Moderate certainty of a small or substantial net health benefit, with high certainty of at least a small net health benefit
- C+ = "Comparable or Incremental" – Moderate certainty of a comparable or small net health benefit, with high certainty of at least a comparable net health benefit
- C- = "Comparable or Inferior" – Moderate certainty that the net health benefit is either comparable or inferior, with high certainty of at best a comparable net health benefit
- C++ = "Comparable or Better" – Moderate certainty of a comparable, small, or substantial net health benefit, with high certainty of at least a comparable net health benefit
- P/I = "Promising by Inconclusive" – Moderate certainty of a small or substantial net health benefit, small likelihood of a negative net health benefit
- I = "Insufficient" – Any situation in which the level of certainty in the evidence is low

Annex 2 – Examples of HTA Agency Methods regarding Uncertainty

This Annex contains descriptions regarding uncertainty for five HTA organizations: NICE, CADTH, ICER, IQWiG, HAS. Information on each was compiled from publicly available English-language documents from respective websites. The descriptions are not comprehensive and intended only to give HTAi Global Policy Forum meeting participants a sense of the variation in how the aspects explored in the background paper have been implemented in well-established HTA agencies.

National Institute for Health and Care Excellence (NICE), England

Summary of key aspects of Current Methods Guide (2013)

INPUT

Sensitivity and scenario analyses (presented separately to the reference case) to be explored. Inputs must be fully justified and uncertainty explored using alternative input values. If the choice of data sources is not clear-cut, the analysis should be re-run using the alternative data source.

Probabilistic sensitivity analyses are preferred (with distributions chosen to represent the available evidence and choice should be justified). If there are alternative plausible distributions, then separate probabilistic analyses of these scenarios should be conducted. Assumptions should be clearly presented and limitations specified.

Regarding structural uncertainty, assumptions should be clearly documented and the evidence and rationale to support them provided. The impact of structural uncertainty on estimates of cost effectiveness should be explored by separate analyses of a representative range of plausible scenarios.

Confidence ellipses and scatter plots on the cost effectiveness plane and cost-effectiveness acceptability curves (including representation and explanation of the cost-effectiveness acceptability frontier). Uncertainty should also be presented in tabular form. Univariate and best- or worst-case sensitivity analysis are useful, but the use of PSA can allow a more comprehensive characterisation of the parameter uncertainty associated with all input parameters.

THROUGHPUT

Uncertainty is a key factor underpinning the judgements of the Committee. The Appraisal Committee is likely to consider more favourably technologies for which evidence on cost effectiveness is underpinned by the best-quality clinical data than those for which supporting evidence is dependant to a large extent on theoretical modelling alone. The Committee is aware that the evidence base will necessarily be weaker for some technologies, such as technologies used to treat patients with very rare diseases.

Above a most plausible ICER of £20,000 per QALY gained, judgements about the acceptability of the technology will specifically take into account the following factors:

- The degree of certainty around the ICER. In particular, the Committee will be more cautious about recommending a technology when they are less certain about the ICERs presented.
- Whether there are strong reasons to indicate that the health-related quality of life has been inadequately captured
- The innovative nature of the technology
- Whether the technology meets the criteria for special consideration as a 'life-extending treatment at the end of life'
- Aspects that relate to non-health objectives of the NHS.

In addition, the Committee will want to be increasingly certain of the cost effectiveness of a technology as the impact of the adoption of the technology on NHS resources increases. The likelihood of

decision error and its consequences are also likely to influence the Committee's judgements on cost effectiveness.

Recommendations can include:

- Only in research
- Data collection within a MEA
- Recommend for use in specific subgroups only
- Research recommendations aligned with NETSCC

OUTPUT

Final appraisal documents are posted on NICE website.

NICE provides explicit explanation of the principles guiding the recommendation documents, including the key issues that have been debated and the rationale for the committee's conclusions. A lay summary is also provided.

CURRENT CONSULTATION (2020/21)

The NICE Methods Guide is being updated and a consultation is currently underway (as at January 2021), with an expected due date for publication of the new Methods Guide by September 2021.

The case for change has been disseminated for consultation, key aspects related to uncertainty within this document are as follows:

- Uncertainty should remain an important consideration in decision making. A greater degree of uncertainty and risk should be accepted in defined circumstances, including:
  - conditions for which it is recognised that evidence generation is complex and difficult, such as rare diseases
  - innovative technologies
  - technologies that provide large benefits
  - when the uncertainty and risks can be monitored and controlled, such as in a managed access arrangement..
- The methods guidance on sourcing, assessing and presenting evidence should be refreshed and clarified, including:
  - no change to the general preference for randomised controlled trials (RCTs), when feasible, to inform estimates of treatment effects
  - an emphasis on the role of a comprehensive evidence base, including non-RCTs and real-world evidence, and the circumstances in which different types of evidence have strengths or limitations
  - additional guidance on the use of RCT and non-RCT evidence, assessment and reporting of study quality, risk of bias and confounding, and presenting evidence.
- Evaluations should include an overall assessment of uncertainty, including the effects of different types of uncertainty, whether uncertainties have been captured in analyses, and whether uncertainties can be addressed by additional evidence.
- Approaches to presenting and visualising uncertainty should continue to be developed.
- Probabilistic analyses should be used as the starting point for economic analyses, including committees' preferred analyses and scenarios, unless there is clear justification not to.
- Proposals about conducting probabilistic analyses (Monte Carlo error, ordering/correlation in parameters) and other sensitivity and scenario analyses (probabilistic univariate analyses, threshold analyses, parametrizing structural uncertainty, EVPI).
- Reviewing key challenges for technologies such as ATMPs, histology-independent cancer

treatments and other emerging technologies allowed us to identify generally applicable methods improvements that are particularly helpful for these challenging topics. We also explored the methods challenges for rare diseases in detail.

- Medical technologies evaluations should now consider, when relevant, unpublished evidence and post-marketing surveillance data.
- Evaluations in which there is uncertainty about long-term health benefits should include scenario analyses that explore the effects of different assumptions about long-term benefits. This might include threshold analysis for the duration of treatment effects.
- Cure-proportion modelling should be considered as an option.

Canadian Agency for Drugs and Technology in Health (CADTH), Canada

Summary of key aspects of Current Procedures and Guidelines

INPUT

Factors that contribute to clinical uncertainty (specifically in the context of rare diseases; https://www.cadth.ca/sites/default/files/pdf/es0326_drugs_for_rare_diseases_appendices.pdf):

- Limited number of clinical studies
- Small sample sizes (e.g., due to rare disease that affects a relatively small number of patients with an incidence of fewer than 5 in 10,000, but typically closer to 1 in 100,000)
- Absence of comparator groups
- Alternative or adaptive trial designs for rare diseases
- Short study durations or follow-up
- Inability to distinguish disease severity in heterogeneous manifested rare diseases • Limited to surrogate end points
- Insufficient evidence on meaningful clinical end points
- Greater uncertainty in statistical analyses

Economic uncertainty – the results of the reference case should be compared with those from a non-reference case analysis, with the impact of departing from the reference case explicitly stated. Non-reference case analyses should still include probabilistic analysis.

Parameter uncertainty should also be addressed using probabilistic analysis taking into account correlation between parameters and consideration (and likelihood) of the critical values that could alter a decision. Deterministic analyses (one-way, multi-way or threshold) are not recommended. The impact of changes in deterministic parameters (e.g. prices) should be assessed using scenario analyses.

Decision uncertainty should be summarized using CEACs, although scatter plots on the cost-effectiveness plane are not part of the reference case and confidence ellipses are not recommended. When the decision problem includes an MEA or OBA (i.e. consideration of further research to inform future decisions), a value-of-information analysis should be presented as part of the reference case with two-stage expected value of perfect parameter information analysis useful to identify all of the correct parameters. One-way probabilistic analysis can also be presented (outside of the reference case). Value-of-sample information can also support decision makers assessments on the return on investment of further research.

Structural uncertainty should be addressed through scenario analyses, but researchers should identify a recommended structural approach for the reference case.

THROUGHPUT

The CADTH drug expert committees may specify that a recommendation in favour of reimbursement is

contingent upon one or more conditions being satisfied. These conditions commonly include initiation criteria, outcomes required, renewal criteria, discontinuation criteria, prescribing criteria, and conditions related to the price of the drug.

Recommendations can include:

- Reimburse
- Reimburse with conditions:
  - Subgroups
  - Listed in a similar manner to one or more appropriate comparators
  - Price reduction (including where there is a greater degree of uncertainty but in a therapeutic area with significant unmet clinical need*)
- Do not reimburse (this could lead to reassessment or continued evidence generation)

*Considerations for Significant Unmet Need - In exceptional cases where there is uncertain clinical and pharmacoeconomic evidence, the CADTH drug expert committees may issue a recommendation to reimburse with conditions, due to practical challenges in conducting robust clinical trials and pharmacoeconomic evaluations and in the presence of significant unmet medical need. In these situations, although there is uncertainty with the clinical evidence, the available evidence must reasonably suggest that the drug under review could substantially reduce morbidity and/or mortality associated with the disease. Significant unmet clinical need is identified on a population basis, including rarity and absence of alternatives.

OUTPUT

Meeting minutes are not made public. Embargoed recommendations are shared in confidence with the manufacturer and drug plans. The manufacturer may file a request for reconsideration and/or the drug plans may file a request for clarification. =

Final recommendations, CADTH review reports, and patient input submissions are made available on the CADTH website. The Implementation Support and Knowledge Mobilisation Team works within Canadian jurisdictions and provides services and knowledge tools to CADTH stakeholders about reports and recommendations.

CURRENT CONSULTATION

Following the 2020 HTAi Global Policy Forum on deliberative processes and the finding of the top three principles being those of transparency, inclusivity and impartiality, CADTH is conducting a two-part patient consultation. This will include a live education session on how CADTH Expert Committees Deliberate on February 4th 2021 followed by a consultation webinar between patient groups and CADTH staff on February 11th 2021.

During the live consultation webinar, and in small groups, participants will explore:

- What needs to occur during deliberations to ensure that patients' needs, expectations and experiences be meaningfully heard and considered by the committee?
- What aspects of the committee deliberation are important for you to see communicated in the recommendations document?
- **How would you suggest CADTH communicates about evidence uncertainties?**

Insights from the discussion will be shared publicly. CADTH will also follow its usual consultation process for any proposed changes to our programs.

In addition, in an effort to improve consistency and transparency (following the HTAi 2020 GPF), CADTH is undertaking a review to harmonize the expert drug committee processes.

Institute for Clinical and Economic Review (ICER), USA

Summary of key aspects of Current Value Framework (2020)

INPUT

Following synthesis of the evidence by qualitative and quantitative techniques, ICER assigns overall evidence ratings to each of the interventions evaluated in its appraisal. ICER developed the ICER Evidence Rating Matrix™ to evaluate the overall strength of evidence for a variety of outcomes, including: the magnitude of the difference of net health benefit between an intervention and comparator, and; the level of certainty in the best point estimate of net health benefit.

The Evidence Report also includes one-way sensitivity analyses, presenting the results in "tornado diagrams" and a table containing the ranges and distributional assumptions around the input parameters varied. PSA is also conducted and the results are presented in tabular fashion in terms of the percentage of simulations that achieve $50,000, $100,000, $150,000, and $200,000 per QALY thresholds, and graphically using scatter plots or cost-effectiveness acceptability curves.

Specific scenario analyses (including one using a modified societal perspective that incorporates estimates such as productivity losses, caregiver burden, and other indirect costs) and subgroup analyses are conducted when appropriate. In addition, the report presents results from threshold analyses which estimate the intervention prices that correspond to cost-effectiveness thresholds extending from $50,000 per QALY gained to $200,000 per QALY gained.

Evidence Reports include a sub-section on "Uncertainty and Controversies" in order to broaden discussion of alternative model structures and assumptions suggested by manufacturers or other stakeholders. One important goal of this section is to provide further elaboration of the rationale behind methodological decisions that underpin the base case. This sub-section also serves as an avenue to discuss how different assumptions or scenarios might affect model results and as a useful tool for decision-makers to understand the issues and uncertainties that may remain controversial. The sub-section provides discussion of different model variations that could be viewed as more conservative or optimistic, addressing alternative model structures or inputs that differ importantly from the base case. This sub-section also consolidates and expands discussion of factors related to uncertainty, including lack of information on natural history, limitations of the data on patient outcomes, difficulties translating existing data into measures of quality of life, and disagreements over the plausibility of certain inputs or assumptions.

ICER review and compare the current model to published models that included the same interventions or comparators of interest, were developed in the last 10 years, and were similar to the current model from a setting and population perspective. ICER recently reviewed methods pertaining to Single and Short Term Therapies and updated the guidance for considering input uncertainty in this context. For further details see the section below ("recent modifications").

THROUGHPUT

ICER only vote on long-term value for money if the committee has voted that the evidence is adequate to determine that the treatment in question has a net positive health benefit compared with the comparator. If there is too much uncertainty for the committee to vote to affirm the comparative clinical effectiveness of the treatment, then there is no vote on value.

The appraisal committee will then vote on whether key model assumptions and/or elements of uncertainty in the findings of the economic model suggests that the findings of the base-case analysis are overly optimistic or pessimistic. The following is rated:

- Uncertainty or overly favorable model assumptions creates significant risk that base-case cost-effectiveness estimates are too optimistic, OR;
- Uncertainty or overly unfavorable model assumptions creates significant risk that base case cost-

effectiveness estimates are too pessimistic

Broad themes of discussion form the foundation for policy recommendations and include:

- Evidence-based insurance coverage policy
- Pricing and payment mechanisms
- Future research needs
- The guidance that clinical specialty societies and patient organizations should provide to their communities

OUTPUT

Draft evidence report is posted to ICER's website for public comment.

All meeting materials made available on ICER website and recordings of entire committee meeting (evidence presentation, deliberation, and voting) posted on YouTube.

Newly developed Patient Engagement Program to provide outreach and debrief to patient groups.

RECENT MODIFICATIONS

ULTRA-RARE DISEASES

ICER will not change its approach to rating evidence according to the ICER EBM matrix, nor will there be different "standards" of evidence. Instead, ICER will provide specific context regarding the potential challenges of generating evidence for these treatments, including considerations of challenges to conducting RCTs, to validating surrogate outcome measures, and for obtaining long-term data on safety and on the durability of clinical benefit. The commonly used approach of evaluating treatments for ultra-rare diseases against historical controls will be highlighted. This added contextual language will be highlighted through special formatting in ICER reports and retained throughout press releases, executive summaries, and other versions of ICER reports.

For report sections on "other benefits and disadvantages" and "contextual considerations," ICER will include a broader frame to seek evidence and perspective on the potential for these treatments to affect the infrastructure for screening and care of the affected individuals. ICER will develop a specific template for the patient community and others to facilitate input on these elements of value that will have a meaningful role in the report and voting in the public meeting.

SINGLE AND SHORT-TERM THERAPIES (SST)

An adapted assessment approach for high-impact SSTs, defined as: "therapies that are delivered through a single intervention or a short-term course (less than one year) of treatment that offer a significant potential for substantial and sustained health benefits extending throughout patients' lifetimes" is used.

This includes:

Cure proportion modeling. ICER will make cure proportion modeling its standard reference case for high-impact SSTs whenever relevant, but to address uncertainty we will also provide survival analysis based on other modeling approaches when feasible.

Optimistic and conservative benefit scenarios. In addition to the base case and associated sensitivity analyses, ICER will develop two specific scenario analyses to reflect an optimistic and a conservative assumption regarding the benefit of SSTs under review. Input for best approaches to modeling the optimistic and conservative scenarios will be sought beginning with the scoping phase and will be included as part of the model analysis plan. These scenario analyses will be presented in conjunction with the base case for consideration by the independent appraisal committees.

Threshold analyses for durability of effect. When the SST price is known or can be estimated, assessments of SSTs will also include a scenario with a threshold analysis determining the duration of beneficial effect (e.g. cure) for those patients receiving short-term benefit that would be needed to achieve standard cost effectiveness thresholds (e.g., $150,000/QALY).

A new economic review section on "Uncertainties and Controversies." ICER will add a new section in the "Long-Term Cost-Effectiveness" section of ICER reports which will discuss "Uncertainty and Controversies" related to the economic evaluation. This new section will be added to all ICER reports, not just those for high-impact SSTs.

Institute for Quality and Efficiency in HealthCare (IQWiG) and Gemeinsamer Bundesausschuss (G-BA), Germany

Summary of key aspects of Current Methods Guide (2020)

INPUT

The Institute uses the following 3 categories to grade the degree of qualitative certainty at the individual study level and outcome level:

- **high qualitative certainty of results:** results on an outcome from a randomized study with a low risk of bias
- **moderate qualitative certainty of results:** results on an outcome from a randomized study with a high risk of bias
- **low qualitative certainty of results:** results on an outcome from a non-randomized comparative study

For dossier assessments (on new drugs), the operationalization for determining the extent of added benefit comprises 3 steps:

1) In the first step the probability of the existence of an effect is examined for each outcome separately (qualitative conclusion). For this purpose, the criteria for inferring conclusions on the evidence base are applied. Depending on the quality of the evidence, the probability is classified as a hint, an indication or proof.

2) In the second step, for those outcomes where at least a hint of the existence of an effect was determined in the first step, the extent of the effect size is determined for each outcome separately (quantitative conclusion). The following quantitative conclusions are possible: major, considerable, minor, and non-quantifiable.

3) In the third and last step, the overall conclusion on the added benefit according to the 6 specified categories is determined on the basis of all outcomes, taking into account the probability and extent at outcome level within the overall picture. These 6 categories are as follows: major, considerable, minor, and non-quantifiable added benefit; no added benefit proven; the benefit of the drug under assessment is less than the benefit of the appropriate comparator therapy.

Steps 2 and 3 are done only for dossier assessments, but step 1 (classification of effect certainty as hint, indication or proof) is performed in all assessments.

For assessments of potential (a specific application-based assessment of new non-drug interventions) – an extended assessment of the qualitative certainty of results of non-randomized studies is performed. In this context, besides the levels mentioned for randomized studies (high or moderate certainty of results) the following grades are used:

- **low qualitative certainty of results:** result of a higher quality non-randomized comparative study with adequate control for confounders (e.g. quasi-randomized controlled studies, non-randomized

controlled studies with active allocation of the intervention following a preplanned rule, prospective comparative cohort studies with passive allocation of the intervention),

- **very low qualitative certainty of results:** result of a higher quality non-randomized comparative study (see point above), but without adequate control for confounders or result of another non-randomized comparative study (e.g. retrospective comparative cohort studies, historically controlled studies, case-control studies),
- **minimum qualitative certainty of results:** result of a non-comparative study (e.g. one-arm cohort studies, observational studies or case series, cross-sectional studies or other non-comparative studies), which allows an indirect comparison with the results of other studies (literature controls).

Health Economic Evaluation (for information: Health economic evaluations are not performed regularly for reimbursement issues [see: Benefit Assessment of Medicinal Products - Gemeinsamer Bundesausschuss (g-ba.de)] Moreover IQWiG is currently in the process of revising its methodology in health economics)- the investigation of uncertainty must be considered in all areas. Parameter uncertainty are quantified with both univariate and multivariate deterministic as well as PSA should be performed and fully documented. Structural sensitivity analyses could be performed to investigate the impact of a variation of assumptions in the model structure, for example, the number or type of the model states.

For the deterministic sensitivity analysis, extreme levels of the input parameters should be provided for which the new intervention possibly saves costs or lies above or below the efficiency frontier. For univariate and multivariate analyses the results must be presented in a table and in a tornado diagram in which the levels of the results are displayed as an interval for the corresponding intervals of the input parameters. For PSAs the proportion of simulations for which cost savings or a position above or below the efficiency frontier arises is provided as a percentage. In the case of PSAs the results are presented as cumulative cost distributions. In addition the net health benefit could be calculated and presented. A budget impact analysis should also be undertaken, with the reference scenario and the predicted new mix of interventions presented as a range.

THROUGHPUT

The Institute's assessments on the benefits and harms of interventions are therefore normally based only on studies with sufficient certainty of results. This ensures that the decisions made by the G-BA, which are based on the Institute's recommendations, are supported by a sound scientific foundation. Moreover, an assessment that includes a literature search for studies with insufficient certainty of results would be costly and time consuming.

If it emerges that studies of the required quality and precision are generally lacking, it is the core task of the Institute to describe the circumstances and conclude that on the basis of the "currently best available" evidence, it is not possible to make reliable recommendations.

It is the G-BA's responsibility to take this uncertainty into account in its decision-making processes. In addition to considering scientific evidence, the G-BA also can consider other aspects in its decisions, such as the efficiency of interventions as well as the needs and values of people. In an uncertain scientific situation, such aspects become more important. In addition, the G-BA also has the option to call for or initiate studies in order to close the evidence gaps identified.

Every new active pharmaceutical ingredient requires an early benefit assessment within six months after it is launched on the German market. During the early benefit assessment, the G-BA examines whether the drug is really something new: if it offers patients greater benefit than comparable treatments that are already available. For example, if the drug causes fewer side effects, or the side effects are less severe, that is an additional benefit. Orphan drugs that are licensed for use in Germany (with turnover less than 50Mio €) have an "assumed additional benefit" , according to a recent report, this means that 73% of orphan drugs are designated with a "non-quantifiable benefit" (additional info: all medicinal

products with market authorisation in Germany are reimbursed

For non-drug interventions, the G-BA determines if they deliver patient benefit, and if they are necessary and economical. For these technologies, if the study findings are too inconclusive, the G-BA can initiate (and fund) studies if the benefit of a technology has not been conclusively researched but there is reason to think it could show potential.

OUTPUT

Meetings conducted in public and resolutions are published online with lay summaries available.

Pharmaceutical companies and device manufacturers may request regulatory scientific advice from the G BA. This advice can address the question, which research design is considered appropriate by G-BA.

Pharmaceutical companies may submit the dossier to the G-BA in advance for the purpose of a check for formal completeness without regard to data contents. Where necessary, the G-BA will then give comments in writing to the company which additional documents or information is required.

Haute Autorité de Santé (HAS), France

Summary of key aspects of Methods guides

INPUT

For reimbursement purposes, all medicinal products are submitted to HAS and its Transparency Committee (TC) for clinical assessment and appraisal; only some of them are also assessed on a medico-economics side by the dedicated committee (economic assessment and public health commission/ (CEESP).

General principles of TC evaluation are set in its "doctrine", whose last version dated on December 2020

https://www.has-sante.fr/upload/docs/application/pdf/2019-07/doctrine_de_la_commission_de_la_transparence_-_version_anglaise.pdf

TC assessment of the clinical benefit (CB) in the event of major uncertainties:

The assessment of some medicinal products is based on the analysis of early or still limited data, in a context of major uncertainty with respect to the real effect of the latter. In addition, the safety profile of these medicinal products is generally associated with high levels of uncertainty, particularly in the medium and long term, given the early nature of the data. In this case, the TC may consider that additional data will be essential to reassess the medicinal product. It may then specify in its opinion the information and additional studies essential for reassessment of the clinical benefit of the medicinal product that must be submitted by the pharmaceutical company, within a period stipulated in the opinion (article R. 163-18 of the French Social Security Code). Pending new data, a CB may be allocated in situations in which the absence of reimbursement in view of the preliminary data is liable to result in a loss of opportunity for patients:

- serious disease, irrespective of its prevalence, and
- unmet medical need, and
- initial data suggestive of a clinical utility for the patient, and
- development plan enabling the elimination of uncertainties in the short term on the basis of: • clinical studies: in this case, the development plan must be predefined by the company and known at the time of the initial assessment to enable the TC to assess whether it will be able to eliminate the uncertainties, • and/or real-life studies capable of eliminating uncertainties.

The maintenance of a sufficient CB level following reassessment may then be envisaged only if the

results eliminate the uncertainties identified in the initial assessment.

During the initial assessment of a medicinal product, the available data primarily comes from clinical trials and data obtained in real-life use conditions is rare. The TC may therefore identify uncertainties or questions concerning the clinical benefit of the medicinal product, its place in the therapeutic strategy in view of the alternatives, misuse, as well as the short or long-term consequences of introduction of the medicinal product on public health. The TC may request that additional data essential for subsequent reassessment of the clinical benefit or clinical added value of the medicinal product be collected. This involves post-registration studies (PRS), which may concern any medicinal product and are generally requested at the time of the initial assessment or an indication extension but may also be requested during a reassessment. The results of PRS must be submitted to the Committee by a date stipulated by the Committee in its opinion.

The results of these studies are systematically assessed and contribute to reassessment of medicinal products by the TC.

Randomised clinical trials remain the reference design to demonstrate the efficacy of a medicinal product. This means that observational studies cannot be a substitute for clinical studies in situations in which the latter are expected or provide evidence of efficacy that clinical studies have failed to demonstrate.

TC provide also recommendations in case of basket trials and/or use of biomarkers.

Regarding medico economics, HAS and CEESP has updated its methodological guidance: https://www.has-sante.fr/upload/docs/application/pdf/2020-11/methodological_guidance_2020_-choices_in_methods_for_economic_evaluation.pdf

Analysing the uncertainty associated with the conclusion of the evaluation: Each evaluation involves a degree of uncertainty, imprecision and methodological inconsistency. The results of the analyses making it possible to explore this uncertainty should systematically be interpreted in the submitted dossier, as the analysis is not limited to summarising quantitative results. A clear, justified discussion should make it possible to estimate the robustness of the conclusion. For the evaluation of cost-effectiveness, the conditions under which the cost-effectiveness frontier* would be altered should be defined.

A systematic exploration of the sources of uncertainty associated with the evaluation's structural choices, the modelling choices and the model parameter estimations should be presented according to an appropriate methodology. Sensitivity analyses should quantify the impact of a different structural choice in the reference case analysis (e.g., perspective, time horizon, population analysed, comparators, discount rate).

Sensitivity analyses should quantify the impact of methodological choices and modelling assumptions (e.g. model structure, data sources, calculation methods or assumptions to estimate the value of parameters not directly observed). The impact of the assumptions used for the extrapolation of treatment effects should be systematically explored.

The uncertainty associated with the parameters of the model should be systematically explored using two complementary approaches: a probabilistic sensitivity analysis, based on a second-order Monte Carlo simulation, and deterministic sensitivity analyses identifying the parameters (or combinations of parameters) which have the greatest influence on the results of the evaluation. For all of the sensitivity analyses presented, the credibility of the options tested should be justified, along with an interpretation of their results and their contribution to the comprehension of the evaluation. When a scenario which is fundamentally different to that used in the reference case analysis is put forward, the presentation of its results should include a thorough exploration of the uncertainty through deterministic and probabilistic sensitivity analyses.

Quantitative results should be presented and interpreted in a way which is consistent with the objective of the economic evaluation. The evaluation of cost-effectiveness requires the identification of the interventions on the cost-effectiveness frontier and the presentation of the results based on the incremental cost-effectiveness ratio (ICER) or net benefit (NB). All of the relevant economic information to aid public decision-making should be extracted from the evaluation. A clear, justified discussion should make it possible to estimate the robustness of the conclusion of the evaluation and define the conditions under which the conclusion would be altered. This discussion should rest on a critical appraisal of the methods and data used, and on the sensitivity analyses conducted. The degree of confidence associated with the results should be detailed.

The economic evaluation should be presented is a well-structured, clear and detailed way. The methodology should be transparent. The data and sources used should be clearly presented. For each intervention, non-discounted values should be presented for each major cost or result component. The total costs and health outcomes obtained on the main criterion should then be calculated and discounted.

THROUGHPUT

Given the available data, the Transparency Committee (TC) assesses the Actual Clinical Benefit (ACB) (criteria for reimbursement opinion) by considering:

- efficacy
- adverse effects
- intended role in the therapeutic strategy in comparison with other available therapies
- severity of the disease/condition
- public health benefits

The 4 ACB levels used for medicinal products are:

- "Sufficient":
  - Substantial ACB: 65% reimbursement
  - Moderate ACB: 30% reimbursement
  - Low ACB: 15% reimbursement
- "Insufficient"

The Clinical Added Value (CAV) (used to negotiate the price) takes into account:

- Comparative efficacy and safety data with regards to available treatments (reference medicinal product or better treatment modalities)
- Particular attention is paid to the following criteria: – the quality of the demonstration, which includes the comparison and the choice of comparator(s), the methodological quality of the study, the appropriateness of the population included for the indication, the relevance of the clinical endpoint and its significance, etc.; – the effect size in terms of clinical efficacy, quality of life and safety in view of the robustness of the demonstration; – the clinical relevance of this effect compared to clinically relevant comparators; in view of the medical need.
- A medicinal product with no CAV can only be included onto the list of medicines for reimbursement if it offers savings in terms of treatment cost.
- The 5 CAV levels used for medicinal products when compared with existing therapeutic interventions are:
  - I: major
  - II: substantial
  - III: moderate
  - IV: minor
  - V: no improvement

OUTPUT

A draft recommendation based on the results of the TC vote is written by the project team. The sponsor is given 10 days to respond with written comments or request for hearing.

The final recommendation is sent to the government ministers, the sponsor, and published on the HAS website. Agenda and meetings minutes are also published on HAS website and include the date of session, list of members present and excused, list of external individuals present, COIs, topics examined, claims of the sponsor, content of deliberations, and result of voting.

The TC and the CEESP report on their work in the HAS annual activity report, including information on the recommendations made in that year.

RECENT MODIFICATIONS

Published in January 2020, HAS have issued an "action plan for innovative medicines" to keep pace with innovation. : https://www.has-sante.fr/upload/docs/application/pdf/2020-03/innovative_medicine_action_plan_27.01.20.pdf

The plan includes the following:

1. Issue conditional reviews until uncertainties are lifted
2. Monitor medicines in real-life conditions to make sure they fulfil initial promises
3. Reinforce HAS agility to better support innovation
    a. Focus on assessments offering high added value
    b. Improve early dialogues process to support clinical developments
    c. Promote accelerated assessment procedures ("fast-tracking")
    d. Improve internal collective efficiency
4. Systematically involve patients and consumers
5. Improve transparency
    a. In terms of timeframes
    b. In terms of post-launch studies
6. Increase European cooperation to pool knowledge

Annex 3 – Case Studies

Case Study One: CAR T therapy

In order to more easily conceptualize how HTA agencies across the world consider uncertainty in the same evidence base, a case study is described. Chimeric Antigen Receptor T-Cell, or "CAR T" therapy is an immunological treatment that uses the body's own immune system to destroy cancerous cells. The first licensed CAR T therapy was Kymriah™ (tisagenlecleucel; Novartis) in 2017, which is currently licensed to treat patients up to 25 years of age with B-cell precursor acute lymphoblastic leukemia (ALL) and for adults with relapsed or refractory large B-cell lymphoma (DLBCL). This type of genetically-engineered therapy was hailed as a "breakthrough" for cancers affecting young people with very few viable treatment alternatives. According to multiple websites, the average list-price for a one-time treatment course ranges from $373,000 to $475,000 USD.

At the time of most appraisals in 2017 and 2018, the clinical evidence for Kymriah™ assessed by HTA agencies came from three single-arm studies which were pooled together by the manufacturer (n=193). Data on long-term benefits and adverse events were not available at the time of appraisal. In reviewing the publicly available documentation and summaries of the HTA appraisals of Kymriah™ from seven HTA agencies across the world, all agencies highlighted the levels of uncertainty present in the evidence base, particularly the lack of a direct comparison with standard of care (instead small single-arm trials using historical comparisons) and no long-term clinical data as major concerns.

The resulting outcomes of the HTAs of Kymriah™ have the following commonalities:

- All agencies recognised the innovative nature of the treatment
- Most recommended rejecting the treatment based on the initial submission and multiple submissions/ assessments were conducted (with 3 re-submissions to MSAC in Australia)
- Most agencies have since recommended Kymriah™ for use in one or both of the licensed indications (noting that some additional evidence became available during this time); the SMC in Scotland has not recommended the use of Kymriah™ for the DLBCL indication as it considered that the treatment cost in relation to health benefits were not sufficiently justified and the economic analysis for this indication was not sufficiently robust
- All agencies that recommended the use of Kymriah™ have done so in the presence of an accompanying risk sharing scheme (whereby the cost is reduced/shared or systems pay only while the treatment is effective) and/or through alternative pathways (for example through the Cancer Drugs Fund in England) and with additional data collection. It was the reduction in costs to the payer and planned generation of comparative evidence that justified the reimbursement recommendations.
- Many agencies have stipulated strict timelines for conducting reviews/reassessments (some with annual reviews, for example NICE in England and HAS in France) to ensure the value is attained.

Agency, country	Recommendation	Key Uncertainties & Other Comments
NICE, England	Recommended for ALL and DLBCL through the Cancer Drugs Fund and with MEA	Approval conditional on collection of additional data from ongoing clinical trials and from RWE
SMC, Scotland	Recommended for ALL with a patient access scheme. Not recommended for DLBCL	For the DLBCL indication the treatment cost in relation to the health benefits were not sufficiently justified and the economic analysis was not considered robust
MSAC, Australia	Recommended for ALL and DLBCL with Managed Entry Agreement	Approved in 2019 after more data were available. Strict governance and criteria for usage in place in the recommendation. Considered by MSAC as a "hybrid" therapy

CADTH, Canada	Recommended for ALL and DLBCL with condition of a reduction in price	Long-term, standardized RWE for both indications to be collected with the set-up of a pan-Canadian registry. A revised process for assessing gene and cell therapies has since been introduced.
ICER, USA	Kymriah™ was considered to provide small-to-substantial (cost effective) net health benefit compared with current salvage chemotherapy	Uncertainty from non-comparative studies, small patient size and short follow-up made it difficult to assess the magnitude of health benefit.
IQWIG/G-BA, Germany	Recommended for ALL and DLBCL with a rebate-based agreement	Met orphan drug status though extent of additional benefit considered "non-quantifiable" due to data scarcity, uncertainty in short follow-up, incomplete patient recruitment, impact of bridging therapy and indirect comparison with historical evidence
HAS, France	Recommended with review in 12 months and with MEA	Kymriah was found to have high actual clinical benefit and moderate clinical added value for ALL. It had high actual clinical benefit and minor clinical added value for DLBCL. Uncertainties remained

Case Study Two: Spinal Muscular Atrophy

Spinal Muscular Atrophy (SMA) is an autosomal recessive neuromuscular disease caused by deletions or mutations on the survival motor neuron gene. SMA leads to loss of lower motor neurons leading to progressive symmetrical muscle weakness and atrophy causing swallowing and breathing difficulties in affected children. In its most common form (Type 1 SMA), symptoms begin in infancy, with later onset SMA (types 2 and 3) starting later in life and now with genetic testing, children can also be diagnosed pre-symptomatically. Without treatment children have reduced life expectancies, with many Type 1 SMA patients dying by age 2. SMA is the most common inherited cause of childhood mortality (though SMA is itself an ultra-rare condition). There are currently two technologies licensed for treatment of SMA; the first was the antisense oligonucleotide Spinraza™ (Nusinersen; Biogen) and the second was the gene replacement therapy Zolgensma™ (onasemnogene abeparvovec; AveXis/Novartis). At the time of launch, Spinraza™ (given as an initial and then annual maintenance doses) was listed at $750,000 USD for the first year and $375,000 USD annually thereafter. Zolegnesma™ (given as a single dose) was launched as "the most expensive drug in the world" with a list price of $2.125 million USD.

The clinical evidence base for Spinraza™ at the time of most appraisals primarily consisted of two studies (ENDEAR, n=122 and CHERISH, n=126) with some ongoing (continuation) studies underway. Major limitations of this evidence base identified by HTA agencies included poorer baseline prognosis in control groups, the trial population not being reflective of the population in practice, differences in the dosing regimens in the trials versus real world and short follow-up periods (although ENDEAR was stopped early on ethical grounds due to a strong treatment benefit). Similar to the first case study of CAR T therapy, at the time of appraisal, the clinical evidence base for Spinraza™ was considered short and so the long-term benefits of the technology were deemed uncertain. However, acknowledging the potential for therapeutic effect for a disease affecting babies and young children with a huge unmet need, almost all jurisdictions recommended the use of Spinraza™ with the implementation of managed entry agreements and careful patient selection (with diversity between recommendations for pre-symptomatic patients, type 1 and types 2&3). The proposed agreements included various risk management strategies, such as patient selection (based on SMA type and age), starting and stopping rules, data collection, patient consent and commercial offer.

Zolgensma™ has recently been approved as the first gene therapy for Type 1 SMA and is given as a single dose. There are no data available to assess the net benefit of Zolgensma™ in presymptomatic or later-onset

SMA populations. The evidence base for Zolgensma™ currently includes two Phase 3 studies (STR1VE-US and SPR1NT) and Phase 1 START trials. While formal HTA recommendations have yet to be issued regarding the use of Zolgensma™, a number of jurisdictions have granted early access to the technology (such as with the ATU Framework in France; a system whereby exceptional use of drugs that do not have a marketing authorisation that treat serious or rare diseases with no other appropriate alternative treatments can be granted). Some insurance providers and payers in the US have also reimbursed the technology for use after ICER stated that the technology could potentially be considered cost effective at the expected price point.

In the meantime, issues and queries that have been raised around Zolgensma™ (and similarly CAR T therapies), including the impracticality of stopping rules and more traditional performance-based agreements given the single-dose nature of the treatment. Some agencies have indicated that other mechanisms could be explored, such as payment by instalments while the patient continues to respond to therapy. In addition, there could be potential challenges and additional uncertainty created by the proximity of licensure of these two innovative "game-changing" treatments for an ultra-orphan condition. In jurisdictions where Spinraza™ has now been recommended for Type 1 SMA, the standard of care technically changes from best supportive care to an alternative, costly therapeutic. In cases such as these, where new comparators are approved prior to the assessment of new interventions there are additional challenges in determining the true value of the new offering.

Considerations of Spinraza™ by country

Agency, country	Recommendation	Key Uncertainties & Other Comments
NICE, England	Recommended for pre-symptomatic SMA and SMA 1-3; with an MEA	Disease rarity, severity and end-of-life criteria were all considered. The implementation of an MEA was considered sufficient to mitigate uncertainties, with a 5 year review window.
SMC, Scotland	Recommended for symptomatic SMA 1. SMA 2 & 3 later recommended under ultra-orphan pathway; both with MEA in place.	The benefit was considered in the context of disease modifiers (absence of other treatments of proven benefit and a substantial improvement in life expectancy and an ultra-orphan medicine); the SMC can accept greater uncertainty in the economic case. The economic case for pre-symptomatic SMA was not presented by the manufacturer.
PBAC, Australia	Recommended for symptomatic SMA 1-3 for children under 18 years with MEA in place. This was extended to pre-symptomatic children in December 2020; not recommended for those over 19 years old.	PBAC noted its requirement for further information regarding: ▪ the Australian SMA patient population, especially patient numbers and variations ▪ the cost effectiveness ratio of Spinraza™ as a treatment option ▪ availability of further clinical evidence ▪ definition of the appropriate adult population to receive Spinraza™ as a treatment ▪ how Spinraza™ performs in the context of rapid developments in SMA treatments and diagnosis.
CADTH, Canada	Recommended for symptomatic SMA 1 for children under 6 months with a price reduction.	Strict eligibility criteria in place, including stopping rules. Evidence for other classes of SMA was not considered sufficient, but another HTA for these groups is underway.

ICER, USA	Spinraza™ was not considered to meet cost effectiveness thresholds.	While there was a net health benefit for Spinraza™ (including improved quality of life), the current price of Spinraza™ "far exceeds common thresholds for cost-effectiveness". Zolgensma™ was also reviewed by ICER and it was noted that as it is administered as a one-time dose, it offers less complexity than Spinraza™. Key recommendations included: ▪ To align reasonably with the benefits for patients and families, the price for Spinraza™ should be far lower than it is, and the price for Zolgensma™ should be lower than the hypothetical $4-5 million price the manufacturer has suggested could be justified. To achieve the needed balance between incentives for innovation and health system affordability, all manufacturers should exercise their monopoly pricing power responsibly, setting prices that do not exceed a reasonable cost-effectiveness threshold. ▪ Payers should negotiate outcomes-based contracts under which a substantial portion of treatment cost is at risk should patients not receive adequate clinical benefit. Outcomes measures should extend beyond death and permanent ventilation, which might not be able to capture near-term lack of benefit for some Type I patients and are inadequate measures for treatment of later-onset or presymptomatic patients.
BeNeLuxA – Belgium, Netherlands, Luxembourg, Ireland and Austria.	Belgium – recommended for all patients eligible for treatment. Netherlands – recommended for subgroups of patients according to age, disease severity and diagnosis. Luxembourg, Austria and Ireland were not part of the decision making.	The evidence was assessed by the Dutch HTA however the coverage recommendations were broader in Belgium than in the Netherlands. The countries had different procedures for providing reimbursement beyond the patients specified in the HTA report (i.e. the Netherlands can only recommend treatment in populations that have undergone formal HTA). MEA with additional data collection requirements were implemented in both countries. The recommendation concerning Spinraza™ was the first positive recommendation from the BeNeLuxA collaborative.
HAS, France	Recommended for all SMA patients (initially available to children under the ATU temporary scheme)	*Temporary access to Spinraza™ (at list price) was arranged under the ATU framework after the EMA approval in 2017.* During the HTA, it was considered that Spinraza™ provides substantial clinical benefit with moderate clinical added value. An MEA is in place (with costs also being recouped from the payment at full list price for the technology under the ATU framework).

Case Study Three: Transvaginal Mesh for Pelvic Organ Prolapse and Stress Urinary Incontinence

Pelvic organ prolapse (POP) and stress urinary incontinence (SUI) are stressful and quality-of-life-limiting dysfunctions. Up to 50% of women who have given birth are estimated to develop POP or SUI during their lives with up to one in five women requiring surgical intervention(90). Mesh augmented repair using synthetic polypropylene mesh was introduced in the 1990s on the basis that it successfully corrected incisional hernias and therefore would have comparable effects in correcting POP and SUI. A number of synthetic mesh products were approved for use by regulatory agencies, however due to the marketing process for surgical device, there was an average of 5 years between the approval of the first urogynecological implant and the first published RCT.

The history of the use of the transvaginal mesh for POP and SUI is detailed below:

- The Food and Drug Administration (FDA) in the US was the first agency to publish a safety report over the use of the transvaginal mesh for POP and SUI in 2008, based on an increasing number of reported complications. A lack of published studies was also noted.
- In 2011, a second safety report was published outlining significant adverse effects without improvements in symptoms or quality of life. The FDA then imposed "Post marketing Surveillance Studies".
- In 2016, the FDA increased the risk of transvaginal meshes from Class II to Class III. In numerous lawsuits filed, only 12% of surgeons were found to be certified for POP and SUI surgeries.

Current usage around the world:

- In 2019, the FDA disallowed the sales of anterior transvaginal mesh implants; a petition for the continued – and targeted- use of the mesh in POP was started.
- In 2017, NICE in the UK recommended that mesh augmented POP repair should be used only in research. In 2018, this use was paused and in 2019, NICE recommended specific types of mesh and procedure for targeted patients, and only as a last resort once the patient has been fully informed of her options. A mandatory register for surgeons is being established.
- In Australia, transvaginal meshes are not recommended for use. However, suburethral slings for SUI are considered safe for use. The same recommendations apply in New Zealand.
- In Canada, the posterior use of the mesh is not recommended. However, transvaginal meshes for anterior and apical repair can be used, but only for patients at high risk of recurrence with no other viable treatment options.
- In Asia, there is a very diverse use of the mesh and self-cut ("ready-made") mesh kits are used (although these are not licensed for use in countries such as Japan).
- In South America there are no limitations on the use of transvaginal meshes for POP or SUI.
- In the EU, transvaginal meshes can be used in the case of POP recurrence or in cases where the patient is at high risk of recurrence, with a certification system for surgeons and patients should be appropriately selected and counseled. However the recommendations still vary within countries, for example, in France the use of transvaginal meshes is not recommended, except in the context of clinical trials. In the Netherlands, transvaginal meshes are recommended for use in certain subgroups and the use and results from the mesh surgery is subject to regular audit.

Generally speaking, the public has become very skeptical and the use of surgical repair for POP and SUI has decreased across the world. While there are variable guidelines and recommendations on the use of transvaginal mesh, it has become apparent that it is crucial that patients are well informed about their options and the uncertainty within the evidence base, outlining the potential advantages and disadvantages of the procedure. Appropriate credentialing of surgeons and the need for longer-term follow up is a unified theme across all of the recommendations reviewed.

References

1. Bond K. 2020 HTAi Global Policy Forum. Deliberative processes in Health Technology Assessment: Prospects, Problems and Policy Proposals 2020.
2. Oortwijn W, Sampietro-Colom L, Trowman R. How to Deal with the Inevitable: Generating Real-World Data and Using Real-World Evidence for HTA Purposes – From Theory to Action. International journal of technology assessment in health care. 2019;35(4):346-50.
3. Husereau D, Henshall C, Sampietro-Colom L, Thomas S. CHANGING HEALTH TECHNOLOGY ASSESSMENT PARADIGMS? International journal of technology assessment in health care. 2016;32(4):191-9.
4. Klemp M, Frønsdal KB, Facey K. What principles should govern the use of managed entry agreements? International journal of technology assessment in health care. 2011;27(1):77-83.
5. Hutton J, Trueman P, Henshall C. Coverage with evidence development: an examination of conceptual and policy issues. International journal of technology assessment in health care. 2007;23(4):425-32.
6. Citation (n.d.). In Camridge English Dictionary 2020 [Available from: https://dictionary.cambridge.org/dictionary/english/uncertain
7. 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. Royal Society open science. 2019;6(5):181870.
8. HTAi I, Avalia-t. HTA Glossary
9. Fischhoff B, Davis AL. Communicating scientific uncertainty. Proceedings of the National Academy of Sciences of the United States of America. 2014;111 Suppl 4(Suppl 4):13664-71.
10. Hutton J, McGrath C, Frybourg J-M, Tremblay M, Bramley-Harker E, Henshall C. Framework for describing and classifying decision-making systems using technology assessment to determine the reimbursement of health technologies (fourth hurdle systems). International journal of technology assessment in health care. 2006;22(1):10-8.
11. O'Rourke B, Oortwijn W, Schuller T. The new definition of health technology assessment: A milestone in international collaboration. International journal of technology assessment in health care. 2020;36(3):187-90.
12. Luft J, Ingham H. The Johari Window: a graphic model of awareness in interpersonal relations. Human relations training news. 1961;5(9):6-7.
13. Chalmers I. Well informed uncertainties about the effects of treatments. BMJ (Clinical research ed). 2004;328(7438):475-6.
14. Thompson M TA, Fu R, et al. A Framework To Facilitate the Use of Systematic Reviews and Meta-Analyses in the Design of Primary Research Studies [Internet]. Agency for Healthcare Research and Quality (US).
15. Vreman RA, Naci H, Goettsch WG, Mantel-Teeuwisse AK, Schneeweiss SG, Leufkens HGM, et al. Decision Making Under Uncertainty: Comparing Regulatory and Health Technology Assessment Reviews of Medicines in the United States and Europe. Clinical pharmacology and therapeutics. 2020;108(2):350-7.
16. Sullivan TR, Latimer NR, Gray J, Sorich MJ, Salter AB, Karnon J. Adjusting for Treatment Switching in Oncology Trials: A Systematic Review and Recommendations for Reporting. Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research. 2020;23(3):388-96.
17. Santesso N, Carrasco-Labra A, Langendam M, Brignardello-Petersen R, Mustafa RA, Heus P, et al. Improving GRADE evidence tables part 3: detailed guidance for explanatory footnotes supports creating and understanding GRADE certainty in the evidence judgments. Journal of clinical epidemiology. 2016;74:28-39.
18. Alonso-Coello P, Schünemann HJ, Moberg J, Brignardello-Petersen R, Akl EA, Davoli M, et al. GRADE Evidence to Decision (EtD) frameworks: a systematic and transparent approach to making well informed healthcare choices. 1: Introduction. BMJ (Clinical research ed). 2016;353:i2016.

19. Higgins JP, Altman DG, Gøtzsche PC, Jüni P, Moher D, Oxman AD, et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. BMJ (Clinical research ed). 2011;343:d5928.

20. Owens DK, Lohr KN, Atkins D, Treadwell JR, Reston JT, Bass EB, et al. AHRQ series paper 5: grading the strength of a body of evidence when comparing medical interventions--agency for healthcare research and quality and the effective health-care program. Journal of clinical epidemiology. 2010;63(5):513-23.

21. Woolacott N, Corbett M, Jones-Diette J, Hodgson R. Methodological challenges for the evaluation of clinical effectiveness in the context of accelerated regulatory approval: an overview. Journal of clinical epidemiology. 2017;90:108-18.

22. Elston J, Taylor RS. Use of surrogate outcomes in cost-effectiveness models: a review of United Kingdom health technology assessment reports. International journal of technology assessment in health care. 2009;25(1):6-13.

23. Grigore B, Ciani O, Dams F, Federici C, de Groot S, Möllenkamp M, et al. Surrogate Endpoints in Health Technology Assessment: An International Review of Methodological Guidelines. PharmacoEconomics. 2020;38(10):1055-70.

24. Groot Koerkamp B, Stijnen T, Weinstein MC, Hunink MG. The combined analysis of uncertainty and patient heterogeneity in medical decision models. Medical decision making : an international journal of the Society for Medical Decision Making. 2011;31(4):650-61.

25. Briggs AH, Weinstein MC, Fenwick EA, Karnon J, Sculpher MJ, Paltiel AD. Model parameter estimation and uncertainty: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force--6. Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research. 2012;15(6):835-42.

26. Sanders GD, Neumann PJ, Basu A, Brock DW, Feeny D, Krahn M, et al. Recommendations for Conduct, Methodological Practices, and Reporting of Cost-effectiveness Analyses: Second Panel on Cost-Effectiveness in Health and Medicine. Jama. 2016;316(10):1093-103.

27. Briggs A, Sculpher M, Buxton M. Uncertainty in the economic evaluation of health care technologies: the role of sensitivity analysis. Health economics. 1994;3(2):95-104.

28. Bujkiewicz S, Jones HE, Lai MC, Cooper NJ, Hawkins N, Squires H, et al. Development of a transparent interactive decision interrogator to facilitate the decision-making process in health care. Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research. 2011;14(5):768-76.

29. Ades AE, Claxton K, Sculpher M. Evidence synthesis, parameter correlation and probabilistic sensitivity analysis. Health economics. 2006;15(4):373-81.

30. Fenwick E, Claxton K, Sculpher M. Representing uncertainty: the role of cost-effectiveness acceptability curves. Health economics. 2001;10(8):779-87.

31. Bala MV, Zarkin GA, Mauskopf J. Presenting results of probabilistic sensitivity analysis: the incremental benefit curve. Health economics. 2008;17(3):435-40.

32. Hatswell AJ, Bullement A, Briggs A, Paulden M, Stevenson MD. Probabilistic Sensitivity Analysis in Cost-Effectiveness Models: Determining Model Convergence in Cohort Models. PharmacoEconomics. 2018;36(12):1421-6.

33. Claxton K, Sculpher M, McCabe C, Briggs A, Akehurst R, Buxton M, et al. Probabilistic sensitivity analysis for NICE technology assessment: not an optional extra. Health economics. 2005;14(4):339-47.

34. National Institute for Health and Care Excellence. CHTE Methods Review - Exploring Uncertainty 2020 August 2020.

35. Jackson C, Stevens J, Ren S, Latimer N, Bojke L, Manca A, et al. Extrapolating Survival from Randomized Trials Using External Data: A Review of Methods. Medical decision making : an international journal of the Society for Medical Decision Making. 2017;37(4):377-90.

36. Latimer N, , . NICE DSU Technical Support Document 14: Undertaking survival analysis for economic evaluations alongside clinical trials - extrapolation with patient-level data. 2011.

37. Strong M, Oakley JE, Chilcott J. Managing structural uncertainty in health economic decision models: a discrepancy approach. Journal of the Royal Statistical Society: Series C (Applied Statistics). 2012;61(1):25-45.

38. Afzali HH, Karnon J. Exploring structural uncertainty in model-based economic evaluations. PharmacoEconomics. 2015;33(5):435-43.

39. Leelahavarong P. Budget impact analysis. Journal of the Medical Association of Thailand = Chotmaihet thangphaet. 2014;97 Suppl 5:S65-71.

40. Danzon PM, Drummond MF, Towse A, Pauly MV. Objectives, Budgets, Thresholds, and Opportunity Costs-A Health Economics Approach: An ISPOR Special Task Force Report [4]. Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research. 2018;21(2):140-5.

41. Bindels J, Ramaekers B, Ramos IC, Mohseninejad L, Knies S, Grutters J, et al. Use of Value of Information in Healthcare Decision Making: Exploring Multiple Perspectives. PharmacoEconomics. 2016;34(3):315-22.

42. Eckermann S, Willan AR. Expected value of information and decision making in HTA. Health economics. 2007;16(2):195-209.

43. Claxton K, Ginnelly L, Sculpher M, Philips Z, Palmer S. A pilot study on the use of decision theory and value of information analysis as part of the NHS Health Technology Assessment programme. Health technology assessment (Winchester, England). 2004;8(31):1-103, iii.

44. Eckermann S, Willan AR. Time and expected value of sample information wait for no patient. Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research. 2008;11(3):522-6.

45. Fenwick E, Steuten L, Knies S, Ghabri S, Basu A, Murray JF, et al. Value of Information Analysis for Research Decisions-An Introduction: Report 1 of the ISPOR Value of Information Analysis Emerging Good Practices Task Force. Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research. 2020;23(2):139-50.

46. Grimm SE, Pouwels X, Ramaekers BLT, Wijnen B, Knies S, Grutters J, et al. Development and Validation of the TRansparent Uncertainty ASsessmenT (TRUST) Tool for Assessing Uncertainties in Health Economic Decision Models. PharmacoEconomics. 2020;38(2):205-16.

47. Annemans L, Makady A. TRUST4RD: tool for reducing uncertainties in the evidence generation for specialised treatments for rare diseases. Orphanet journal of rare diseases. 2020;15(1):127.

48. Nicod E, Berg Brigham K, Durand-Zaleski I, Kanavos P. Dealing with Uncertainty and Accounting for Social Value Judgments in Assessments of Orphan Drugs: Evidence from Four European Countries. Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research. 2017;20(7):919-26.

49. Rothery C, Claxton K, Palmer S, Epstein D, Tarricone R, Sculpher M. Characterising Uncertainty in the Assessment of Medical Devices and Determining Future Research Needs. Health economics. 2017;26 Suppl 1:109-23.

50. McCabe C, Claxton K, Culyer AJ. The NICE cost-effectiveness threshold: what it is and what that means. PharmacoEconomics. 2008;26(9):733-44.

51. Campbell B, Campbell M. NICE Medical Technologies Guidance: a novel and rigorous methodology to address a new health technology assessment challenge. Applied health economics and health policy. 2012;10(5):295-7.

52. Single ANV, Facey KM, Livingstone H, Silva AS. Stories of Patient Involvement Impact in Health Technology Assessments: A Discussion Paper. International journal of technology assessment in health care. 2019;35(4):266-72.

53. Mandeville KL, Barker R, Packham A, Sowerby C, Yarrow K, Patrick H. Financial interests of patient organisations contributing to technology assessment at England's National Institute for Health and Care Excellence: policy review. BMJ (Clinical research ed). 2019;364:k5300.

54. de Wit M, Guillemin F, Grimm S, Boonen A, Fautrel B, Joore M. Patient engagement in health technology assessment (HTA) and the regulatory process: what about rheumatology? RMD Open. 2020;6(3):e001286.

55. Boothe K. "Getting to the Table": Changing Ideas about Public and Patient Involvement in Canadian Drug Assessment. Journal of health politics, policy and law. 2019;44(4):631-63.

56. Oortwijn W. MR, Tummers M., Tayler C.,. Wat leren we van de introductie van de Da Vinci robot? . Radboud University Medical Centre 2020.

57. HTAi Patient Citizen Involvement Group. International Summary Information of Products (SIP Template) 2020 [Available from: https://htai.org/wp-content/uploads/2020/10/Project-Proposal-SIP.pdf.

58. Garrison LP, Jr., Towse A, Briggs A, de Pouvourville G, Grueger J, Mohr PE, et al. Performance-based risk-sharing arrangements-good practices for design, implementation, and evaluation: report of the ISPOR good practices for performance-based risk-sharing arrangements task force. Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research. 2013;16(5):703-19.

59. A Vreman R, F Broekhoff T, GM Leufkens H, K Mantel-Teeuwisse A, G Goettsch W. Application of Managed Entry Agreements for Innovative Therapies in Different Settings and Combinations: A Feasibility Analysis. International journal of environmental research and public health. 2020;17(22):8309.

60. 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.

61. Gladwell D, Bullement A, Cowell W, Patterson K, Strong M. "Stick or Twist?" Negotiating Price and Data in an Era of Conditional Approval. Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research. 2020;23(2):191-9.

62. Claxton K, Palmer S, Longworth L, Bojke L, Griffin S, Soares M, et al. A Comprehensive Algorithm for Approval of Health Technologies With, Without, or Only in Research: The Key Principles for Informing Coverage Decisions. Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research. 2016;19(6):885-91.

63. Longworth L, Youn J, Bojke L, Palmer S, Griffin S, Spackman E, et al. When does NICE recommend the use of health technologies within a programme of evidence development? : a systematic review of NICE guidance. PharmacoEconomics. 2013;31(2):137-49.

64. Holland S, Hope T. The ethics of attaching research conditions to access to new health technologies. Journal of medical ethics. 2012;38(6):366-71.

65. Mundy L, Trowman R, Kearney B. Improving Access to High-Cost Technologies in the Asia Region. International journal of technology assessment in health care. 2019;35(3):168-75.

66. Allen N, Walker SR, Liberti L, Salek S. Health Technology Assessment (HTA) Case Studies: Factors Influencing Divergent HTA Reimbursement Recommendations in Australia, Canada, England, and Scotland. Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research. 2017;20(3):320-8.

67. Council NC. NICE Citizens Council Reports. Departing from the Threshold. London: National Institute for Health and Care Excellence (NICE)

68. Adkins EM, Nicholson L, Floyd D, Ratcliffe M, Chevrou-Severac H. Oncology drugs for orphan indications: how are HTA processes evolving for this specific drug category? ClinicoEconomics and outcomes research : CEOR. 2017;9:327-42.

69. Garrison LP, Jackson T, Paul D, Kenston M. Value-Based Pricing for Emerging Gene Therapies: The Economic Case for a Higher Cost-Effectiveness Threshold. Journal of managed care & specialty pharmacy. 2019;25(7):793-9.

70. Cookson R, McCabe C, Tsuchiya A. Public healthcare resource allocation and the Rule of Rescue. Journal of medical ethics. 2008;34(7):540-4.

71. Lakdawalla DN, Doshi JA, Garrison LP, Jr., Phelps CE, Basu A, Danzon PM. Defining Elements of Value in Health Care-A Health Economics Approach: An ISPOR Special Task Force Report [3]. Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research. 2018;21(2):131-9.

72. Jönsson B, Hampson G, Michaels J, Towse A, von der Schulenburg JG, Wong O. Advanced therapy medicinal products and health technology assessment principles and practices for value-based and sustainable healthcare. The European journal of health economics : HEPAC : health economics in prevention and care. 2019;20(3):427-38.

73. Kuchenbecker RS, Mota DM. Miracle drug: Brazil approves never-tested cancer medicine. Journal of oncology pharmacy practice : official publication of the International Society of Oncology Pharmacy Practitioners. 2017;23(5):399-400.

74. National Institute for Health and Care Excellence. Reviewing our methods for health technology evaluation: consultation 2020 [Available from: https://www.nice.org.uk/about/what-we-do/our-programmes/nice-guidance/chte-methods-consultation.

75. Griffiths EA, Hendrich JK, Stoddart SD, Walsh SC. Acceptance of health technology assessment submissions with incremental cost-effectiveness ratios above the cost-effectiveness threshold. ClinicoEconomics and outcomes research : CEOR. 2015;7:463-76.

76. Broekhuizen H, Groothuis-Oudshoorn CG, van Til JA, Hummel JM, MJ IJ. A review and classification of approaches for dealing with uncertainty in multi-criteria decision analysis for healthcare decisions. PharmacoEconomics. 2015;33(5):445-55.

77. Baltussen R, Marsh K, Thokala P, Diaby V, Castro H, Cleemput I, et al. Multicriteria Decision Analysis to Support Health Technology Assessment Agencies: Benefits, Limitations, and the Way Forward. Value in Health. 2019;22(11):1283-8.

78. Facey KM, Rannanheimo P, Batchelor L, Borchardt M, de Cock J. Real-world evidence to support Payer/HTA decisions about highly innovative technologies in the EU-actions for stakeholders. International journal of technology assessment in health care. 2020:1-10.

79. European Patients Forum. Patient Involvement in Health Technology Assessment in Europe 2013.

80. Buck D, Gamble C, Dudley L, Preston J, Hanley B, Williamson PR, et al. From plans to actions in patient and public involvement: qualitative study of documented plans and the accounts of researchers and patients sampled from a cohort of clinical trials. BMJ open. 2014;4(12):e006400.

81. Broomell SB, Kane PB. Public perception and communication of scientific uncertainty. Journal of experimental psychology General. 2017;146(2):286-304.

82. Gibbs P, Hanlon M, Hardaker P, Hawkins E, MacDonald A, Maskell K, et al. Making sense of uncertainty: why uncertainty is part of science. 2013.

83. Lasswell H. The Structure and Function of Communication in Society. The Communication of Ideas. New York: Institute for Religious and Social Studies. p.117. 1948.

84. Patient-Centered Outcomes Research Institute. Patient-Centered Outcomes Research Translation Center 2020 [updated 2020. Available from: https://www.pcori.org/research-results/2016/patient-centered-outcomes-research-translation-center.

85. U.S. Food & Drug Administration. FDA's approval of Veklury (remdesivir) for the treatment of COVID-19 - The Science of Safety and Effectiveness 2020 [updated 2020. Available from: https://www.fda.gov/drugs/drug-safety-and-availability/fdas-approval-veklury-remdesivir-treatment-covid-19-science-safety-and-effectiveness.

86. Institute for Clinical and Economic Review. ICER provides Second Update to Pricing Models for Remdesivir as a Treatment for COVID-19 2020 [Available from: https://icer-review.org/announcements/remdesivir-icer-covid-second-update/.

87. Ramsey AT, Proctor EK, Chambers DA, Garbutt JM, Malone S, Powderly WG, et al. Designing for Accelerated Translation (DART) of Emerging Innovations in Health. Journal of clinical and translational science. 2019;3(2-3):53-8.

88. Abrams EM, Shaker M, Oppenheimer J, Davis RS, Bukstein DA, Greenhawt M. The Challenges and Opportunities for Shared Decision Making Highlighted by COVID-19. The journal of allergy and clinical immunology In practice. 2020;8(8):2474-80.e1.

89. Smolenyak SEMLJAK. Coronavirus misinformation: quantifying sources and themes in the COVID-19 'infodemic'. Cornell Alliance for Science 2020.

90. Ugianskiene A, Davila GW, Su T-H, Urogynecology tF, Committee PF. FIGO review of statements on use of synthetic mesh for pelvic organ prolapse and stress urinary incontinence. International Journal of Gynecology & Obstetrics. 2019;147(2):147-55.