



Methods to Analyze and Present Uncertainty in Economic Evaluation to Support HTA Decisions

Lotte Steuten*, Kednapa Thavorn*, Lauren E. Cipriano, Jason Robert Guertin, Brett McQueen, Gillian Schmidler, Sean Sullivan, Allan Wailoo, Mark Sculpher*

**Denotes membership of the writing group*

Author Affiliations

Lotte Steuten*, PhD, Office of Health Economics, London, UK

Kednapa Thavorn*, BPharm, MPharm, PhD, Ottawa Hospital Research Institute, Ottawa, Ontario, Canada

Lauren E. Cipriano, PhD, Ivey Business School, Schulich School of Medicine & Dentistry, Western University, London, Ontario, Canada

Jason Robert Guertin, PhD, Department of Social and Preventive Medicine, Université Laval, Quebec City, Quebec, Canada

Brett McQueen, PhD, University of Colorado Skaggs School of Pharmacy and Pharmaceutical Sciences, Anschutz Medical Campus, Aurora CO, USA

Gillian Schmidler, PhD, Department of Population Health Sciences and Duke-Margolis Institute for Health Policy, Duke University, Durham NC, USA

Sean Sullivan, BScPharm, MSc, PhD, The Choice Institute, University of Washington, Seattle WA, USA

Allan Wailoo, PhD, School of Medicine and Population Health, Sheffield Centre for Health and Related Research (SCHARR), University of Sheffield, UK

Mark Sculpher*, PhD, Centre for Health Economics, University of York, UK

*Denotes membership of the writing group

HEMA's policies for managing potential conflicts of interest, and individual disclosures for all Working Group members, are available here: <https://hemamethods.org/hema-working-group/>

Table of Contents

Plain-Language Summary	ES1
Executive Summary	ES3
Abbreviations	ES8
Glossary	ES9
1. Introduction	1
1.1. <i>Economic Evaluation for HTA</i>	1
1.2. <i>Uncertainty in Economic Evaluation</i>	2
1.3. <i>Aims and Objectives</i>	4
1.5. <i>Scope of the Report</i>	5
2. Principles of Uncertainty Analysis	6
2.1. <i>Introduction</i>	6
2.2. <i>Uncertainty versus Sensitivity</i>	7
2.3. <i>Informing Decisions</i>	10
2.3.1 <i>How EVPI Informs Decisions</i>	10
2.3.2 <i>Presenting Cost-Effectiveness and EVPI at Population Level</i>	11
2.3.3 <i>Comparing the Population Health Impacts of Access and of Research</i>	11
2.3.4 <i>Price Reductions and Uncertainty</i>	12
2.4. <i>The Concept of Risk Aversion in Decision Making</i>	12
2.5. <i>Issues in Practical Application</i>	14
3. Sources of Uncertainty and Their Characterization	15
3.1. <i>Introduction</i>	15
3.2. <i>Distribution of Uncertainty Across Population Groups</i>	16
3.3. <i>Boundaries of Uncertainty</i>	16
3.4. <i>Model Inputs, Model Structure, and Assumptions</i>	17
3.5. <i>Documenting Uncertainty: The Uncertainty Assessment</i>	20
3.6. <i>Uncertainty by Type of Evidence</i>	21
3.6.1 <i>Treatment Effects and Safety</i>	21

3.6.2 Baseline Risk, Natural History, and Transferability	21
3.6.3 HRQoL and Patient-Relevant Outcomes	22
3.6.4 Resource Use and Costs	23
3.7 Appraisal of the Evidence and the Probabilistic Analysis	24
3.8 Expert Elicitation and Lived-Experience Expertise	25
3.9 Summary	26
4. Relevance of Uncertainty and Methods of Quantification	27
4.1 Introduction	27
4.2 The Purpose of Uncertainty Analysis	27
4.3 Probabilistic Versus Deterministic Analysis	28
4.4 Parameter Uncertainty and PSA in Practice	28
4.5 Bringing Non-Parameterized Uncertainty into the Uncertainty Analysis	29
4.6 Integrating Multiple Sources of Uncertainty in the PSA	30
4.7 Summary	31
5. Linking Uncertainty Analysis to Decision Options	32
5.1 Introduction	32
5.2 The Expanded Decision Menu	33
5.3 Mapping Uncertainty Profiles to Decision Options	35
5.4 Value of Information as Decision Support	36
5.5 Managed Access and CED: What Good Looks Like	37
5.6 Linking the Uncertainty Root Causes to Decision Options	37
5.7 Stylized Examples	39
6. Communicating Uncertainty to Support HTA Decisions	47
6.1 Introduction	47
6.2 Principles for Communicating Uncertainty	48
6.3 Users of Uncertainty Information and a Layered Reporting Approach	49
6.4 Standard Outputs and Their Limitations	52
6.5 Emerging Tools and Presentation Formats	53
6.6 Minimum Reporting Expectations and Committee Deliberation	54
6.7 Communication and Engagement with Patients and the Public	55

6.8 Summary	58
7. Discussion and Recommendations	60
7.1 Key Messages	60
7.2 Further Considerations	61
7.3 Suggestions for Further Work	63
7.4 Recommendations	64
References	67
Appendix A. A Worked Example of Layered Reporting	A1
A.1 Layer 3: Technical Report (excerpt from the uncertainty assessment; Section 3.5)	A1
A.2 Layer 2: The Committee Decision Debrief	A2
A.3 Layer 1: The Public and Decision Summary	A3
Appendix B. How Patient Representatives Informed This Report	B1
B.1 Purpose	B1
B.2 Who Took Part and The Session	B1
B.3 What We Heard	B1
B.4 What Changed in the Report	B2
B.5 What Did Not Change, and Why	B4
B.6 Limitations	B4

Plain-Language Summary

What is this report about?

Organizations, such as the Institute for Clinical and Economic Review (ICER, USA), Canada's Drug Agency (CDA-AMC), and the National Institute for Health and Care Excellence (NICE, England), advise health systems on whether they should pay for new treatments and other health technologies. This process is called health technology assessment, or HTA. These organizations often have to make recommendations before all useful evidence is available. For example, studies may have small numbers of patients, may not follow people for long enough, or may not include patients who are similar to those who will receive the treatment once it is available. This creates uncertainty about how well a treatment works, how long its benefits last, what it costs, and whether its expected health benefits justify the resources needed to provide it.

This report is mainly for HTA organizations and the people who prepare economic evaluations for them. It explains how uncertainty can be identified, analyzed, used to inform decisions, and communicated clearly. It does not tell organizations which treatments to fund or how they should run their decision-making processes.

Why does uncertainty matter?

Health systems have limited resources. Funding one treatment may mean fewer resources are available for other care. HTA organizations therefore consider the potential health benefits of a new treatment alongside what providing it could mean for other patients and services.

Because the evidence is uncertain, a decision that seems right today may look different when more evidence becomes available. A treatment that is funded may provide less benefit than expected. A treatment that is delayed or not funded may later prove beneficial. Either situation can affect patients' health. For some patients, health lost while waiting for an effective treatment cannot be regained. Other patients may also be affected if resources are shifted away from services they need.

The report therefore argues that uncertainty should be considered in terms of the decisions it could affect. Decision makers need to understand not only how uncertain the evidence is, but also how likely it is that another option could be better, what that could mean for patients and health services, and whether collecting more evidence would be worthwhile. This is sometimes called "decision uncertainty."

What does the report recommend?

The report recommends that HTA organizations:

- Clearly report the expected differences in health benefits, costs of a treatment, as well as the incremental cost-effectiveness ratio, and explain how uncertainty could affect whether patients receive access now or whether more evidence is needed first;
- Identify the important sources of uncertainty and explain why they matter;

- Ensure errors or methodological problems are corrected where possible instead of treating them as uncertainty;
- Consider important uncertainties together and clearly explain which uncertainties could not be included in the analysis;
- Use “probabilistic” analysis to estimate how uncertainty could affect the decision, while clearly explaining which uncertainties could not be included;
- Consider both how likely it is another option could be better and what that could mean for patients and health care services;
- Consider options beyond a simple yes-or-no funding decision, such as changing the price, providing access while more evidence is collected, providing access only as part of research;
- Ask for more evidence when it could realistically reduce an important uncertainty and when the potential benefit justifies the cost, delay, and burden of collecting it;
- Present the same underlying results consistently in technical reports, committee materials, and plain-language summaries;
- Involve patients and caregivers when further research is planned; and
- Invest in the people, tools, and resources needed to analyze and communicate uncertainty clearly.

What does this mean for patients and the public?

Information about uncertainty should begin with the recommendation and explain what it means for patients. It should clearly distinguish between uncertainty about whether a treatment benefits patients and uncertainty about whether those benefits justify the resources needed to provide it. Numbers should be presented in familiar terms, and technical language should be explained or replaced with plain language.

Summaries should also explain who was included in the evidence, which groups were not well represented, where evidence is limited, and what additional research could help. They should explain what taking part in that research could require from patients. If access to a treatment could change when new evidence becomes available, patients should be told this from the beginning.

Patients and health systems may view uncertainty differently. For example, someone with a serious condition and few treatment options may be willing to accept more uncertainty about a treatment. The report does not propose changing the economic analysis to reflect individual risk preferences. Instead, it recommends clearly explaining these different perspectives and their possible consequences so they can be considered openly when decisions are made.

How did patients help shape the report?

Seven patient representatives from Canada, the United Kingdom, and the United States took part in a workshop in September 2026. They reviewed parts of the draft, identified wording that was difficult to understand, and highlighted consequences for patients that had not been sufficiently considered. Their feedback informed revisions to the report, including how uncertainty should be communicated to patients and the public.

Executive Summary

The Institute for Clinical and Economic Review (ICER, USA), Canada's Drug Agency (CDA-AMC), and the National Institute for Health and Care Excellence (NICE, England) convened Health Economics Methods Advisory (HEMA) to provide independent guidance on methods for economic evaluation. This second HEMA report considers uncertainty in economic evaluation and how it should inform health technology assessment (HTA) decisions.

Economic evaluations combine evidence on treatment effects, health outcomes, resource use, costs, and other aspects of care to estimate the expected costs and benefits of alternative interventions. These estimates are uncertain for many reasons. Clinical studies use finite samples. Evidence may come from populations, comparators, or settings that differ from those relevant to the HTA decision. Follow-up may be too short to observe long-term outcomes. Studies may be affected by bias, and important evidence may be missing. Economic models also require assumptions about disease, treatment pathways, and future outcomes.

HTA organizations already use several methods to examine uncertainty, including deterministic sensitivity analysis (DSA), probabilistic sensitivity analysis (PSA), scenario analysis, and value of information (VOI) analysis. These methods are useful, but their findings are not always connected clearly to the decisions HTA organizations need to make. A tornado diagram can show which inputs affect model results. A cost-effectiveness acceptability curve can show how the probability that an option is cost-effective changes across thresholds. Scenario analyses can show what happens to cost-effectiveness under alternative assumptions. None of these, on its own, shows the extent to which uncertainty matters to the decision or what should be done about it.

Uncertainty analysis should therefore help answer several practical questions: could uncertainty lead to an incorrect decision about funding an intervention? What is the probability of an incorrect decision? What are the consequences if the option selected using current evidence later proves not to be the best choice? Can uncertainty be reduced? If so, would further evidence be worth obtaining? These questions keep the analysis focused on the purpose of economic evaluation: supporting decisions about how limited health care resources can generate the greatest health benefit.

The importance of a source of uncertainty depends on its effect on the decision in combination with others. An input may be estimated imprecisely and have little influence on the recommendation. Another may be estimated with considerable precision but still provide a poor basis for decision making because the evidence is biased or poorly applicable to the population or setting of interest. Standard errors and confidence intervals describe sampling uncertainty, but they do not show whether an estimate is biased, whether the evidence applies to the decision setting, or whether an important outcome has been omitted. Bias,

limited relevance or transferability, extrapolation beyond observed data, evidence gaps, and assumptions about model structure may all matter.

Some problems should be corrected during the assessment and should not be carried forward as uncertainty. These include coding errors, inadequate reporting, and avoidable methodological shortcomings. Heterogeneity is also distinct from uncertainty. Known differences between identifiable patient groups may be relevant to the decision, while uncertainty may be greater for groups that are poorly represented in the evidence.

A structured uncertainty assessment can help keep these issues visible throughout the evaluation. Important uncertainties should be identified before decisions are made about how to analyze or respond to them. For each source of uncertainty, the assessment should record why it arises, how it affects the model or evidence, and how it has been represented in the analysis.

Where uncertainties can be represented probabilistically, PSA provides the main approach for bringing them together. It produces distributions of expected (mean) costs and health outcomes for each intervention. The means of this distribution provide the best estimate of an incremental cost-effectiveness ratio (ICER). Conditional on a suitable cost-effectiveness threshold, expected cost-effectiveness can also be expressed in terms of expected incremental net benefit (e.g., expected incremental net health benefit), which can also be scaled to the relevant patient population. PSA can also be used to estimate the probability that each option has the highest expected net health benefit, which is a measure of decision uncertainty. PSA can also generate the expected consequences of this uncertainty in health or monetary terms. DSA has a narrower role. It can help analysts understand model behavior, identify influential sources of uncertainty, and detect coding or conceptual problems, but it does not provide an overall measure of decision uncertainty or its consequences.

PSA also has limits. Its results reflect the uncertainties that have been included and how they have been specified. Concerns about bias, transferability, extrapolation, model structure, or missing outcomes usually remain outside the analysis. Greater integration of these sources of uncertainty into PSA is desirable using methods such as model averaging, but it is not practical in every assessment. When uncertainty cannot be parameterized credibly, it is usually examined through scenario analysis. Reports should make clear what sources of uncertainty have been integrated into PSA and which have been reflected using scenario analysis.

Structured expert elicitation can be useful when relevant empirical evidence is unavailable or cannot reasonably be obtained within the time available. It should not substitute data that could be made available and analyzed. Structured elicitation will often involve relevant clinical experts, but patients and caregivers also contribute expertise from lived experience when it is relevant to the quantity or assumption being elicited. Their input may inform how

treatment pathways are represented, including treatment uptake, adherence, discontinuation, and the time or resources required to receive care. It can also help verify model assumptions or identify outcomes that may otherwise be overlooked.

How uncertainty should affect an HTA recommendation depends on the decision context. The same probability of an incorrect decision may have very different consequences across conditions and technologies. The potential for additional research to resolve uncertainty and the cost of undertaking research will also differ. Delaying access may have limited consequences where disease progresses slowly and effective alternatives exist, but it may lead to irreversible health loss for patients with rapidly progressing disease and few treatment options. Permitting immediate access can also create irrecoverable consequences. For example, a new intervention may require high setup costs that cannot be recovered if further research indicates the intervention is not cost-effective.

The available decision options are also broader than a simple choice between recommending and rejecting a technology. Depending on the circumstances, options may include recommendation at the submitted price, immediate access at a lower price, managed access while additional evidence is collected with scheduled reassessment, access only to patients as part of research or no access. The appropriate response depends on the consequences of the decision, whether research can reduce uncertainty, how long further research would take, whether the decision can later be changed, and the cost and feasibility of collecting additional evidence. The burden placed on patients and health services by additional research also matters.

VOI analysis is a way of quantifying the consequences of uncertainty by indicating the population health (or cost) impact that research could, in principle, generate. By also considering the likely cost and feasibility of research, it can determine whether further research is worthwhile. It can estimate the value of reducing uncertainty and identify which uncertainties are most valuable to resolve. Formal VOI is not routinely used in all HTA organizations, but its underlying logic provides a framework for considering the most appropriate decision option. Further evidence is worthwhile when the expected benefit of reducing uncertainty justifies the resources, time, and patient burden involved.

The consequences of an incorrect decision should be considered in both directions. Funding a technology that later proves not to be cost-effective can reduce population health because resources could have produced more health elsewhere. Delaying or denying a technology that later proves cost-effective can also result in health loss. For some patients, that loss cannot be recovered when the decision is revisited.

Price can form part of the response to uncertainty. A lower price may improve expected net benefit and reduce the consequences of an uncertain decision. It does this by reducing the value of information and making further research non-economic, with the implication that the underlying uncertainty in the evidence remains.

Communication is part of uncertainty analysis. Reporting expected cost-effectiveness while placing most of the uncertainty analysis in an appendix can give readers a stronger impression of precision than the evidence supports. The report proposes three levels of reporting: a technical report for analysts and reviewers, a decision brief for the HTA committee, and a plain-language summary for patients and the public. The level of detail can differ, but the results and interpretation should remain consistent.

No single output provides a complete account of uncertainty. For those sources of uncertainty integrated into PSA, cost-effectiveness acceptability curves and a cost-effectiveness acceptability frontier should be presented alongside expected consequences in the form of the expected value of perfect information. Those sources of uncertainty which are assessed using scenario analysis outside a PSA should present the impacts of alternative scenarios on these measures of uncertainty as well as expected cost-effectiveness, such as an incremental cost-effectiveness ratio. Patient-facing materials on uncertainty analysis should begin with the recommendation and what it means for patients. They should distinguish uncertainty about clinical benefit from uncertainty about value for money, and present probabilities in accessible terms. They should also identify which groups were poorly represented in the evidence and explain how this affects confidence in the findings. Where further evidence collection is proposed, the potential burden on patients should be made clear, including additional appointments, travel, and the possibility that access may change over time.

The recommendations below assume that the full range of decision options is available. Where the scope to require or support further research is limited, the range of available options is narrower, which may have negative implications for population health. The report makes the following recommendations:

- Report expected incremental net health benefit alongside the incremental cost-effectiveness ratio, at both individual and population levels, to show the expected health impact of immediate access and the potential value of resolving remaining uncertainty;
- Keep the focus of uncertainty analysis on decision uncertainty and its drivers and consequences to help identify the most appropriate decision option;
- Identify and report important sources of uncertainty, including sampling uncertainty, bias, transferability, extrapolation, model structure, and missing patient-relevant outcomes;
- Correct methodological and reporting problems where possible, rather than treating them as sources of uncertainty;
- Use PSA to quantify joint decision uncertainty by integrating as many sources of uncertainty as possible, while stating clearly which sources of uncertainty remain outside the PSA;
- Communicate PSA results using cost-effectiveness acceptability curves, or where appropriate, cost-effectiveness acceptability frontiers, and report VOI analysis

results, at a minimum, the population expected value of perfect information. When scenario analysis is used outside PSA, the results of PSA should be reported for each scenario;

- HTA processes should consider greater use of methods to integrate all sources of uncertainty into PSA and consider both the probability and the consequences of incorrect decisions based on current evidence. Decision options, including access as part of further research, should be informed by the uncertainty analysis and by whether further research can address and resolve important uncertainty, and whether that research is feasible, and whether its expected value justifies the cost, delay, and burden of obtaining it;
- Present a consistent account of uncertainty across technical reports, committee materials, and public summaries;
- Involve patients and caregivers in plans for further evidence generation;
- HTA organizations should consider investing in capacity building, training, and resources to support the use of the methods and considerations recommended in this report.

Abbreviations

Abbreviation	Term
ART; ARCH	Assessment of Risk Table; Appraisal of Risk Chart
CDA-AMC	Canada's Drug Agency
CEAC	Cost-effectiveness acceptability curve
CEAF	Cost-effectiveness acceptability frontier
CED	Coverage with evidence development
CHEERS	Consolidated Health Economic Evaluation Reporting Standards
DAGWOOD	Directed Acyclic Graphs with Omitted Objects Displayed
DAMWOOD	Decision Analytic Models with Omitted Objects Displayed
DSA	Deterministic sensitivity analysis
EMA	European Medicines Agency
ENBS	Expected net benefit of sampling
EVPI	Expected value of perfect information
EVPII	Expected value of partial perfect information
EVSI	Expected value of sample information
GRADE	Grading of Recommendations Assessment, Development and Evaluation
HEMA	Health Economics Methods Advisory
HTA	Health technology assessment
ICER	Incremental cost-effectiveness ratio; also, in Section 1, the Institute for Clinical and Economic Review
ISPOR	Professional Society for Health Economics and Outcomes Research
NICE	National Institute for Health and Care Excellence
OIR	Only in research
PSA	Probabilistic sensitivity analysis (probabilistic analysis)
QALY	Quality-adjusted life year
ROBVALU	Risk of Bias in studies of Values and Utilities
RWE	Real-world evidence
SAVI	Sheffield Accelerated Value of Information
SHELF	Sheffield Elicitation Framework
TRUST	Transparent Uncertainty Assessment tool
VOI	Value of information

Glossary

This glossary explains terms used in the report in everyday language. The technical definitions are given where each term is first used.

Term	What it means
Bias	A systematic difference between an estimate and the quantity needed for the decision, arising from how a study was designed or conducted or from the people, comparators or setting it covered. Unlike sampling uncertainty, it does not shrink with a larger study.
Cost-effectiveness acceptability curve	A chart that shows how the probability that each option is good value changes as the amount a health system is willing to pay for a unit of health changes.
Cost-effectiveness acceptability frontier	A chart showing which option gives the greatest expected health benefit at each level of willingness to pay, which is not always the option most likely to be good value.
Cost-effectiveness threshold	The benchmark used to judge whether the health gained from a technology is worth what the health system must give up to pay for it.
Cost of uncertainty (expected loss)	The health that is expected to be lost by having to decide on current evidence, combining how likely a wrong choice is with how much would be lost if it were made.
Decision uncertainty	The chance that the option chosen on current evidence is not the one that would give the greatest health benefit.
Deterministic sensitivity analysis	Changing one input at a time to see how much the results move. Useful for checking how a model behaves, but it cannot say how likely a wrong decision is.
Evidential uncertainty	Uncertainty about how well the available evidence supports the model inputs and assumptions used in an economic evaluation.
Expected value of perfect information	The health that could be gained if all the uncertainty in a decision were removed. It places an upper limit on what further research could be worth.
Extrapolation	Projecting what happens beyond the period a study observed, for example, estimating survival over a lifetime from a few years of trial follow-up.
Heterogeneity	Real, explainable differences between identifiable groups of patients, such as differences by age or disease severity. This reflects differences between people.
Incremental cost-effectiveness ratio	The extra cost of one option compared with another, divided by the extra health it produces.

Term	What it means
Incremental net health benefit	The difference in net health benefit between an intervention and its comparator. It represents the additional health gained from the intervention after accounting for the health opportunity cost of any additional resources required.
Indirect treatment comparison	A way of comparing two treatments that have not been tested against each other directly, by linking them through studies of a common comparator.
Irrecoverable costs	Money or resources committed that cannot be recovered if the decision is later reversed, such as equipment, training or set-up costs.
Managed access and coverage with evidence development	Funding a technology for a limited period while agreed evidence is collected, with the decision revisited once that evidence is available.
Model (decision-analytic model)	A structured representation of a disease, its treatments and their consequences, used to estimate the costs and health outcomes of each option over time.
Model structure	How a model represents the decision problem: the health states people move between, the care pathways, and the assumed relationships between inputs.
Net health benefit	The health gained from an option after subtracting the health expected to be lost elsewhere because the money is spent here.
Only in research	Funding a technology solely for patients taking part in a study, so that evidence is generated before wider access is decided.
Opportunity cost	The health that could have been produced if the same resources had been used for other care.
Parameter (model input)	A number used in a model, such as a treatment effect, an event rate, a cost or a utility value.
Parameterized	Represented in the model as a number with a range of plausible values, so that the uncertainty about it can be included in the probabilistic analysis.
Probabilistic (sensitivity) analysis	Running a model many times, each time drawing the inputs from their plausible ranges together, to show the range of possible results and how likely each option is to be the best choice.
Quality-adjusted life year	A measure that combines how long people live with how well they live, so that a year in full health counts as one and a year in poorer health counts as less.
Real-world evidence	Evidence from routine care rather than a clinical trial, such as registries, health records or administrative data.

Term	What it means
Relevance and transferability	How closely the people, treatments, comparators and setting in a study match those of the decision being made, and whether its findings can reasonably be applied to them.
Sampling uncertainty	The uncertainty that arises because a study observes a limited number of people. Results would differ somewhat if the study were repeated, and standard errors and confidence intervals describe that variation.
Scenario analysis	Re-running the analysis under a different assumption, such as an alternative way of projecting survival, to see whether the conclusion changes.
Structured expert elicitation	A formal process for asking experts, who may include people with lived experience, for their judgment about a quantity that evidence does not provide, and for how certain they are.
Uncertainty assessment	A table recording what is uncertain in an evaluation, why it arises, how it has been handled in the analysis, and whether further evidence could reduce it.
Utility (health-state value)	A number representing how good or bad a state of health is judged to be, used to calculate QALYs.
Value of information analysis	Methods that estimate how much health could be gained by collecting further evidence, so that the value of research can be compared with its cost.
Variability	Differences in outcomes between individual patients that cannot be explained or predicted based on baseline knowledge, as distinct from uncertainty about the average effect.

1. Introduction

In brief. Economic evaluation helps health systems decide whether the health benefits of a new intervention justify the resources needed to provide it. These decisions use evidence about how well a treatment works, its effects on health, its costs, and how care is delivered. All evidence has some uncertainty. For example, studies may be small, may not include patients or health care settings that closely match those affected by the decision, or may not follow patients long enough to see important outcomes. Studies can also have limitations that affect their results, and economic models must make assumptions about what may happen beyond the time covered by available studies.

HTA organizations use several approaches to examine uncertainty. Depending on what they find, they may consider changes in price, collecting more evidence, or placing other conditions on access. However, uncertainty analysis is not always clearly connected to these decisions. This report examines how uncertainty can be identified, analyzed, and communicated so decision makers can better understand why it matters, what it could mean for patients and health services, and what options are available.

1.1. Economic Evaluation for HTA

In the field of health care, economic evaluation is generally focused on supporting health systems to improve population health using the financial resources available, which are typically limited by budgets or other constraints. In the context of health technology assessment (HTA), these methods are applied to inform recommendations about whether new medicines and other technologies are cost-effective. This means that the additional health benefits they offer to those receiving the interventions are not outweighed by any health opportunity costs. In simple terms, therefore, economic evaluation is about checking that the benefits to the gainers are at least as great as the “disbenefits” to the losers. If that is not the case, investment in the new intervention will reduce population health. As shown in the Health Economics Methods Advisory’s (HEMA’s) first report (1), how economic evaluation defines “benefit” can be extended; for example, to include the impact of changes in health outcome on productivity. However, the same challenge of comparing the benefits gained from a new intervention to its opportunity costs still exists and should be at the heart of how these methods are used in HTA.

In estimating the costs and health outcomes of new interventions compared to existing alternatives, economic evaluation uses different types of information which are typically drawn from a range of sources. Key evidence relates to the comparative clinical effectiveness of the intervention (intended and unintended – “efficacy” and “side effects”), which is generally taken from clinical studies, ideally randomized controlled trials (RCTs).

To express these clinical effects as changes in health outcomes, as perceived by patients and which are comparable across different clinical areas, health benefit is generally quantified in terms of a composite measure such as a quality-adjusted life year (QALY), disability-adjusted life year (DALY), equal-value life year (evLY) or health years in total (HYT). As in HEMA's first report (1), here we use the term "lifetime health gain" to cover all these general measures. To quantify the impact of a new intervention using one of these metrics requires evidence on mortality risks and a suitable health-related quality of life (HRQoL) measure, which can be sourced from RCTs or other clinical studies.

Depending on the intervention, economic evaluations may require evidence on the absolute risk of clinical events with existing interventions over time, the long-term natural history of disease, treatment discontinuation, changes in treatment effectiveness over time, and the accuracy of diagnostic and screening tests. These estimates may come from clinical and epidemiological studies using a range of designs, or from routinely collected health system data. Estimating costs also requires evidence on resource use, such as clinic visits, days in hospital, medicines, and diagnostic tests. These data may be sourced from clinical studies, surveys or routine data. The resource used must then be valued using appropriate unit costs or prices, which should be specific to the setting and jurisdiction in which the decision is being made.

Economic evaluation is generally used to support decisions about how limited health care resources can be used to improve population health. Consistent with this objective, the analysis typically focuses on the *expected* or *mean* health effects and costs of alternative interventions for a specific population. Cost-effectiveness can be expressed using an incremental cost-effectiveness ratio (ICER) or the expected incremental net health benefit, which represents the health gained after accounting for the health opportunity cost associated with additional spending. Where an appropriate cost-effectiveness threshold is available, an intervention with an incremental cost-effectiveness ratio below the threshold is generally considered cost-effective. If the threshold is a measure of the health system's opportunity cost, this is equivalent to a non-zero expected net health benefit. The maximum price at which a new technology would remain cost-effective (sometimes called a "value-based price") is the price at which the incremental cost-effectiveness ratio equals the threshold, and the expected net health benefit is zero.

1.2. Uncertainty in Economic Evaluation

Establishing whether a new intervention is cost-effective *on average* cannot ignore the fact that the mean inputs (clinical efficacy, costs, etc.) are estimated with uncertainty. In general, the less evidence that exists and the lower the quality and relevance of that evidence, the greater the uncertainty in cost-effectiveness and the appropriate funding decision. This second HEMA report focuses on how this uncertainty should be reflected in economic evaluation to inform the decision options available to HTA organizations.

A range of methods can quantify uncertainty in economic evaluation to identify which types of uncertainty most influence expected incremental net health benefit and to present these to decision makers. These approaches are reflected in the methods guidelines issued by HTA organizations. However, translating such uncertainty analyses into better decisions has proven more of a challenge. Different sources of uncertainty in economic evaluation can generate *decision uncertainty*. That is, the estimated expected (or mean) net health benefits can result in two forms of incorrect decisions with undesirable consequences: (i) agreeing to fund a new intervention based on a non-negative expected incremental net health benefit that ultimately turns out to generate less benefit for its recipients than its additional costs impose on other patients, with a consequent reduction in population health; and (ii) deciding not to fund a new intervention because of its negative expected incremental net health benefit, which turns out to be cost-effective so that funding it would have increased population health. Uncertainty can be expressed as the likelihood of making either of these wrong decisions but understanding the consequences in terms of forgone health is also important.

Approaches to characterizing uncertainty in economic evaluation can be complex, and HTA organizations often place considerable emphasis on assessing uncertainty in economic evaluation on the incremental cost-effectiveness ratio or expected incremental net health benefits associated with the “most plausible” single set of assumptions and evidence. This does not directly consider the extent of the uncertainty or its consequences because it effectively assumes these are the only correct assumptions and evidence, which imposes the assumption that the likelihood of others being correct is zero. Even when a new intervention is cost-effective *on average*, its value to a health system is negatively affected by uncertainty. That is, when there is high decision uncertainty, the value of the intervention to the health system is less than when decision uncertainty is low. An extreme example would be that when the evidence for a new treatment for a common disease is based on averaging costs and benefits over 10 patients, a positive expected net health benefit cannot hide the weakness of the evidence. That uncertainty should therefore be made explicit and considered alongside the expected cost-effectiveness results.

The challenge is for HTA organizations to identify the levers they have at their disposal in their decision making or guidance to reflect directly the magnitude and consequences of the uncertainty in terms of population health. One approach has been to seek price reductions in interventions in the face of uncertainty. For example, Medicaid has proposed a requirement for higher rebates for medicines that have received accelerated regulatory approval, which generally results in less mature evidence and uncertainty (2). Another is NICE’s use of a lower cost-effectiveness threshold when it considers uncertainty to be significant and where additional evidence could have been generated which is reflected in its methods manual (3) and specific appraisals (4). Another lever HTA organizations have used when facing uncertainty is agreements with manufacturers to pay based on treatment outcomes — that is, health systems pay less or nothing when patients’ treatments are unsuccessful. These arrangements — sometimes called risk sharing schemes, outcomes-

based agreements, or “payment by results” — are effectively reducing the price of a new technology (5).

As additional research can, in principle, reduce uncertainty, a third lever HTA organizations can use in the face of uncertain evidence is to make funding available (often with price negotiation), but to complement this with an agreement that the manufacturer will undertake additional research to address important uncertainties, after which the intervention may be re-evaluated. NICE’s managed access agreements are an example of such an arrangement. In some situations, HTA organizations may decide that there is value in additional research for an intervention (and may seek to facilitate this) but not with routine funding — sometimes called an “only in research” decision — which, depending on the evidence generated, may require re-evaluation.

The problem with these existing HTA responses to uncertainty is, in general, their limited association with the economic evaluation that guided the decision and the quantification of uncertainty in those studies. Although uncertainty in the evidence may be described and emphasized in documenting HTA decisions, the scope for economic evaluation to quantify decision uncertainty, the consequences for population health and the potential value of research is rarely used. Indeed, some of this information is requested by HTA organizations — for example, the use of probabilistic sensitivity analysis — but there are few examples of where this is used directly to inform decisions and to specify responses to address uncertainty. Methods have been suggested to augment standard economic evaluation approaches to identify suitable responses in the form, for example, of additional evidence development (6). However, these have had limited impact on the types of analyses submitted to HTA organizations and the way decision makers have used those studies. This may be due to their perceived complexity and the associated burden of analysis, rejection of some aspects of the methods, a wish to retain flexibility in decision making, or other considerations.

1.3. Aims and Objectives

The aim of this report is to review available methods to characterize uncertainty and present it to decision makers with an emphasis on guiding decisions.

Specific objectives are:

- To taxonomize the sources of uncertainty in economic evaluation;
- To summarize key methods available to characterize uncertainty in economic evaluation and its consequences for population health;
- To consider the challenges for a fully integrated analysis of uncertainty and potential ways to address them;
- To describe proposed tools and frameworks which can be used with a fully integrated analysis of uncertainty;

- To propose “pragmatic” approaches to respond to uncertainty when economic evaluation does not provide a fully integrated approach to uncertainty;
- To consider methods for communicating uncertainty analysis to stakeholders;
- To make recommendations regarding best practice in economic evaluation to support decisions under uncertainty and regarding research.

1.5. Scope of the Report

Several methodological issues are closely related to uncertainty in economic evaluation and HTA but beyond the scope of this report. These include:

- Estimation of opportunity cost and the cost-effectiveness thresholds used in economic evaluation studies;
- Methods for identifying and analyzing heterogeneity and best methods for implementing the latter in economic evaluation. One approach HTA organizations may take to uncertainty is to change the decision problem, for example by looking at the cost-effectiveness of an intervention only in a sub-group of patients. This report does not view this as uncertainty analysis, but as a means of exploiting heterogeneity; even within the subgroup, uncertainty analysis remains important;
- A detailed review of recent technical developments and enhancements to uncertainty analysis; and
- The detailed design and implementation of policy responses intended to manage uncertainty for specific technologies. The report considers how uncertainty analysis can inform such responses, but not their operational implementation.

The remainder of the report is structured as follows. [Section 2](#) outlines the key principles of decision making under uncertainty. [Section 3](#) considers the main sources of uncertainty and how they can be identified and characterized. [Section 4](#) sets out the methods of analyzing or quantifying uncertainty and its consequences. [Section 5](#) looks at how uncertainty analysis can inform the decisions available to HTA organizations. [Section 6](#) considers the communication of uncertainty to different audiences. [Section 7](#) draws together the main implications and offers recommendations and sets out priorities for further methodological development.

2. Principles of Uncertainty Analysis

In brief. Economic evaluations help decision makers consider whether the health benefits of a new treatment justify the resources needed to provide it. Because evidence is never complete, these decisions always involve some uncertainty. This section explains how to look beyond the single best estimate and consider how likely it is that another option could be better, what that could mean for patients and health services, and whether collecting more evidence would be worthwhile.

Methods that bring together uncertainty from different parts of the analysis can show how confident decision makers can be and what possible health consequences could be. This information can help them weigh the benefits of giving patients access now against the value, cost, delay, and patient burden of collecting more evidence. It can also inform decisions about prices. Some important uncertainties are difficult to include in numerical analysis. These still need to be clearly described and considered in other ways.

2.1. Introduction

This section sets out some key building blocks of uncertainty analysis and ways in which they can inform decision making. As mentioned in [Section 1](#), the way economic evaluation is designed, implemented and communicated to inform decisions in HTA reflects an underlying health system goal of maximizing the relevant objective (here assumed to be population health) subject to financial and other constraints. This is consistent with estimating an incremental cost-effectiveness ratio and comparing it with a threshold representing opportunity cost or presenting this as an expected incremental net health benefit (7, 8).

This focus on supporting decision making contrasts economic evaluation from other areas of health research in terms of how the uncertainty reflected in statistical analyses is handled. In most clinical research, statements about how uncertainty in evidence is communicated use the conventional methods of statistical inference and the language of “statistical significance” and “p-values”. Determining whether evidence is sufficiently compelling to reject a “null hypothesis” of, for example, no difference in efficacy between two interventions relies on judgments about levels of proof that are inconsistent with decisions about resource allocation with a view to improving population health (9). Instead, decision analysis methods provide a more suitable framework for assessing how uncertain evidence should inform resource-allocation decisions in HTA (10). This framework underpins the material presented in this section.

2.2. Uncertainty versus Sensitivity

Two approaches to uncertainty analysis have commonly been used in economic evaluation, but the terminology sometimes fails to communicate key differences between them. The first has been used routinely from the earliest studies in the field of health and focuses on how the measure of cost-effectiveness (e.g., the incremental cost-effectiveness ratio) varies as changes are made in the numerical value of one or more uncertain parameters (i.e., how sensitive cost-effectiveness is to variation in specific parameters). For example, such analyses can show how an incremental cost-effectiveness ratio for a new intervention changes as the mean length of hospital stay for patients undergoing the comparator intervention is varied. This is a deterministic *sensitivity analysis (DSA)* which can help decisions if, for example, the incremental cost-effectiveness ratio suggests the new intervention remains cost-effective for a range of “plausible” values of the mean input parameter — that is, it is *robust* to this source of uncertainty. A variant of DSA to assess decision robustness under uncertainty is threshold analysis, which, in this example, might show how low the comparator's mean length of hospital stay can be for the new intervention to remain cost-effective.

Most economic evaluations have many input parameters that are subject to uncertainty, so DSA of this type would need to be undertaken on each of these. Using DSA, it is possible to show how cost-effectiveness results change when more than one uncertain mean parameters is varied, but this is challenging for more than three parameters. There are several limitations of DSA, including the generally incomplete and arbitrary basis of how the range of variation in uncertain parameters is chosen, that it ignores the fact that values across that range are not equally likely to be correct, it provides little guide to decisions when robustness cannot be demonstrated, the impossibility of handling DSA simultaneously for all parameters and the lack of a metric to communicate the level of overall uncertainty and its consequences. DSA can provide some insights into robustness and those sources of uncertainty that are likely to “drive” cost-effectiveness, but this is partial and does not quantify decision uncertainty or the consequences of uncertainty.

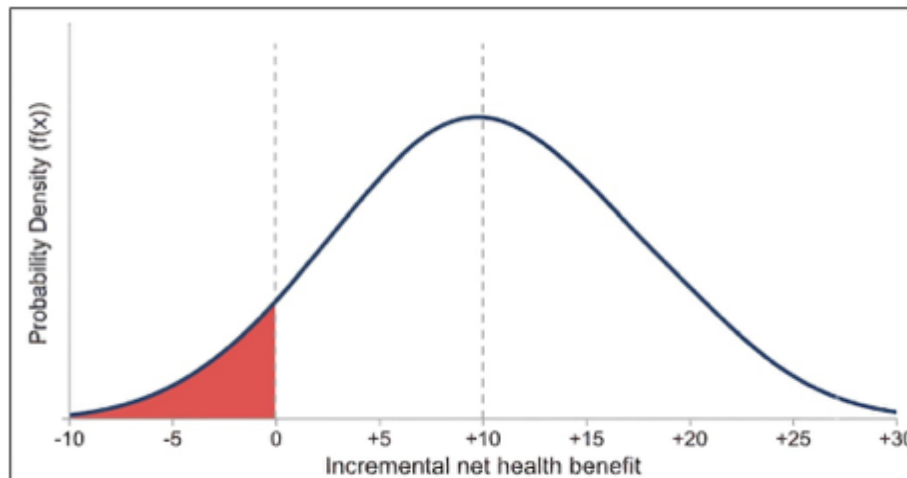


Figure 2.1. A probability density function showing the uncertainty in expected incremental net health benefit. The mean value is at 10 (cost-effective), but a proportion of the distribution (shaded area) is less than 0.

The second approach — which can more accurately be called *uncertainty analysis* — seeks to reflect the combined uncertainty of all parameters as a distribution of possible cost-effectiveness results. This is generally implemented using probabilistic sensitivity analysis (PSA) which involves characterizing uncertainty in each input parameter as a distribution of possible mean values. These are based on available evidence and reflect the characteristics of each parameter (e.g., the fact that the value of a probability lies between 0 and 1). PSA involves effectively running the model many times to generate a distribution of cost-effectiveness results. Figure 2.1 shows a simple way of presenting uncertainty in incremental expected net health benefits, based on PSA, using a probability density function. The mean value of this distribution represents overall expected cost-effectiveness (shown with the dotted line at 10), but the distribution also shows that other values are possible, with the height of the distribution at any given value indicating the likelihood of that value being correct. The tails of the distribution show markedly lower and higher estimates than the expected value but have low likelihoods of being correct.

Figure 2.1 is consistent with positive overall expected net health benefits (the intervention is cost-effective on average), but some proportion of the density is less than zero (not cost-effective), which is shaded in red. This distribution can provide important information to support decision making. Assume that the decision is taken to fund this intervention given that the expected incremental net health benefit is positive. The total area of the distribution equals 1, but the proportion of that area showing a negative incremental net health benefit (shaded red) can be interpreted as the probability that a decision to fund would be incorrect. This is known as decision uncertainty: the probability of a wrong decision which, in this case, represents the probability that the intervention is not cost-effective although the expectation (mean) is positive.

Figure 2.2 shows a new intervention which, based on expected net health benefits, is not cost-effective (shown as a mean at -10). This would suggest that the appropriate decision would be not to fund this intervention. The shaded area under the function above 0 represents the decision uncertainty associated with that decision — that is, the probability that the intervention would actually generate a positive net health benefit and, by not funding it, patients who would benefit from using it are not given access, with a negative impact on population health.

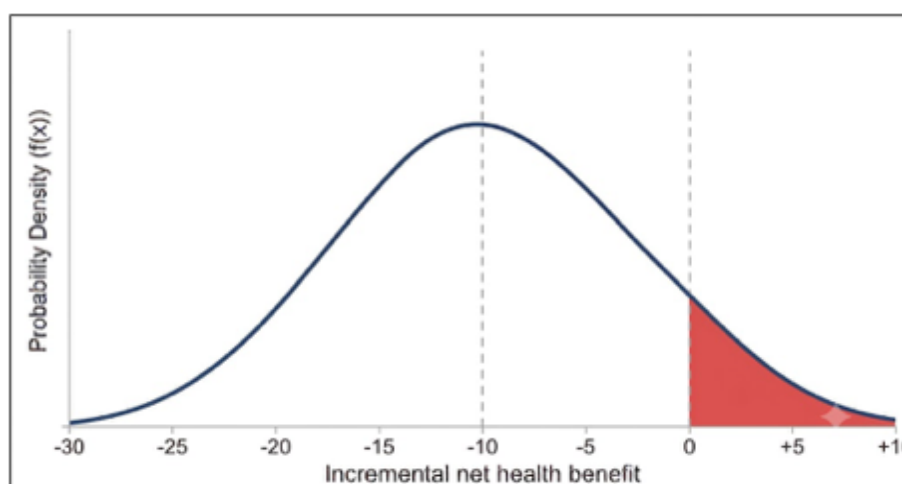


Figure 2.2. A probability density function showing the uncertainty in expected incremental net health benefit. The mean value is at -10 (not cost-effective on average) but a proportion of the distribution (shaded area) is greater than 0.

This distribution can also communicate something about the consequences of a wrong decision, expressed in terms of health. The distributions in Figures 2.1 and 2.2 include the full range of values of expected incremental net health benefit that have some probability of being correct. These can be interpreted as a very large number of possibilities which are consistent with the evidence, but which vary in their likelihood of being the “true” value.

Table 2.1 shows a simplified version of this which, rather than a very large number of possibilities, includes only 5. For each, the expected net health benefit for two treatments (A and B) are presented. The “best” choice of treatment is the one with the highest expected net health benefit. As for the two distributions in Figures 2.1 and 2.2, the best estimate of cost-effectiveness based on existing evidence is the average of all the possibilities which is shown in the final line and suggests that Treatment B is the more cost-effective and may be considered suitable for funding, with an incremental expected net health benefit of 1.0. The 4th column shows that, across the 5 possibilities shown, B is the best option in only 3/5 (60%) cases. Therefore, 1 minus this value (0.4) can be interpreted as the decision uncertainty associated with a positive funding decision: the probability that a decision to fund is wrong and will reduce population health.

The final column of Table 2.1 shows how the consequences of the decision uncertainty (which can also be thought of as the *cost of uncertainty*) are calculated. For those possibilities which turn out to indicate that Treatment B is more cost-effective, the decision to fund is correct, which accords with the decision based on the overall average. In Possibilities 2 and 4, however, the true values suggest Treatment A is more cost-effective and that funding it would have generated more health. Averaging the values in Column 5 (as it is unknown in advance which of these possibilities is the correct one) indicates the expected net health benefit if no uncertainty existed and it was always possible to pick the best treatment from the range of possibilities. Subtracting the expected net health benefit with uncertain evidence (13) from the expected net health benefit if there was no uncertainty in Column 5 (13.6) represents the cost of uncertainty (0.6); that is, the consequence for population health of having to make the decision under uncertainty.

This measure of the cost of uncertainty is also interpreted as the expected value of perfect information (EVPI). That is, if research was possible to remove all the many sources of uncertainty from the economic evaluation, this would be the gain in the expected net health benefit. Of course, research can never be undertaken that removes all aspects of uncertainty: even if such research could be designed, it may not be affordable. So EVPI can be seen as the maximum value of further research in relation to a given decision problem.

Table 2.1. A Simplified Version of Figures 2.1 and 2.2 Showing just 5 Possibilities from a Distribution for Two Treatments, A and B (Claxton, 2008) (11) .

How things could turn out	Net health benefit			Best we could do if we knew
	treatment A	treatment B	best choice	
Possibility 1	9	12	B	12
Possibility 2	12	10	A	12
Possibility 3	14	17	B	17
Possibility 4	11	10	A	11
Possibility 5	14	16	B	16
Average	12	13		13.6

2.3. Informing Decisions

2.3.1 How EVPI Informs Decisions

These analytical steps, including EVPI, provide building blocks to guide decisions about funding new interventions under uncertainty, which are built on in subsequent sections of this report. A starting point is that the expected incremental net health benefit of a technology represents the best estimate, *given available evidence*, of the benefit current patients will gain from immediate access to an intervention, net of opportunity costs. In Table 2.1, this is 1.0 unit of lifetime health gain from Treatment B. In principle, however, the manufacturer of the new technology could have undertaken research before HTA submission which would have reduced the uncertainty faced by health systems. As such,

manufacturers' decisions about the development of the technology, including the types and extent of research to show its effectiveness and cost-effectiveness, can have direct impacts on population health. The calculation of EVPI provides a measure of the maximum value of the additional research that could have been undertaken by the manufacturer prior to HTA submission. Research consistent with perfect information would never have been undertaken in practice for a range of reasons, including its infeasible cost, but it provides an upper bound of the consequences of the uncertainty implied by manufacturers' research and development decisions. In the simple analysis in Table 2.1, this is 0.6 units of lifetime health gain.

2.3.2 Presenting Cost-Effectiveness and EVPI at Population Level

The results of economic evaluations are conventionally presented for the average patient. Because the ultimate impact of decisions is at the population level, a valuable starting point for decision makers is to present results at the population level to understand the magnitude of uncertainty and its implications for decisions. The expected impact of immediate access to a new intervention on expected incremental net health benefit scaled to represent the population of current patients can then be compared to the expected value of future research at the population level. Together with some assessment of how long the research will be valuable, the latter can also be seen as the maximum value of potential research which could, in principle, be undertaken in the future. This research could be in many forms, including further follow-up in existing clinical studies and entirely new studies.

2.3.3 Comparing the Population Health Impacts of Access and of Research

In situations where the expected impact on population health of uncertainty (research which could be undertaken) is high compared to the expected impact on population health of immediate release, decision makers may need to consider whether funding the new technology can be linked to a requirement for additional research. In other words, HTA decisions or guidance can broaden to include agreements with manufacturers on generating additional evidence on a new technology alongside its funding.

The feasibility and value of that future research would involve additional considerations, including:

- identifying the key uncertainties in the economic evaluation based on existing evidence;
- establishing whether these areas of uncertainty can be reduced through research;
- considering how long any feasible research activities would take to report;
- assessing whether the cost of undertaking research would be less than the EVPI and acceptable given the specific feasible research possibilities;
- considering whether irreversible costs would be incurred in offering immediate funding which would not be recoverable (sunk costs) if future research indicated that the new intervention should not be funded.

2.3.4 Price Reductions and Uncertainty

The price of the new intervention directly impacts its expected incremental net benefit. Indeed, if the incremental cost-effectiveness ratio of an intervention is equal to the opportunity cost threshold (the “value-based price”), the expected incremental net health benefit is zero. When there is little uncertainty regarding whether the intervention is effective relative to its comparators, but uncertainty exists in the magnitude of benefit, its price will also directly impact the cost of uncertainty/EVPI. As mentioned in [Section 1](#), decision makers will often seek to secure lower prices when faced with uncertainty about the cost-effectiveness of a new intervention. The principles outlined here can link the results of an economic evaluation and its uncertainty analysis to the magnitude of potential price reduction. As the price of the intervention reduces, the expected incremental net health benefit of immediate access increases, and the cost of uncertainty and the value of feasible research falls. As such, these analyses provide a framework giving decision makers and manufacturers an opportunity to consider the trade-offs between reducing an intervention’s price and/or conducting additional research to reduce uncertainty to facilitate immediate access and funding.

In some circumstances, there may be no agreement on price and future research, in which case HTA organizations might decide not to fund the technology. Such a decision may be justified even when the expected incremental net health benefit of the new intervention is positive, based on existing evidence. For example, withholding funding may create incentives for all manufacturers to reduce the cost of uncertainty associated with future products by doing more extensive research prior to HTA submissions. This may be more compelling for larger markets or if there were similar responses across countries. However, this decision involves a clear trade-off. Current patients may lose access to a technology that is expected to provide health benefits, and this loss may be irreversible if the intervention would only be effective earlier in their disease course. At the same time, the decision may influence future evidence generation, the cost of uncertainty associated with other interventions, and the health benefits for future patients.

The considerations necessary to address these potential trade-offs, and the nuances associated with the decision options, are considered in [Section 5](#) of the report.

2.4. The Concept of Risk Aversion in Decision Making

How economic evaluation can better guide decisions while reflecting uncertainty has long been debated. This relates to whether decision makers can be or should be risk averse (12, 13). In this context, risk aversion means that decision makers are not assumed to focus on enhancing population health from available health system resources. Instead, they are assumed to maximize something different (technically, a “collective utility function”) which places an additional penalty on uncertainty arising from limitations in the evidence base (i.e., evidential uncertainty), including concerns about whether the evidence is sufficiently credible, direct, and applicable to the decision.

Introducing risk aversion has important implications for pricing as the maximum acceptable price or value-based price is no longer the one at which the incremental cost-effectiveness ratio equals a cost-effectiveness threshold based on opportunity cost. Rather, the acceptable price would be lower, with the size of the reduction depending on the degree of decision makers' risk aversion and the extent of uncertainty in evidence (12).

Bringing risk aversion into economic evaluation to address evidential uncertainty is usually justified in terms of the preferences of decision makers. Even if this is a suitable objective, the literature has long debated whether decision makers *should* be risk averse given their remit. Arrow and Lind's seminal work argued that risk neutrality would be appropriate if, as with HTA, numerous decisions are spread over many people across interventions which have independent costs and outcomes (14). However, over time, this position has been critiqued based on arguments including whether investments in services are independent of aggregate consumption and macroeconomic risk (15). Much of the debate about risk aversion in public funding decisions has related to the appropriate discount rate (16).

However, this debate has taken place assuming decision makers' objectives are rather different. As discussed in HEMA's earlier report on defining benefits in economic evaluation (1), a focus on individual preferences rather than socially valuable attributes of health care (such as improved health) is a normative judgment rather than a technical one. Health systems and HTA organizations are given objectives and face constraints, and these have generally shaped the methods used in economic evaluation for HTA (17), with a focus on how a new technology impacts net population health. Allowing an additional layer of decision-maker risk aversion to be applied, which would effectively reduce the "value-based price" given uncertain evidence, does not remove the need also to consider the costs of uncertainty in terms of potentially forgone research opportunities. Furthermore, with a focus on impacts on population health, it is hard to see how a "risk aversion factor" could be conceptualized and quantified, and how HTA organizations would make themselves accountable for such judgments.

Importantly, the concept of risk aversion on the part of decision makers is distinct from risk aversion in patients or the public about health, which was discussed in HEMA's first report (1). The key consideration is whether the measure of health benefit (lifetime health gain) should reflect the risk preferences of those whose preferences constitute the HRQoL weights and, if so, how this should be implemented. In that context, the potential for individuals to exhibit risk aversion relates to their preferences about the risks of potential health states associated with a given disease with or without interventions. This is not uncertainty in the sense discussed in this report, which relates to populations. Rather, it relates to the risks individuals might face or imagine — sometimes called *variability*. The concepts are related in that the greater the uncertainty in evidence, the higher the variability facing individuals regarding the health states they may experience. However, even if risk preferences are incorporated into elicitation exercises to generate HRQoL weights, this does not address the implications for population health in terms of potentially valuable research which is forgone.

2.5. Issues in Practical Application

Economic evaluation can support decision making by doing more than estimating the expected incremental net health benefit per patient of interventions. It can also estimate the magnitude of the uncertainty associated with the decisions, and the potential population health consequences of that uncertainty, including the value of further research that might be forgone if decisions are made immediately. These elements provide building blocks for better decision making. This framework has been recognized for several years in published methods and applied research (6). It is also reflected, to some extent, in HTA methods guidelines — for example, supporting the use of PSA to translate parameter uncertainty into a distribution of expected cost-effectiveness estimates. However, economic evaluation has not been routinely used to link uncertainty to decisions about suitable prices and feasible research requirements.

Several reasons may explain this. One is that not all uncertainty can be easily represented in a PSA, so some elements of evidential uncertainty may not be fully reflected in a distribution of expected incremental net health benefit. In practice, funding decisions often involve price negotiation in the face of significant evidential uncertainty, and may include requirements for additional research, but the details of such agreements are rarely shaped by the uncertainty analysis in economic evaluation. The remainder of the report considers how to address this disconnect.

3. Sources of Uncertainty and Their Characterization

In brief. Uncertainty in economic evaluation can arise for many reasons. Study results may not be precise, may be affected by bias, or may not fully apply to the patients, treatments, or health care setting involved in the decision. Long-term outcomes may need to be estimated beyond the available evidence, and important information may be missing. Economic models also require assumptions about how a disease, treatment pathways, and outcomes are represented. These uncertainties should be clearly identified and documented, including how they have been addressed, whether they could affect the decision, and whether more evidence could reduce them. Problems that should be corrected, such as coding errors, unclear reporting, or avoidable problems with methods, should be fixed rather than treated as uncertainty.

3.1. Introduction

This section addresses uncertainty in model-based economic evaluation, the approach used by most HTA organizations to inform funding recommendations. The principles also apply to evidence used in other forms of economic analysis. We describe uncertainty according to its source: sampling uncertainty, bias, limited relevance of the evidence to the decision setting, extrapolation beyond observed data, evidence gaps that require assumptions, and choices about how the model represents the disease, treatment pathways, and decision problem. In principle, any of these sources could be parameterized and incorporated into a probabilistic analysis. In practice, some are routinely parameterized and others rarely are ([Section 4](#)).

Two further concepts are used throughout this report. *Sampling uncertainty* arises because studies observe finite samples: estimates would vary if the study were repeated in another sample, and standard errors and confidence intervals describe this variation. It affects every empirical estimate, including estimates from well-conducted and directly relevant studies, and is usually straightforward to parameterize. *Bias* is a systematic difference between an estimate and the quantity needed for the decision. It may arise from the design or conduct of a study, or because its population, comparators, outcomes, or setting do not adequately reflect the decision problem. Sampling uncertainty affects precision, whereas bias affects validity. An estimate can therefore be precise but biased, and a wide confidence interval does not itself indicate bias ([Section 3.7](#)). Other factors, including measurement error and heterogeneity between studies, can also affect precision. *Decision uncertainty* describes the combined implications of these evidence limitations for the choice between available options, including the risk that the option selected is not the one that would maximize expected net benefit ([Section 2](#)) (18).

Heterogeneity is distinct from uncertainty. It refers to systematic and explainable differences between identifiable subgroups. Methods for modeling heterogeneity are outside the scope of this report ([Section 1.5](#)), although heterogeneity and uncertainty can interact. Reports should make clear whether observed variation reflects heterogeneity, uncertainty, or both (19).

3.2 Distribution of Uncertainty Across Population Groups

Much of the evidence informing economic evaluations is derived from clinical trials, whose participants may differ from patients treated in routine practice in age, sex, race and ethnicity, comorbidity, and care setting. Older adults (20) and people with multiple conditions (21), for example, are often underrepresented in trials. Similarly, the racial and ethnic composition of trials supporting recent cancer drug approvals has differed substantially from that of the populations receiving these treatments (22). When particular groups are poorly represented, estimates for these populations are based on less direct evidence and may rely on assumptions about whether outcomes observed in trial participants are applicable to them.

Routinely collected data may help address these evidence gaps, but they can introduce other limitations. People who use health care services less often, including those who avoid care because of previous negative experiences, may be less well represented in health records. Analyses based on these data may therefore provide an incomplete picture of their health care experiences and needs (23, 24). For example, a UK parliamentary inquiry into sickle cell care received evidence that some patients avoided hospital because of previous experiences of care (25).

Reports should therefore identify which groups are poorly represented in the evidence and what has been assumed about outcomes for them, without assuming the direction or size of any difference.

3.3 Boundaries of Uncertainty

Four types of problems are frequently described as uncertainty in HTA practice but fall outside the definition used in this report, and each requires a different response. A transparency problem arises when an analysis cannot be followed or reproduced because information has not been reported, and the appropriate response is more complete reporting. Methodological shortcomings, such as coding errors or unjustified simplifications, are flaws in the analysis that should be corrected during assessment rather than carried forward through conservative assumptions, price adjustments, or general caution (26). Flaws in the design or conduct of the underlying clinical studies should be identified in the clinical review; uncertainty analysis can indicate how much such flaws might matter for the recommendation, but it cannot correct them. Finally, affordability and budget impact are relevant to HTA decisions but do not represent uncertainty about cost-effectiveness.

Any shortcoming that cannot be corrected in the time available should be recorded for deliberation, together with its likely direction of effect, rather than absorbed into an assumption. Evidence limitations, in contrast, cannot be corrected within the assessment; they should be characterized, incorporated into the uncertainty analysis, and linked to further explicit deliberation and potentially to further evidence generation.

3.4 Model Inputs, Model Structure, and Assumptions

Uncertainty in model inputs, such as treatment effects, event rates, costs, and utilities, is commonly represented through probability distributions. Where inputs are correlated, their joint uncertainty should be represented (18). Uncertainty about model structure concerns how the decision problem is represented, including health states and transitions, functional forms, care pathways, extrapolation methods, and relationships between model components (27, 28).

Whether uncertainty is parameterized often depends on how the model is specified. Waning of a treatment effect, for example, may be represented by a waning-rate parameter, and the choice between alternative survival curves by weights assigned to each. The important point is to state clearly what is uncertain, whether it has been parameterized and, if not, how it has been characterized.

Models necessarily involve assumptions. Some assumptions are required because relevant evidence is limited or absent, such as assumptions about the duration of treatment effects or the representation of treatment discontinuation, switching, and subsequent therapies. These become important sources of uncertainty when several plausible alternatives exist, and the choice between them could materially affect the results. Other assumptions simplify the analysis, such as using a general-population life table or average costs. These simplifying assumptions should be documented and justified but do not always need to be treated as sources of uncertainty. Some simplifications, however, may obscure heterogeneity, for example by representing mild, moderate, and severe disease with a single weighted-average cost and survival estimate. Where clinically or decision-relevant heterogeneity can be identified and supported by the evidence, modeling subgroups separately may better represent differences in costs, outcomes, and cost-effectiveness that would otherwise be obscured by population averages (29).

The relevant comparator is determined by the decision problem and the requirement of the HTA organization. Differences in standard care between jurisdictions or settings therefore generally represent heterogeneity rather than uncertainty about the technology itself. However, uncertainty may remain about how current practice is represented and, importantly, about the relative effectiveness of the technology compared with the selected alternatives, particularly where relative effects must be estimated using indirect comparisons, real-world evidence, or assumptions ([Section 3.6.1](#)).

Most uncertainty in economic evaluations can be attributed to a limited set of evidence situations, ranging from relevant, well-conducted studies whose estimates are affected by sampling uncertainty to decision problems for which little or no direct evidence exists. These situations can arise with any study design, and their implications depend on the relevance, design, conduct, and limitations of the evidence, not on study design alone. Table 3.1 summarizes these situations, indicating whether they are typically parameterized in current practice, and describes how they can be characterized when they are not. [Sections 4](#) and [5](#) outline the corresponding analytical and decision responses.

Table 3.1. Sources of Uncertainty in the Evidence and Approaches to their Characterization

Evidence Situation	How it Appears in the Model	Can it be Parameterized?	Characterization Where Not Parameterized
Sampling uncertainty from well-designed, relevant studies	An interval around a treatment effect, event rate, cost, or utility that reflects the size of the source study	Yes, routinely, using a distribution fitted to the data, with correlations preserved	Not applicable
Estimates from studies at risk of bias (e.g., single-arm studies with external controls; observational studies with residual confounding)	An effect, cost, or utility that may be systematically overestimated or underestimated	Sometimes, through bias adjustment or quantitative bias analysis; rarely undertaken	Scenarios across plausible magnitudes and directions of bias; elicitation of likely bias
Estimates that do not generalize to the decision setting (different population, standard of care, jurisdiction, or period; indirect evidence against the relevant comparator)	An estimate that may not apply to the relevant patients, comparators, or health system	Sometimes, by re-estimation for the target population; often not undertaken	Re-analysis using local data; scenarios; elicitation on transferability

Evidence Situation	How it Appears in the Model	Can it be Parameterized?	Characterization Where Not Parameterized
Assumptions in the absence of evidence (e.g., effects in an unstudied subgroup; surrogate-to-final-outcome relationships; treatment switching)	An input or relationship with no empirical source	Sometimes, by representing the assumption as an elicited parameter	Alternative assumptions as scenarios; explicit statement for deliberation
Judgments from structured expert elicitation	A distribution derived from expert judgment	Yes; the elicited distribution can be included in the probabilistic analysis	Reporting of between-expert variation and its causes; sensitivity to expert selection and pooling method
Extrapolation beyond observed follow-up (e.g., long-term survival; duration of treatment effect)	Lifetime outcomes depend on a period not observed in any study	Partly; parameters of a selected curve are included, but the choice of curve rarely is unless model averaging is used	Alternative extrapolations as scenarios; model averaging with justified weights; external or registry data
Omission of aspects of the disease or treatment (e.g., patient-relevant outcomes; caregiver effects; treatment sequencing)	Consequences not represented in the model	Rarely, only if the omitted element is incorporated	Statement of the omission and its likely direction of effect; scenarios incorporating it where data allow
Changes over time (standard of care, new comparators, prices, software versions)	Evidence available at assessment may not reflect later practice	Rarely	Scenarios for future practice; prespecified triggers for reassessment

3.5 Documenting Uncertainty: The Uncertainty Assessment

Uncertainties should be identified systematically and documented before analytical methods, approaches or decision options are selected. An uncertainty assessment table can provide a structured record of uncertainties that may be important to the decision. Its purpose is to make potentially important uncertainties explicit, document how they are presented, and provide a basis for subsequent analysis and committee deliberation.

For each uncertainty, the table should record:

- its source in the evidence and the reason it arises;
- which model inputs, features of the model structure, or assumptions it affects, and its relationship to other uncertainties;
- its likely direction and plausible magnitude of effect, and whether it could change the recommendation;
- how it is represented in the analysis (for example, through probabilistic analysis, scenario analysis, or diagnostic check, or not quantitatively represented) and where it has already been reflected, such as in a base-case assumption, probability distribution, scenario, or other decision consideration;
- whether it is likely to resolve over time, could be reduced through feasible additional research, or is likely to persist and, if further research could help, what evidence would be required.

Recording how uncertainty has already been reflected can help avoid counting the same underlying evidence limitation more than once. For example, a single limitation may influence the choice of a conservative base-case assumption, the width of a probability distribution, a pessimistic scenario, and subsequent consideration of price or cost-effectiveness threshold. Each response may be reasonable individually, but their combined effect may disproportionately influence the same underlying uncertainty if the connections between them are not made explicit.

Uncertainties should also be considered jointly where appropriate. Some may have little effect individually but become important in combination with others, and model outputs rarely change in direct proportion to individual inputs ([Section 4.3](#)).

Several published tools can support systematic identification and documentation of uncertainty. The TRansparent Uncertainty ASsessmentT tool (TRUST) provides a structured approach to identifying and rating uncertainties (26). The Assessment of Risk Table (ART) and the Appraisal of Risk Chart (ARCH) distinguish uncertainties represented within the probabilistic analysis from those outside it and relate them to possible policy responses (30). Directed Acyclic Graphs with Omitted Objects Displayed (DAGWOOD) (31) and Decision Analytic Models with Omitted Objects Displayed (DAMWOOD) (32) make structural choices and omissions explicit. This report does not endorse a particular tool.

Different tools serve different purposes, and the central requirement is a systematic and transparent account of what is uncertain, why it matters, and how it has been addressed.

3.6 Uncertainty by Type of Evidence

3.6.1 Treatment Effects and Safety

Treatment-effect estimates are subject to sampling uncertainty, which is routinely represented in probabilistic analysis. Bias and limited relevance to the decision setting may be more consequential but are less often parameterized ([Section 3.7](#)). Related outcomes, such as progression-free and overall survival, may be correlated; treating them as independent can understate variation for some combinations while producing implausible combinations for others (18).

Single-arm studies are common in accelerated oncology and rare-disease programs, and some regulatory approvals rest on small or single pivotal trials (33, 34). When no concurrent comparator is available, comparative effectiveness must be estimated using external controls or indirect evidence. These approaches can strengthen the evidence base but may not fully address differences in populations, outcome definitions, follow-up, study conduct, or residual confounding (35).

Real-world evidence, including target trial emulation, can complement trial evidence and address some of these gaps (36). Its value depends on the quality and relevance of data, study design, conduct, and analysis, not the data source alone. A well-designed observational study may be more applicable to the decision setting than a trial conducted in a different population, whereas a poorly designed observational study may introduce selection bias, measurement error, or residual confounding. NICE (37), CDA-AMC (38), and the French Haute Autorité de Santé (39) have published guidance on the conduct and reporting of such studies.

Indirect comparison may also be needed when the trial does not include the comparator relevant to the jurisdiction. Sampling uncertainty from these analyses can usually be incorporated into probabilistic analysis, but uncertainty arising from assumptions about exchangeability, population adjustment, model specification, or residual confounding may remain outside it. The population considered in an HTA may also correspond to a subgroup, such as patients with a specific biomarker, for which the trial was neither designed nor powered. In addition, rare or late adverse effects may be poorly characterized because of limited sample size or follow-up.

3.6.2 Baseline Risk, Natural History, and Transferability

Baseline event rates, disease progression, and mortality determine the absolute benefits associated with a relative treatment effect. These inputs are commonly derived from observational studies, registries, or administrative data and, like treatment-effect estimates, are subject to sampling uncertainty. Large datasets may produce very small standard

errors, but this does not establish that the estimates are relevant to the decision population. Applicability to the population and care setting of interest may therefore be the more important consideration ([Section 3.7](#)).

At a minimum, reports should document differences between the source and target populations, differences in standard of care, changes in clinical practice over time, and limitations arising from sparse data in rare diseases or underrepresented groups ([Section 3.2](#)). Estimates derived locally may still become outdated if treatment patterns, diagnostic practices, or access to subsequent therapies change.

Evidence from other countries or regions may not transfer directly to the decision setting because of differences in baseline risk, clinical practice, treatment pathways, or health-system characteristics (40). Where appropriate, local observational data can be used to adapt trial evidence to the decision population; for example, by applying a trial-based treatment effect to a local baseline risk. Repeating an observational analysis in the decision jurisdiction may also be valuable when differences in population or practice could materially affect the results, even if the analysis offers limited methodological novelty.

3.6.3 HRQoL and Patient-Relevant Outcomes

Utility values used to represent HRQoL can be affected by several sources of imprecision and bias (41). Values estimated from patient or general-population samples are subject to sampling uncertainty and may also differ in their relevance to the population of interest. Patient-reported health status may be affected by measurement error, while the value sets used to convert health states into utilities are themselves estimated from data and therefore subject to sampling uncertainty. Generic instruments may be insensitive to changes that matter in a particular condition, and mapping between measures introduces additional uncertainty. Estimates of caregivers' quality of life raise similar issues and are often based on a more limited evidence base.

A separate issue is whether the model includes the outcomes that matter to patients. Outcomes, such as independence, cognitive function, treatment burden, the ability to work or provide care, and participation in family and social life, may not be adequately captured by standard health-related quality of life measures. Their absence from the evidence does not mean they are unimportant. If an outcome that could affect the decision is omitted from the model, the resulting uncertainty will not be reflected in probabilistic analysis unless that outcome is incorporated. Approaches, such as DAMWOOD (32), can help make such omissions explicit.

Two roles for patients and caregivers should be distinguished. In HTA organizations that use general-population preferences to value health states, patient and caregiver involvement does not replace that convention. Patients and caregivers provide responses to HRQoL instruments used to estimate utility values, and they also contribute first-hand knowledge of living with a condition and its treatment, including how treatment is accessed, monitored, switched, and discontinued, and treatment burdens that may not be captured in

clinical or administrative data. This knowledge can help identify outcomes absent from the model, assess whether the modeled care pathway reflects actual care, and determine whether proposed data collection is feasible. These contributions are particularly valuable during scoping, when the model can still be revised. Important omissions identified through this process should be recorded in the uncertainty assessment and, where appropriate, inform further analysis or research recommendations.

3.6.4 Resource Use and Costs

Cost inputs may be affected by uncertainty in resource use, unit costs, care setting, treatment duration, adherence, wastage, monitoring, subsequent care, and the management of adverse events. Resource use and unit costs should be considered separately because underlying issues differ. Resource use is commonly estimated from trials, observational studies, administrative data, or clinical opinion, and may be subject to sampling uncertainty, bias, limited transferability across settings, and changes in practice over time. Resource use quantities estimated from samples can be assigned probability distributions and, where appropriate, sampled jointly with the clinical events that generate them so that relationships among effectiveness, treatment duration, resource use, and cost are preserved (18). Unit costs are often obtained from published fee schedules, formularies, tariffs, or costing databases. They may involve little or no sampling uncertainty but may not reflect the opportunity cost of providing the service in the setting of interest and may be incomplete for new services or models of care. Because unit costs taken from fixed schedules do not have sampling distributions, they may be better examined through scenarios that reflect plausible alternative settings and costing sources.

Costs should be linked to clinical events, health states and treatment pathways that generate resource use (42). This helps ensure that expected costs consistently represent the modeled clinical pathway. Assumptions about treatment duration, adherence, wastage, and the extrapolation of time on treatment should be examined through scenarios when plausible alternatives could materially affect the results. For targeted therapies, diagnostic and testing costs may also be important. Where many individuals must be tested to identify one eligible patient, relevant uncertainty may include the testing strategy, uptake, test performance, resource use, and which costs fall within the perspective of the analysis (43).

The submitted price of the technology is usually known, but confidential discounts or patient access schemes may mean that the net prices of comparators or subsequent treatments are unavailable to the analyst. This is not uncertainty in the evidence; the information exists but has not been disclosed. An assumption must nevertheless be made in the analysis. The assumption should be stated explicitly and its effect on the results examined over a plausible range. A price reduction may reduce the consequences of uncertainty for the decision, but it does not reduce uncertainty in the underlying evidence or change other costs in the care pathway. The relationship between price, the consequences of uncertainty, and the value of further research is considered in [Section 5](#).

3.7 Appraisal of the Evidence and Probabilistic Analysis

Every estimate and assumption used in an economic evaluation should be critically appraised, regardless of how the associated uncertainty is represented. A standard error or confidence interval describes sampling uncertainty under the assumptions of the source study; it does not establish that the estimate is unbiased, relevant to the decision population, or transferable to the decision jurisdiction. A precise estimate may therefore still be highly uncertain for decision making.

Three aspects of evidence are particularly important: risk of bias, relevance to decision context, and indirectness. *Risk of bias* concerns whether the design or conduct of a study could systematically distort the estimated effect. Relevance concerns how well the study population, setting, standard of care, and time correspond to the decision context. Indirectness arises when the evidence does not directly address the comparison or outcomes required for the decision, for example because it relies on surrogate outcomes or indirect treatment comparisons. GRADE (Grading of Recommendations Assessment, Development, and Evaluation) provides a framework for assessing these considerations for clinical evidence and also considers imprecision, which largely reflects sampling uncertainty (44). ROBVALU (Risk of Bias in studies of Values and Utilities) supports assessment of studies used to derive health-state values and utilities (45).

Some limitations can be explored quantitatively through bias analysis (46) or scenario analyses using alternative data sources, assumptions, or subgroups. These limitations may affect model inputs, model structure, or both; others may remain outside the quantitative analysis altogether. Probabilistic analysis reflects them only when they have been explicitly characterized.

Where a concern cannot be represented credibly in the model, it should be documented explicitly for deliberation instead of being absorbed into an arbitrary assumption. Uncertainties identified during regulatory assessment may also differ from those most relevant to reimbursement decisions, and HTA organizations should therefore assess evidence limitations in relation to the decision at hand, including whether they could materially affect expected costs, health outcomes, or the choice between available options (47).

Probabilistic analysis should also be critically appraised. Its results are conditional on the model structure, evidence, distributions, correlations, and assumptions specified by the analyst ([Section 4.4](#)). Where uncertainties are examined through scenarios, reviewers should assess whether the alternatives are credible, span a plausible range, and have a clear basis for selection.

The same standards apply whether the analysis is submitted by a manufacturer or rerun by a review group. Reviewers should examine which uncertainties are included, how probability distributions are specified, whether important correlations are preserved, whether enough simulations have been run to produce stable results, and what remains outside the

analysis. No agreed checklist currently exists for appraising the completeness and adequacy of probabilistic analysis ([Section 7.3](#)).

3.8 Expert Elicitation and Lived-Experience Expertise

When empirical evidence is absent or insufficient, structured expert elicitation provides a transparent way to represent informed judgment about an uncertain quantity, including its plausible range. Structured approaches are preferable to informal consensus or unexplained assumptions because unstructured judgment may be affected by overconfidence, anchoring, and other cognitive biases (48). Elicitation may inform missing inputs, long-term extrapolation, the applicability of evidence across settings, alternative model structures, rare events, or the weights assigned to alternative assumptions represented as scenarios, which makes the judgment underlying structural choices explicit ([Section 4.5](#)). Established approaches include the Sheffield Elicitation Framework (SHELF) (49), the reference protocol for health care decision making (48), the EXPLICIT spreadsheet-based approach (50), and NICE Decision Support Unit guidance on long-term survival (51).

Expert elicitation should not substitute for evidence that could reasonably be obtained. Its appropriateness depends on why the evidence is missing. It may be useful where relevant evidence cannot yet exist, as with long-term outcomes of a new technology, or when obtaining empirical evidence is not feasible within the required timeframe. Greater caution is warranted when relevant data already exist and could reasonably be analyzed or made available, including data held by the sponsor or other evidence holder. The choice of quantities to elicit should therefore be systematic and justified. Reports should explain why elicitation was required, how experts were selected, what information they received, how judgments were elicited and combined, and whether participants had financial or professional interests relevant to the exercise.

Patients, caregivers, and family members contribute expertise derived from lived experience and should be involved when that expertise is relevant to the quantity or process being elicited. Examples may include treatment burden, time required to access or receive care, symptom patterns, adherence, and reasons for treatment discontinuation. Caution is needed, however, when asking individuals to estimate population-level frequencies or probabilities that they may not be able to observe directly. In these situations, lived-experience input may be more useful for identifying relevant factors, informing plausible ranges or scenarios, or assessing whether model assumptions reflect actual care.

Participants should receive a clear explanation of the purpose of the elicitation, the quantity being estimated, and how their judgments will be used. Appropriate training and accessible materials should be provided, and the time and resources required for meaningful participation should be anticipated and funded.

Elicitation should capture uncertainty within individual judgments and variation between participants. Reasons for disagreement should be explored and documented because they may reflect different interpretations of the question, differences in experience, or genuine uncertainty about the underlying quantity. A narrow elicited distribution indicates that experts were relatively confident and in agreement; it does not establish that the quantity itself is known precisely. This limitation should be stated when elicited inputs materially influence the results.

3.9 Summary

This section has five main implications for those preparing or reviewing economic evaluations. First, uncertainties should be identified systematically and documented in a structured assessment that records their source, how they are represented in the analysis, and their potential relevance to the decision. Second, genuine uncertainty should be distinguished from problems requiring correction, including inadequate transparency, methodological errors, and avoidable flaws in study design or analysis. Third, the assessment should show where each uncertainty has already been accounted for so that the same underlying limitation is not counted more than once. Fourth, uncertainty arising from limited or unrepresentative evidence for underrepresented populations should be made explicit without assuming its direction or consequences. Fifth, for uncertainties that could materially affect the decision, the assessment should indicate whether they are likely to diminish over time, could be reduced through feasible further research, or are likely to persist, and what evidence would be required to address them.

4. Relevance of Uncertainty and Methods of Quantification

In brief. Uncertainty analysis should help decision makers understand whether uncertainty could change a decision and what could happen if another option later proves better. Looking at many uncertainties together can show how confident decision makers can be and what the possible health consequences could be. Looking at one assumption at a time can also be useful, but it does not provide a complete picture. Some uncertainties are harder to include in numerical analyses, such as questions about how a model is designed or whether evidence from one population or health care setting applies to another. When uncertainty cannot be included in the numerical analysis, reports should clearly explain what was left out, why it matters, and how it was considered in other ways.

4.1 Introduction

This section sets out the methods available to quantify uncertainty, with emphasis on what decision makers need to interpret each method's output. Technical derivations are coupled with references to relevant literature rather than dealt with in detail here, as the main focus is the path from the range of uncertainties in a given analysis, through the methods that characterize them, to the single quantification of decision uncertainty and the cost of uncertainty, the implications of which are considered in [Section 5](#). How uncertainty is then presented is the subject of [Section 6](#).

4.2 The Purpose of Uncertainty Analysis

The common objective of economic evaluation for HTA, which was established in [Sections 1](#) and [2](#) — to inform decisions that seek to maximize population health subject to financial and other constraints — determines the purpose of uncertainty analysis. Its job is not to describe how widely a model's outputs vary, but to establish whether that variation could lead to a decision that fails the objective with a cost (financial or health) which can be quantified. The operative question is not “how uncertain is the incremental cost-effectiveness ratio?” but “is the decision uncertain, what would it cost if it was wrong and how can this be addressed?” (9). A useful analysis, therefore, integrates the distinct sources described below into a single characterization of the expected incremental net health benefit, the decision uncertainty around that, including the option with the highest expected net health benefit, the probability that it is optimal, and the cost of the decision uncertainty that remains — rather than a catalogue of separate uncertainties. The decision is rarely a binary accept-or-reject: there is a menu of options (described in [Section 5](#)), and the task is to identify the one that does most for population health given what is and is not known.

4.3 Probabilistic Versus Deterministic Analysis

Probabilistic analysis should generally be the default for characterizing uncertainty for two main reasons. First, decision models are non-linear in their inputs. As a result, evaluating the model at the mean value of each input does not necessarily produce the expected costs and health effects. Probabilistic analysis addresses this by averaging model outcomes over the joint distribution of uncertain inputs, even when decision uncertainty itself is not the primary focus. Second, probabilistic analysis integrates all the characterized uncertainty and presents their combined effect on outcomes, such as incremental net health benefit (52) (discussed in [Section 1](#)). DSA has a useful but narrower role: varying one or two inputs at a time can help analysts understand how the model behaves, exposes coding and conceptual errors, and identifies influential inputs. However, it cannot quantify the uncertainty in a decision because it considers only selected inputs and does not capture their joint variation. Where evidence-based uncertainty bounds are unavailable, DSA may also rely on ranges chosen by convention rather than derived from empirical evidence. As discussed in [Section 2](#), varying an input by, for example, $\pm 20\%$ therefore shows how sensitive the model is to that change, but not how uncertain the input itself is.

When ranges are derived from the evidence, a one-way analysis can provide useful information about the influence of a particular input. Its interpretation should nevertheless remain limited: it does not capture the joint effects of uncertainty across inputs and does not provide measures such as the probability of a different decision or the expected loss associated with decision uncertainty.

These limitations do not imply that probabilistic analysis is free from judgment. The distributions and correlations entering a probabilistic analysis, and therefore any VOI metric computed from them, rest on the same underlying evidence and the modeling assumptions. The case for probabilistic analysis rests on its ability to propagate characterized uncertainty through the model, not on any inherent superiority of its inputs. Analysts must still justify how those inputs, distributions, and correlations have been specified. Finally, we note that HTA decision makers in practice often weigh a few selected (frequently structural) uncertainties more heavily than the integrated probabilistic analysis; [Section 5](#) returns to this.

4.4 Parameter Uncertainty and Probabilistic Sensitivity Analysis in Practice

While PSA is more appropriate than DSA as a method of quantifying uncertainty, it is only as credible as the distributions and dependencies it encodes, and a few pitfalls recur. Correlations between inputs are routinely ignored, and pruning the analysis to a convenient handful of “key” parameters produces an illusion of a coherent uncertainty analysis while leaving it materially incomplete. A particular — and increasingly common — challenge is the use of a standard error of a sample mean lifted from a single study, which reflects sampling

in that study's setting and not the genuine uncertainty about the mean parameter in the population, place and period to which the decision applies. When treated as if it were the latter, it does not represent what is actually unknown. The remedy is not more computation, but better characterization — distributions and correlations chosen to represent the decision-relevant uncertainty, to a defensible standard of completeness. Of the two, the choice of distribution is usually the more constrained: the form follows from the parameter's support and the way it was estimated — beta for probabilities and utilities, gamma or lognormal for costs, multivariate normal for sets of regression coefficients (53). Correlation structure is more often neglected. It is also, in practice, estimable when correlated parameters are measured in the same dataset. Published evidence typically supplies only partial correlations and the remainder is supplied by assumption. Where correlations are assumed rather than estimated, they are an elicited quantity, and the expectations set out in [Section 3.7](#) apply to them — they should be reported and justified.

4.5 Bringing Non-Parameterized Uncertainty into the Uncertainty Analysis

[Section 3](#) describes which sources of uncertainty are typically parameterized and which are not. This section deals with the methods available for the latter.

In current practice, scenario analysis is the most common approach. Its limitation is that it is a form of DSA in that it explores discrete alternatives without weighting them, so it shows whether the decision changes but not how likely each structure is to be correct — which is where other methods could add value. These include model averaging or weighting across alternative structures (including parameterizing the structural choice so it can be varied like any other input), and formal expert elicitation for quantities on which no data are available (54). Which method is appropriate depends on the extent to which the alternatives can be enumerated and whether they can be credibly weighted.

Such uncertainty can also be made decision-relevant without being fully quantified: running several independently built models against the same question (as in the CISNET cancer models (55) and comparable work in other diseases) shows whether different defensible structures point to the same policy conclusion or to different ones — which shows whether structural uncertainty becomes decision uncertainty.

Because scenario analysis carries most of what is not parameterized, how the scenarios are chosen matters as much as the result. A defensible set is specified before the results are seen, states the basis for each alternative, and spans the range that the evidence and the available judgment can support rather than a pair of convenient extremes. Each scenario should correspond to an identified uncertainty in the assessment ([Section 3.4](#)), so that a reader can see which concern it addresses and what would have to be true for it to hold. Where the alternatives can be enumerated and weighted, model averaging is preferable, because it returns a single distribution rather than a set of separate results that still have to be combined by judgment. Our recommendation is that analysts make greater

use of these methods — model averaging, switch parameters and structured elicitation — to bring into the probabilistic analysis sources of uncertainty that are currently handled by scenarios alone, wherever this is feasible.

We further recommend that every scenario analysis be accompanied by a probabilistic analysis of that scenario and, at a minimum, the expected value of perfect information. Computing EVPI alongside a probabilistic analysis is no longer a substantial undertaking, and freely available tools such as the Sheffield Accelerated Value of Information (SAVI) application (56) make it a routine step rather than a specialist exercise. Parameter-level EVPI and the expected value of sample information are valuable where a research decision is in prospect, but we do not propose them as a general requirement.

4.6 Integrating Multiple Sources of Uncertainty in the PSA

The hard problem is integrating multiple sources of uncertainty into a common PSA to generate the distribution of expected incremental net health benefit presented in [Section 2](#). In this report, by full integration we mean placing every characterized uncertainty into a single probabilistic analysis. This would include sources of uncertainty such as extent of bias, choice of survival model for extrapolation, effectiveness waning and relevance of evidence to the local population. As described in [Section 2](#), a fully integrated cost-effectiveness distribution provides the building blocks to enumerate the trade-offs presented to decision makers for deliberation, rather than expecting them to guess at decision uncertainty and its consequences. Two terms are used from this point on. *Tier 1* describes the case in which every characterized uncertainty has been placed in one probabilistic analysis, so that decision uncertainty and its consequences can be read directly from that analysis. *Tier 2* describes the common case in which integration is incomplete, with the analysis supplemented by explicit reasoning about what it has not captured, typically using scenarios relating to the sources of uncertainty which are not easily parameterized. The value of *Tier 1* is that it shows an ideal to aspire to. *Tier 2* acknowledges that in practice integration is usually partial, and what options are justified in such a case. Generally, our recommendation is that the assessment report states which parameters, correlations or other sources of uncertainty have been omitted from the probabilistic analysis, and not merely that the analysis is incomplete. What has been left out is decision-relevant in its own right: uncertainty about clinical benefit and uncertainty about costs point to different options, and a reader who knows only the size of the residual cannot tell which applies.

Methods for full integration of uncertainty in PSA are developing, but none is yet routine, and an analysis that integrates some sources while ignoring others may under- or overestimate decision uncertainty in a way that can be worse than candid incompleteness — also for communication reasons.

This also raises the question of where the analytic burden falls, which differs across the three sponsoring agencies according to whether assessment is conducted in-house or commissioned, and which differs again between pharmaceuticals and devices within the

same agency. Any expectation of completeness therefore lands on a different party in each setting, and a defensible minimum standard has to balance the value of doing this with the time and resources required.

Finally, integration has its place in so far as it changes the presentation of decision uncertainty and its cost — the elements [Section 5](#) acts on — and the methods must not privilege what is easy to quantify over what matters to decisions impacting population health.

4.7 Summary

Uncertainty analysis earns its place by answering a decision question. Probabilistic analysis is the default on two independent grounds — unbiased expectations under non-linearity and the characterization of decision uncertainty and its cost. Deterministic and scenario analysis retain a narrower role as “uncertainty diagnostics” and as the main current route to structural uncertainty, but neither yields a probability of error or a loss. The credibility of a probabilistic analysis rests on the distributions and the correlations it encodes, and where these are assumed rather than estimated, they should be reported as the judgments they are. Where integration is incomplete — the Tier 2 case, which is the common one — the analysis should say what has been left out, because what is missing bears on which decision option is indicated. [Section 5](#) takes those outputs and links them to the options available.

5. Linking Uncertainty Analysis to Decision Options

In brief. Uncertainty does not always call for the same response. What matters is how much could affect the decision, what the consequences could be if another option later proves better, and whether more evidence could reduce the uncertainty. Delaying access can mean patients lose health while waiting for an effective treatment. Funding a treatment that later proves less beneficial than expected can also affect other patients if resources are shifted away from other care. Both consequences should be considered. Options may include providing access now, changing the price, providing access while more evidence is collected, limiting access to a research setting, or not funding the treatment. The choice depends on the expected health benefits and costs, what remains uncertain, whether useful new evidence can realistically be obtained, how long and costly that research would be, and what it would require from patients and health services. A lower price can reduce the consequences of uncertainty, but it does not provide better evidence about how well the treatment works.

5.1 Introduction

Each decision option is a different way of acting on the distribution of expected net health benefit and on the expected cost of uncertainty (i.e., the potential of forgone research in terms of population health). In this report, two principles from the decision-theoretic foundation have governed the link. The first is that the response should turn on the consequences of error — the expected loss from a wrong decision, the cost of uncertainty, and the extent to which this can be addressed by additional research — not on the level of uncertainty as such.

The second is that those consequences arise on both sides of the decision, though not in the same way. To that extent, we distinguish reversibility from recoverability. “*Reversibility*” is a property of the decision. It is asymmetric, and the asymmetry favors a negative recommendation: a technology not adopted can be revisited when evidence arrives, whereas one already embedded in practice is difficult to withdraw. Furthermore, there are limited incentives for all parties to generate the evidence that may trigger review of a positive recommendation, and agencies have limited capacity to reassess at all.

Recoverability is a property of the consequences. In the time period before a decision is revisited, health forgone is irreversible whichever way the decision went: health forgone by patients who deteriorate or die while an effective technology is withheld — or in chronic disease, who live for that period with symptoms, disability or treatment burden that could have been avoided — cannot be recovered later any more than health displaced by funding a cost-ineffective intervention can be, and that loss may fall harder on patients already

underserved. Note that we use “*irrecoverability*” more broadly than some other published work, which tends to emphasize the sunk costs of adoption such as training, capital, and disinvestment elsewhere, which are all relevant too. “Irrecoverable” costs include health forgone, which cannot be recovered on either side of the recommendation, and committed costs such as capital, training and disinvestment elsewhere, which are recoverable to varying degrees and only on one side. As such, all wrong decisions (whether positive or negative) carry costs that cannot be recovered, and should be identified, quantified and factored into the decision.

In summary, decision uncertainty imposes a cost we can potentially act on — through more evidence development, waiting for uncertainty to resolve, a lower price, or several of these (e.g., a discount and evidence generation are frequently combined, as in managed access). So, it is the response, not the level of uncertainty as such, that the rest of this section addresses.

Communicating this to various audiences is also critical for agencies that are increasingly dealing with substantial decision uncertainty and will be further developed in [Section 6](#).

5.2 The Expanded Decision Menu

Before the menu itself, six considerations determine which option fits, and each applies on both sides of the decision:

- the expected cost of a wrong decision, in terms of the maximum value of the evidence potentially forgone;
- the probability that a wrong decision could be reversed, together with the time a review would take;
- the extent to which the research under contemplation would resolve the uncertainty — a matter of degree, and the same quantity as the decision uncertainty that would remain at the second decision;
- the cost of the research itself;
- the burden the option places on patients and on services, which can render an option that is sound unworkable on paper in practice ([Section 3.6](#)); and
- whether the research would in fact be carried out and reported.

It is these, not the magnitude of the uncertainty itself, that should drive the choice of option.

The menu runs from recommending at the submitted price to not recommending, with several conditional options in between. What makes each option appropriate is a particular profile of uncertainty — its size, whether it can be resolved, and the consequences, on both sides, of being wrong.

Two general points apply across the menu. Some uncertainty resolves without anyone acting on it: research underway elsewhere reports, or a comparator goes generic, and its price falls. Where resolution is expected within a time-period the decision can tolerate,

waiting is itself an option and should be considered alongside the others. And a change in price does not simply select one option; because it alters the expected cost of being wrong and therefore the value of resolving the uncertainty, it shifts the balance across the whole menu.

- **Recommend at the submitted price** — expected net health benefit is non-negative, and the cost of decision uncertainty is small relative to the cost and delay of seeking more information.
- **Recommend with an outcome-based payment arrangement in place** — where the decision uncertainty cannot be resolved economically, tie payment to realized outcomes so that expected net benefit stays non-negative. No uncertainty discount should be applied, as that would penalize the manufacturer for uncertainty that is not worth reducing or is unfeasible. Outcome-based payment models belong here, wherever realized outcomes can be observed at a reasonable cost and tied to payment. The agreed price would be unchanged, which matters where that price is referenced by other jurisdictions, but what changes is the payer's realized cost across the covered population, since the manufacturer earns the agreed price only where the agreed outcome is achieved, or rebates it where it is not. The arrangement shifts risk rather than resolving the underlying uncertainty at the point of implementation but allows for real-world evidence generation that can inform a stable value-based price in the future.
- **Recommend at a reduced price in lieu of refused research** — where the decision uncertainty would be worth resolving but the manufacturer declines to undertake or fund research that would be feasible and worthwhile, a price reduction may be justified. It is set at the research-indifference price: the price at which the expected net benefit of the best available study falls to zero. At that price, the forgone research is no longer worth doing, so no value is lost by not doing it. Hence this price reduction is not a penalty for uncertainty and is not an estimate of the value of research measured at some other price.
- **Recommend with managed access/coverage with evidence development (CED)** — provide funding while additional research is undertaken, where some of the decision uncertainty is resolvable, the value of the additional information exceeds its cost, and the costs of adoption are considered recoverable. A further condition is a credible expectation that the research will be done and reported over a reasonable period ([Section 5.6](#)). Finally, if fit-for-purpose evidence later shows the technology to be more cost-effective than previously estimated, the price adjusts upward; and vice versa if it turns out to be less cost-effective.
- **Recommend only in research (OIR)** — restrict use to a research setting, where the value of resolving the decision uncertainty is high, but the irrecoverable costs of routine adoption are too large to risk while the evidence is gathered.
- **Do not recommend** — either expected net benefit is negative and cannot be made non-negative at an acceptable price; or expected net benefit is non-negative

and further research is efficient, but no agreement on that research can be reached and funding the technology regardless would undermine incentives for appropriate evidence development. The second case is a dynamic-incentive judgment, not a static VOI result: it must weigh the system-wide incentive effect against the expected — and potentially irrecoverable — health forgone by current patients, which falls hardest on underserved groups (see [Section 3](#)). Where the incentive case does not clearly outweigh the health loss, the reduced-price option above is preferred to rejection.

Current availability of the menu options above differs between jurisdictions. OIR, in particular, has been used at scale in only one of the three settings represented here. Some options therefore appear because the logic requires them rather than because they are presently available everywhere. While any option is only as good as its enforceability, decisions should be guided by the cost in terms of population health stemming from not offering it on the menu (and going with a suboptimal approach instead) versus the cost of implementing the option.

5.3 Mapping Uncertainty Profiles to Decision Options

Decision options should be chosen by starting from the source of the uncertainty rather than from a single summary statistic, because the same headline uncertainty can call for different responses depending on where it comes from and whether it can be resolved. The chain of reasoning is that the economic evaluation identifies the sources of uncertainty, characterizes them as uncertain parameters or as alternative scenarios, and uses probabilistic analysis to establish decision uncertainty and EVPI, under a range of scenarios where necessary. What then passes to the HTA organization is the residual decision uncertainty, together with the feasibility of evidence generation, the irrecoverable costs on each side, and the timing of any research. Judging those and selecting the option is typically the HTA organization's task rather than the analysts', and in most settings it is work the secretariat and committee already do.

We set out the position in principle first. Where a fully integrated probabilistic sensitivity analysis is available (Tier 1), the characterization of decision uncertainty is complete and the coverage-decision logic of Claxton et al. (2016) (6) can inform decisions directly, with no separate reasoning by type of uncertainty required. Table 5.1 sets out the Tier 2 fallback for the common case in which integration of all uncertainty into the PSA is incomplete, and it is necessary to reason explicitly about what the PSA has not captured.

Two contrasts show why the starting point matters. Sampling uncertainty and non-generalizable evidence can have the same impact on the probability of cost-effectiveness, yet the first can, in principle, be resolved by more evidence — the territory of VOI and managed access — while the second is not resolved by more of the same data, and points to evidence generated in the decision setting or, failing that, a lower price. And high uncertainty about which of several similar options is best is not uncertainty about whether to

fund the class at all: where substitutes cannot be separated, the better response may be to make several available and let clinical choice operate. Uncertainty that is not parameterized may in some cases be resolved by incorporating a dimension of treatment or benefit that has been left out, particularly one that matters to patients, and so also generates a research recommendation (5, 6).

5.4 Value of Information as Decision Support

VOI analysis quantifies the cost of uncertainty and provides information to support decisions on research prioritization. The expected value of perfect information (EVPI) estimates what it is worth paying to resolve all uncertainty; partial perfect information (EVPPI) points to which parameters where the value of information is highest; sample information (EVSI) and the expected net benefit of sampling (ENBS) value the information generated in a specific study against its cost (57). Importantly, VOI weights a “wrong yes” the same as a “wrong no”: further research (OIR, or managed access where routine adoption is compatible) is warranted only where the expected value of resolving the uncertainty exceeds the cost of that research and of the health forgone by waiting — where the expected net benefit of sampling is positive ($ENBS > 0$). What differs across settings is the type of VOI undertaken and the route by which its result reaches a decision.

Our recommendation, set out in [Section 4](#), is that EVPI be reported alongside every probabilistic analysis as a matter of course, with EVPPI and EVSI used where a research decision is actually in prospect. The report does not present VOI as uniformly applicable. In England it has an established, if intermittent, role in research prioritization and in OIR guidance. In other settings there is at present no established mechanism by which a VOI result enters coverage decisions, and we recommend that agencies without such a route consider establishing one. Where a formal route exists, the relevant VOI metrics can be estimated and used directly. Where it does not, VOI remains a framework for reasoning about whether uncertainty is worth resolving. It operates as a decision principle applied through structured judgment rather than routine quantification (58). A low value of information despite high uncertainty is a reason not to demand more evidence. The choice is therefore to recommend where expected net benefit is non-negative at the relevant price, and to reject outright where it is negative and cannot be made otherwise. Two qualitative judgments have to be made alongside any value-of-information result, because the methods do not capture them: whether the research would in fact be undertaken and reported, and whether the system has the capacity to reassess the decision when it reports.

5.5 Managed Access and CED: What Good Looks Like

Recent experience shows what separates a workable managed-access agreement from a failed one (59):

- clearly specified uncertainty that the agreed additional research can resolve;
- a data-collection plan that is realistic for the services expected to run it and for the patients expected to take part, with feedback timely enough to inform the reassessment it is meant to support;
- pre-specified review and exit;
- arrangements for continuity of treatment, and a defined transition, for patients already receiving the technology should access later be narrowed or withdrawn; and
- a shared understanding of what each party is signing up to, in which the parties include the patients whose participation the agreement presupposes and who should have had an opportunity to shape requirements that fall on them ([Table 3.1](#)).

An uncertainty analysis intended to support such an agreement must, therefore, specify which sources of uncertainty are at issue, what evidence would resolve it, over what period, and what decision each result would trigger. The alternative-payment-model framework (60) maps these problems to instruments; and an analysis of how uncertainties present at EMA approval correlate to HTA outcomes (47) shows that the uncertainties regulators flag are not automatically the ones that drive reimbursement. For instance, Bloem et al. (47) flagged various types of uncertainty at the time of EMA assessment and found that medicines scoring high on these were more likely to receive a negative relative effectiveness assessment from HTA bodies, despite a positive EMA approval. They suggest this may be because regulators tolerate relatively high uncertainty on elements that matter less to their own approval decision, such as non-preferred comparators or surrogate endpoints, but these can weigh heavily in HTA's assessment. This is why the HTA must do its own characterization ([Section 3.7](#)) (12).

5.6 Linking the Uncertainty Root Causes to Decision Options

Agencies already link decision uncertainty to options in overlapping but distinct ways — different conditional-reimbursement vehicles, different roles for price, different reassessment machinery. This report proposes a taxonomy of decision options and the underlying logic, recognizing that implementation will remain jurisdictional (the specific instruments and their legal basis).

Table 5.1. From Source of Uncertainty to Indicated Option, for the Tier 2 Case.

Source of Uncertainty (as Table 3.1)	What is Carried into the Decision	Indicated Option(s)
Sampling uncertainty from well-designed, relevant studies	Fully captured in the probabilistic analysis; EVPI and EVPPI measure the whole residual.	Low cost of uncertainty: recommend at the submitted price. High and worth resolving: managed access or CED where routine use is compatible with the research, only-in-research where it is not.
Estimates from studies at risk of bias	Captured only where a bias adjustment has been parameterized; otherwise EVPI understates the cost of uncertainty and the shortfall should be stated.	Resolvable by better-designed evidence: CED or only-in-research. Not resolvable: price adjustment so that expected net benefit is non-negative, or an outcome-based arrangement where the outcome can be observed.
Estimates that do not generalize to the decision setting	Partly quantified at best. More evidence of the same kind does not resolve it.	Evidence generated in the decision setting: CED with a defined data plan. Otherwise, price adjustment. Check first whether relevant local data are already held but unreleased.
Assumptions in the absence of evidence	Enters through scenarios unless elicited as a parameter; the swing across credible scenarios is the decision-relevant quantity.	Robust across assumptions: recommend. Decision swings: CED or only-in-research to obtain the missing evidence, or a price at which the swing no longer changes the decision.
Judgments from structured expert elicitation	Quantified, but conditional on who was elicited and how judgments were pooled.	Recommend where the decision is stable across expert selection and pooling method; otherwise treat the elicited quantity as a research priority.
Extrapolation beyond observed follow-up	Parameters of the selected curve are captured; the choice of curve usually is not. Time rather than sample size is the missing ingredient.	Robust across extrapolations: recommend. Decision swings but follow-up will mature: time-limited recommendation with reassessment, or payment tied to realized durability. Large swing with irrecoverable costs: price adjustment, CED or only-in-research.
Omission of aspects of the disease or treatment	Not captured at all. The likely direction of effect can usually be stated even where the magnitude cannot.	A pivotal missing outcome: CED or only-in-research to collect it. Where that is not feasible: restriction to the population in which the omission does not bite, or price adjustment.
Changes over time	Not captured; the analysis is conditional on practice at the time of assessment.	Time-limited recommendation with pre-specified reassessment triggers, and provision for review where a comparator's price or availability is expected to change.

Net benefit at the relevant price is assessed throughout, and a change in price moves the assessment for every row rather than selecting one of them. Where a fully integrated probabilistic analysis is available (Tier 1) no separate reasoning by source is needed: the coverage-decision logic selects the option directly from expected net benefit and the cost of uncertainty. The rows above are the Tier 2 fallback. The

Assessment of Risk Table and Appraisal of Risk Chart, introduced in [Section 3.4](#), can be used to carry quantified and non-quantified uncertainty into a single summary.

5.7 Stylized Examples

To show how the reasoning in this section translates into decisions, we developed five stylized examples that span different types of technology and are constructed to vary the features that determine the choice of option:

- expected net health benefit and how it shifts under alternative non-parameterized assumptions;
- decision uncertainty and the value of further research;
- whether that research is feasible, with or without access;
- the irrecoverable consequences of a wrong yes and of a wrong no.

For each, Table 5.2 sets out the option(s) indicated from the menu in [Section 5.2](#), while Figure 5.1 summarizes the set.

Figure 5.1 below shows expected population incremental net health benefit on the horizontal axis. The vertical axis shows EVPI relative to the cost of the proposed research with points above the line being candidates for research. Filled bubbles show the results under the most plausible assumptions (base case) and hollow bubbles for the alternative assumptions, with bubble size indicating absolute EVPI. The quadrants reproduce Tier 1 logic. The arrows show why — in absence of full uncertainty integration in the base case — Tier 2 reasoning is needed as in three of the five examples (A, C, E), plausible alternative assumptions would move the technology across the threshold, and in two (A, C) they would also reduce the value of research. Note that all positions are ordinal and illustrative.

Table 5.2. Stylized Examples: From Uncertainty Profile to Decision Option

	Example A	Example B	Example C	Example D	Example E
New intervention	Medicine for obesity in a large population	Medicine for a rare complication of surgery; small population	Oncology medicine for advanced disease with a moderate-sized population	An advanced therapy medicinal product for a moderately rare cancer	A new biologic therapy for a prevalent skin complaint in young people, licensed in moderate to severe disease
Is this technology cost-effective on average?	Most plausible assumptions: incremental cost-effectiveness ratio below threshold; moderate positive population expected incremental net health benefit (INHB). Some key non-parameterized assumptions: incremental cost-effectiveness ratio above threshold; moderate to large negative population expected INHB.	Most plausible assumptions: incremental cost-effectiveness ratio ~ same as threshold; small positive population expected INHB. Some key non-parameterized assumptions: incremental cost-effectiveness ratio above threshold with small to moderate negative population expected INHB.	Most plausible assumptions: incremental cost-effectiveness ratio below threshold; moderate positive population expected INHB. Overall survival evidence is immature and the regulatory RCT has crossover after progression. Some key non-parameterized assumptions result in an incremental cost-effectiveness ratio above threshold with moderate negative population expected INHB.	Most plausible assumptions: incremental cost-effectiveness ratio above threshold; so there is a small to moderate negative population expected INHB. Some alternative assumptions suggest an incremental cost-effectiveness ratio below the threshold with small to moderate positive population expected INHB.	Most plausible assumptions: incremental cost-effectiveness ratio above threshold; given the population size, this translates to moderate to large negative population expected INHB. Some alternative assumptions result in an incremental cost-effectiveness ratio similar to the threshold with a small positive population expected INHB.
What is the probability of a wrong decision given current evidence?†	Most plausible assumptions: 25% (that recommending is wrong). Alternative assumptions: up to 40% (that not recommending is wrong).	Most plausible assumptions: 50%. Alternative assumptions: up to 45% (that not recommending is wrong).	Most plausible assumptions: 30% (that recommending is wrong). Alternative assumptions: up to 60% (that not recommending is wrong).	Most plausible assumptions: >80%. Alternative assumptions: 30–50% (that recommending is wrong).	Most plausible assumptions: ~70%. Alternative assumptions: 40–50% (that recommending is wrong).
What is the approximate magnitude of the EVPI?	Most plausible assumptions: large. Alternative assumptions: small to moderate.	Most plausible assumptions: small. Alternative assumptions: small to moderate.	Most plausible assumptions: moderate. Alternative assumptions: small to moderate.	Small under most plausible assumptions given the low decision uncertainty and relatively small population. For some alternative assumptions this increases to moderate.	Moderate to large under most assumptions.
What are the major sources of decision uncertainty?	1. Duration of effect on BMI, CV disease and diabetes. 2. Impact of non-modeled benefits (e.g., cancer). 3. Others.	1. High risk of bias in effectiveness due to single-arm study. 2. Comparator may become generic, with an effect on its price, in ~3 years (this would raise the incremental cost-effectiveness ratio of the new medicine).	1. Most suitable extrapolation of overall survival. 2. Appropriate correction for crossover.	A single-arm regulatory study provided the key effectiveness evidence, with uncertainty related to the choice of comparator evidence and adjustment for selection bias.	Regulatory trials and network meta-analysis to incorporate relevant comparators result in uncertainty in the magnitude of HRQoL benefit. The trials are short and small. There is uncertainty in whether the full health effects on patients are reflected in the HRQoL measure.

	Example A	Example B	Example C	Example D	Example E
Are these potentially researchable? What is the duration and cost?	1. Yes: RWE or longer randomized follow-up in regulatory trial, 3–5 years; cost less than EVPI under most assumptions. 2. Yes: secondary sources, 1–2 years; cost less than EVPI under most plausible assumptions. (Cost below EVPI is a first hurdle; ENBS > 0 still to be shown.)	1. Yes, new trial but very slow recruitment (over 8 years). 2. No, will resolve without research.	Yes, continued follow-up of regulatory RCT, although complicated by crossover: follow-up will do more to resolve (1) than (2). Cost limited, 2–3 years.	A randomized study is possible, but recruitment is considered challenging with access, so the duration would be long and cost high. There are concerns about whether real-world data on comparator therapies are reliable.	A larger and longer RCT is considered feasible and of high value under all scenarios. This may need to incorporate an updated measure of HRQoL.
Is the research feasible with immediate access?	1. Yes for RWE; no for extended randomized follow-up, as control-arm patients would be expected to switch once the medicine is routinely available. 2. Yes.	No, RCT unlikely to recruit.	Yes, existing trial.	Concerns about recruitment levels to an RCT with access; recruitment may be more achievable if use is restricted to a research setting.	This is considered feasible.
Irrecoverable consequences of a wrong yes (access given; later shown not cost-effective)‡	Few committed costs (no new infrastructure). Health displaced from other patients by funding a large population is irrecoverable for as long as the decision stands, and withdrawal once widely prescribed is difficult, so a wrong yes is unlikely to be reversed quickly.	Small population, so health displaced elsewhere is small; few committed costs. Once the comparator is generic, the cost of a wrong yes rises unless the decision is reviewed.	Irrecoverable investment in biomarker testing and training (partly offset if the infrastructure serves other therapies); health displaced elsewhere while funded.	Irrecoverable investment in specialist delivery infrastructure. As a one-off therapy with high upfront cost, the health displaced at the point of treatment cannot be recovered even if the decision is later reversed.	Few committed costs, but health displaced elsewhere by funding a large prevalent population is irrecoverable and accrues quickly; widespread use would make reversal difficult.
Irrecoverable consequences of a wrong no (access withheld; later shown cost-effective)‡	Health forgone by a large eligible population while access is withheld. Part of the benefit (long-term CV and diabetes risk reduction) may be deferred rather than lost, but harms accumulated in the interim are irrecoverable.	Acute condition: patients who would have benefited cannot do so later, so the health forgone is permanent for each patient affected.	Health forgone by patients with advanced disease while access is withheld, including survival gains that cannot be recovered later.	Irreversible health effects, including survival, for relevant patients without access.	HRQoL forgone by young people with moderate to severe disease during the research period. Irrecoverable, but not permanent: access can follow a positive result, and existing comparators remain available.
Who can act:	Manufacturer & clinical trialists: Design and fund RWE collection to an agreed specification; report non-modeled outcomes (e.g., cancer) from existing sources; consider a price that keeps expected INHB non-negative under a wider weighting of scenarios. HTA agency / health system payer: State explicitly how	Manufacturer: Provide all available patient-level and external-control data; offer an outcome-based arrangement where outcomes are routinely observed; accept a price review when the comparator becomes generic. HTA agency / health system payer: Weight scenarios explicitly	Manufacturer: Continue and report trial follow-up to an agreed schedule; share patient-level data to allow independent crossover adjustment; agree price terms that adjust with mature overall survival. (Trial design choices such as crossover determine how far the uncertainty can be resolved.) HTA agency / health system payer: Define the overall survival maturity	Provide the best available external-control analyses with transparent selection-bias adjustment; offer an outcome-based arrangement (durable response is observable) and/or a price that brings expected INHB to non-negative; co-fund a research-setting RCT if OIR is pursued.	Undertake or fund the larger, longer RCT, including an updated HRQoL measure; agree a price at which interim access has non-negative expected INHB, or accept that access is restricted to the trial. HTA agency / health system payer: Offer managed access/CED with a clear commitment to adjust the price upwards if the trial favors the technology; specify the HRQoL

	Example A	Example B	Example C	Example D	Example E
	scenarios are weighted; specify the RWE data plan, review date and exit; ensure capacity to reassess and to act on a negative result, including withdrawal, given the scale of use.	and transparently; pre-specify a review triggered by generic entry of the comparator; enable an outcome-based arrangement where routine data allow; avoid requiring research that cannot recruit.	and analysis that will trigger review, and commit capacity to reassess then; consider whether the biomarker infrastructure has value beyond this indication; specify continuity arrangements for treated patients if access is later narrowed.	HTA agency / health system payer: Weight scenarios explicitly; make an outcome-based arrangement available, with the data to operate it; if OIR is considered, confirm a trial can be delivered and that ENBS is positive; reassess if comparator real-world data improve.	measure and review date; if the manufacturer declines, consider a publicly funded RCT, and set any reduced price on the research-indifference principle rather than as a penalty.
Considerations for decision making	- Under the most plausible assumptions, average cost-effectiveness is positive, with moderate population net health benefits, reflecting an incremental cost-effectiveness ratio close to the threshold. But this changes for other scenarios, which suggest not recommending. This is also reflected in decision uncertainty. - Under the most plausible assumptions, more research is good value and feasible with access (RWE). Other scenarios suggest the value of research is less (EVPI lower, as the medicine is very unlikely to be cost-effective). - The consequences of error are material on both sides: a wrong yes displaces health across the system at scale and is hard to reverse; a wrong no forgoes health in a large eligible population. The more weight on the alternative scenarios, the more the former dominates. - A price reduction raises expected INHB under every scenario and so shifts the balance towards access.	- Under the most plausible assumptions, expected cost-effectiveness is positive but only just below the threshold, so minimal population net health benefits. But this changes for other scenarios, under which there would be negative net health benefits. This is also reflected in decision uncertainty. - If more weight is given to the most plausible scenario, expected INHB is positive and the intervention should be recommended without further research or discount. The more weight given to the alternative scenarios, the more expected INHB becomes negative, and recommendation would require a lower price. - EVPI is low due to the small population, and research would only be worth doing under some less plausible assumptions. Given the small population, recruitment to a new trial would be slow, so results would not be available for a lengthy period. - Some uncertainty would be resolved without research, though in a direction unfavorable to the new medicine.	- Under the most plausible assumptions, average cost-effectiveness is positive with positive population net health benefits. But this changes for other scenarios under which there would be negative net health benefits. This is also reflected in decision uncertainty. - Moderate EVPI with further research feasible over 2–3 years at reasonable cost. - Irrecoverable consequences arise on both sides: committed costs of biomarker testing and training with a wrong yes; survival forgone with a wrong no. Because adoption costs are partly irrecoverable, managed access needs a price that reflects them. - Need to balance population net health benefits from immediate access (which will depend on how the assumptions are weighted), the proportion of uncertainty expected to be resolved by continued follow-up and the irrecoverable consequences on each side. Price negotiation will impact this balance.	- If the most plausible incremental cost-effectiveness ratio is the focus, a decision not to recommend is the most likely result. - Weighting the various scenarios, some of which suggest an incremental cost-effectiveness ratio below the threshold, may alter this decision. - A price discount or outcome-based arrangement could influence these considerations. A discount here brings expected INHB to non-negative; it is not an uncertainty discount. - An RCT with access will be challenging due to expected difficulty of recruitment; it would also be costly and probably not economic under the most plausible assumptions. - Irrecoverable consequences arise on both sides: health effects of withholding access, and committed infrastructure and treatment costs of providing it. - Real-world evidence is unlikely to resolve uncertainty due to expected continued problems with comparator evidence with access.	- Based on most plausible assumptions there is a stark difference between the health effects of immediate access (large and negative) and the value of research (moderate to large). - Depending on how other scenarios regarding key non-parameterized assumptions are weighted, the incremental cost-effectiveness ratio may be below the threshold but only just, so population expected INHB will be small but EVPI will remain moderate to large. - A price discount would reduce the expected loss from immediate access but, given its scale, would need to be substantial to make expected INHB non-negative; it does not change the high value of research. - Further research would be of high value and is considered feasible with access, economic and without major concerns about irrecoverable consequences on either side. - Because irrecoverable consequences are small on both sides, the incentive case for not recommending where research is refused is stronger here than in B–D.

	Example A	Example B	Example C	Example D	Example E
		- The consequences of error are asymmetric here: a wrong no imposes permanent losses on individual patients with an acute condition, while a wrong yes displaces little health given the population size. - There would need to be an explicit assessment of expected INHB based on weighted scenarios, probably in the context of price negotiation.		The focus of the decision is likely to rest on the weighting of the scenarios and price negotiation.	
Indicated option(s) from the decision menu (Section 5.2)	Most plausible assumptions weighted highly: <i>recommend with managed access/CED</i> (RWE-based), with pre-specified review and price adjustment in either direction. As weight moves to the other scenarios: <i>do not recommend</i>, unless a price reduction restores non-negative expected INHB, in which case managed access/CED remains available.	<i>Recommend at the submitted price</i> if the most plausible scenario dominates; otherwise recommend at a negotiated price that keeps weighted expected INHB non-negative, or <i>recommend with an outcome-based payment arrangement</i>. In either case, pre-specify review on generic entry of the comparator. No research condition: research is neither worthwhile nor feasible.	<i>Recommend with managed access/CED</i>, time-limited to overall survival maturity (2–3 years), with pre-specified review and price adjustment in either direction; a price reduction may be needed to reflect the irrecoverable infrastructure costs. If weight moves to the other scenarios: managed access at a price that keeps expected INHB non-negative; otherwise <i>do not recommend</i>. OIR is not indicated: the research is follow-up of an existing trial, which restricting use would not help.	Most plausible assumptions weighted highly: <i>do not recommend</i> at the submitted price; <i>recommend with an outcome-based payment arrangement</i>, or at a reduced price that restores non-negative expected INHB. If weight moves to the other scenarios so that a trial becomes worthwhile (ENBS > 0): <i>recommend only in research (OIR)</i>, since irrecoverable costs make routine adoption hard to undo and an RCT is more likely to recruit when access is restricted to it.	At the submitted price: <i>recommend only in research (OIR)</i>, with a manufacturer-funded RCT, as expected INHB of routine use is negative. If the manufacturer agrees a price giving non-negative expected INHB and undertakes the RCT: <i>recommend with managed access/CED</i>, with assurance of upward price adjustment after the trial if warranted. If the manufacturer declines research: <i>recommend at a reduced price in lieu of refused research</i> (research-indifference price), or at a price reduction sufficient to fund a publicly funded RCT. With neither price reduction nor research: <i>do not recommend</i>.

† For each set of assumptions, gives the probability that the decision that is correct under those assumptions (the option with the highest expected net health benefit) is in fact incorrect. For a choice between two options this typically reaching around 50% where the incremental cost-effectiveness ratio is close to the threshold. At that point it can be higher or lower than 50% where the distribution of INHB is markedly skewed, so that the decision based on expected value is wrong more or less often than not but by smaller margins, or where more than two options are compared. Ranges for the most plausible and alternative assumptions are stated on this basis and are stylized.

‡ Irrecoverable consequences include health forgone on either side of the decision in the period before it is revisited, and committed costs (e.g., infrastructure, training) that cannot be recouped on reversal (Section 5.1).

Option names in italics follow the decision menu in Section 5.2. INHB, incremental net health benefit; EVPI, expected value of perfect information; ENBS, expected net benefit of sampling; OIR, only in research; CED, coverage with evidence development.

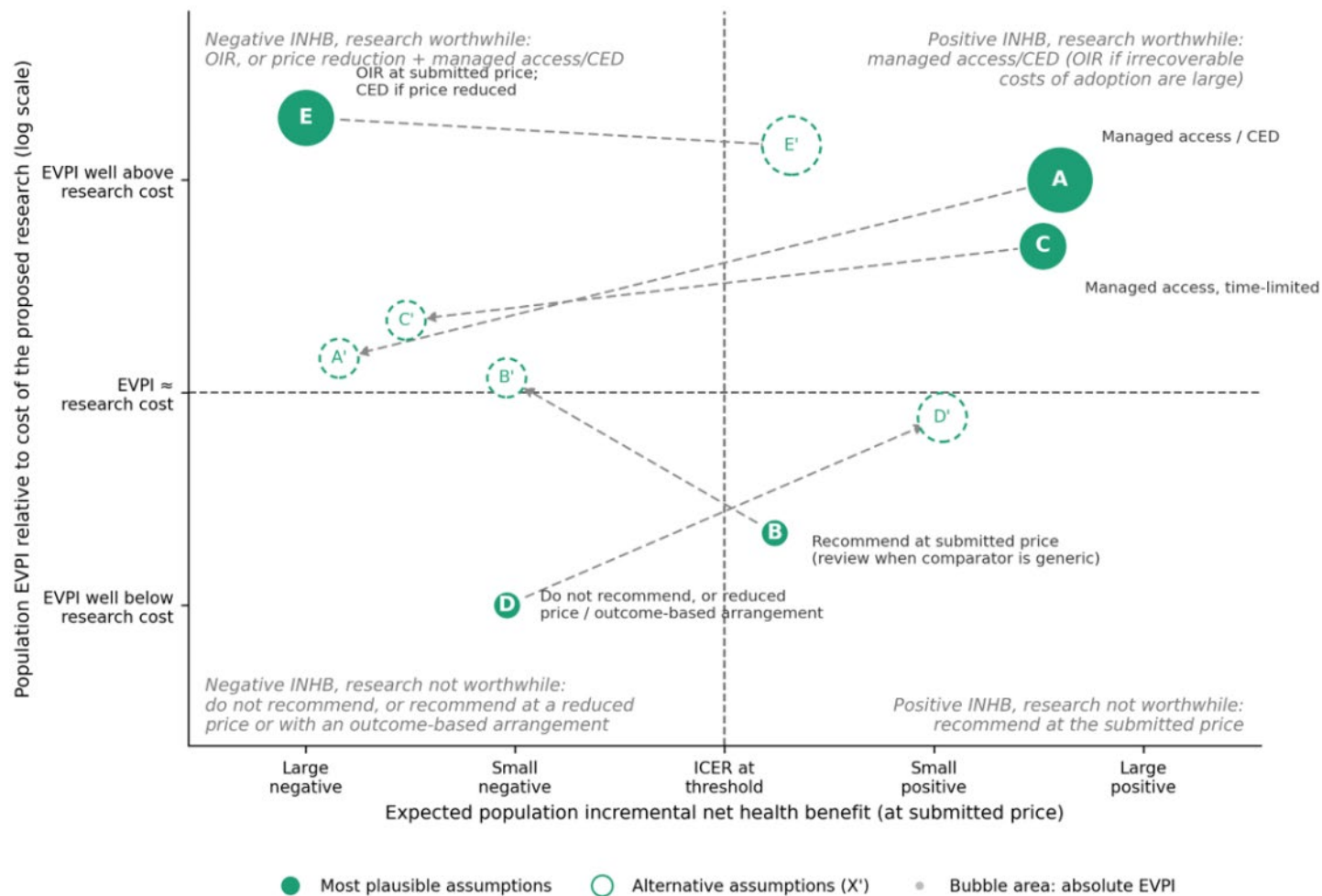


Figure 5.1. Stylized examples positioned by expected population INHB and by EVPI relative to the cost of research

Note: positions are ordinal and stylized, derived from the qualitative descriptions in Table 5.2. Filled bubbles show the most plausible assumptions and hollow bubbles (A'–E') the alternative assumptions; the arrows show the swing from non-parameterized uncertainty. Bubble area indicates absolute EVPI. Quadrant labels summarize the decision logic under Tier 1 ([Section 5.3](#)); irrecoverable consequences on each side and feasibility of research with access then determine the choice within a quadrant (Table 5.2).

Several points emerge from the examples and the figure.

- **A (obesity medicine):** a large population makes both kinds of error costly, and research is feasible alongside access, pointing to managed access/CED.
- **B (rare surgical complication):** research is neither worthwhile nor feasible, and a wrong no is permanent for patients with an acute condition. This favors recommendation at a price consistent with the weighted scenarios, or an outcome-based arrangement, with review when the comparator becomes generic.
- **C (oncology medicine):** managed access is indicated, but the irrecoverable costs of biomarker infrastructure and the limits that crossover imposes on what follow-up research can resolve both bear on the price.
- **D (advanced therapy):** the most plausible assumptions point to rejection or an outcome-based arrangement. OIR becomes relevant only if alternative scenarios make a trial worthwhile, as restricting access would likely make a trial feasible.
- **E (biologic for skin condition):** research is of high value and feasible but expected net health benefit is negative under the most plausible assumptions. OIR is therefore indicated unless a price reduction makes managed access viable; if research is refused, a reduced price in lieu of research, or rejection, applies.

The examples also make it explicit that responsibility for reducing decision uncertainty is shared. Manufacturers and clinical trialists shape how resolvable uncertainty is, through choices in trial design (for example, crossover, single-arm studies or short follow-up), through willingness to fund further research and through price. HTA agencies shape it through making options such as managed access and outcome-based arrangements available, through transparent weighting of scenarios, and through the capacity to reassess and act when evidence arrives.

5.8 Summary

Decision options are responses to the expected cost of being wrong, not to the amount of uncertainty. Reversibility of a decision and recoverability of its consequences are different properties: the first is asymmetric and favors a negative recommendation, while the second holds on both sides within the period before a decision is revisited.

The decision menu runs from recommendation at the submitted price, through price adjustment, disclosure of evidence already held, managed access and only-in-research, to rejection. Which option fits depends on how far research would resolve the uncertainty, what it would cost, how long it would take, and what burden it places on patients and services. Where the uncertainty is worth resolving and research cannot be agreed, a price adjustment is the fallback, and the basis on which it should be derived is set out in the discussion. Where a probabilistic analysis is incomplete, Table 5.1 supplies the fallback reasoning by source of uncertainty. The availability of these options is jurisdictional and the constraints on them are acknowledged in the discussion. Five stylized examples (Table 5.2 and Figure 5.1) illustrate how technologies with different uncertainty profiles map to different

options, and how both manufacturers and HTA agencies can act to reduce the cost of uncertainty.

6. Communicating Uncertainty to Support HTA Decisions

In brief. Uncertainty can support good decisions only when it is communicated clearly. Decision makers need to understand how confident they can be in the results, what uncertainties could change the recommendation, and what the consequences could be if another option later proves better. No single number or graph tells the whole story. Reports should explain both how likely the decision is to change and what that change could mean. The same findings should be presented at an appropriate level of detail for analysts, HTA committees, patients and the public, while keeping the underlying results and meaning consistent. Patient-facing information should begin with the recommendation and what it means for patients, explain uncertainty in familiar terms, and clearly distinguish uncertainty about treatment's benefits and uncertainty about whether those benefits justify the resources needed to provide it.

6.1 Introduction

Economic evaluations typically report a central estimate, such as an incremental cost-effectiveness ratio, and some indication of its variability. Reporting only the base-case results, while moving the uncertainty analysis to an appendix or summarizing it in a single cost-effectiveness acceptability curve (CEAC), can give a misleading impression of precision. HTA organizations increasingly expect quantitative results to be accompanied by a clear statement of the limitations of the evidence and of the confidence that can be placed in the conclusions (61). In a survey conducted by one HTA organization, stakeholders reported that sensitivity analyses were generally presented clearly but wanted greater explanation of how structural and parameter assumptions affected the results, with communication tailored to the needs of different audiences (62).

Communication should therefore be clear about how uncertain the results are, where that uncertainty comes from and whether it could affect the recommendation. Consistent with Sections 3–5, uncertainties should be represented in the analysis wherever feasible. Reporting should focus most on uncertainties that could affect the decision, rather than those that simply produce the largest changes in a one-way sensitivity analysis. When analytical capacity is limited, priority should be given to uncertainties most likely to affect the recommendation, while others should still be recorded in the uncertainty assessment ([Section 3.5](#)) (63).

Communicating uncertainty serves three related purposes. First, it supports decision making by showing how much confidence can be placed in the results, the consequences if the selected option proves not to be optimal, and which evidence or assumptions could

change the conclusion. Second, it informs priorities for further evidence generation by identifying what additional information would be most useful, whether it can realistically be obtained, and what uncertainty it could resolve ([Section 5](#)). Third, it supports transparency by showing how the recommendation was reached, which uncertainties were considered, and how relevant stakeholder input, including evidence from patients and clinicians, informed the assessment. This transparency is important because the consequences of an HTA recommendation may differ across affected groups. Uncertainty in the economic evaluation is one input to an HTA decision and should be considered alongside clinical, ethical, implementation, and other relevant considerations.

6.2 Principles for Communicating Uncertainty

The implication of uncertainty depends on the decision context. The same probability that a technology is not cost-effective may carry different implications depending on the severity and rate of progression of the condition, the level of unmet need, the availability of alternatives, and whether a decision can be reversed if further evidence changes the assessment (6). Delaying access may have limited consequences for a slowly progressing condition with effective alternatives but could result in irreversible health loss for a rapidly progressing condition with few or no alternatives ([Section 5.1](#)). The relevant question is therefore how the available decision options, including funding with further evidence generation or at a lower price, address the uncertainty and consequences.

Two distinctions are important for maintaining consistency with the population health objective described in [Section 2](#). First, uncertainty and variability are different concepts. Uncertainty concerns what is known about expected treatment effects, whereas variability describes differences in outcomes among individual patients (18). An individual patient's willingness to accept risk relates primarily to variability and was addressed in HEMA's first report (1). Second, the consequences of a recommendation may fall on different groups. Some consequences of accepting greater uncertainty may be borne by people other than those receiving the technology, including future patients and people whose care is affected by the resources committed. These perspectives may not be directly represented in deliberation. Rather than adjusting the analysis for the risk attitudes of decision makers or patients ([Section 2.4](#)), reports should make the expected consequences for affected groups explicit.

The following three principles apply to the reporting of uncertainty:

1. **Decision relevance.** Reporting should begin with the decision question, including the alternatives, population, perspective, time horizon, and the threshold, threshold range, or other benchmark used to judge value. The probability of cost-effectiveness should not be used as a stand-alone decision rule. Informal cut-offs, such as 80%, have no clear decision-theoretic basis. The probability should be reported together with expected net benefit and expected loss ([Section 6.4](#); Figure 6.1).

2. **Comprehensive and transparent reporting.** Reporting should address uncertainty captured through model parameters and uncertainty that remains outside the probabilistic analysis, including potential bias, limitations in the relevance or transferability of the evidence, assumptions needed to address evidence gaps, and unresolved modeling choices. Reports should state how each uncertainty is represented, as recorded in the uncertainty assessment ([Section 3.5](#)), and whether relevant relationships between model inputs have been preserved. Precise results should not conceal these limitations.
3. **Tailoring to the reader.** The depth, language, and format of reporting should reflect the needs of different users, while the underlying evidence, results, and interpretation remain consistent across audiences ([Section 6.3](#)).

6.3 Users of Uncertainty Information and a Layered Reporting Approach

A single uncertainty analysis may need to serve users with different roles, technical backgrounds and information needs. Analysts and reviewers require sufficient detail to assess the model, assumptions, probability distributions, relationships between inputs, diagnostics, scenarios, and the uncertainty assessment described in [Section 3](#). Committee members require sufficient methodological detail to understand and challenge the analysis, and a clear account of the option with the highest expected net benefit, the uncertainties that could affect the recommendation, and the consequences of making the wrong decision.

Other users have different needs. Payers and price negotiators need information on the financial implications of the decision, including uncertainty in budget impact (a separate analysis from cost-effectiveness), the reversibility of funding arrangements, and the options for further evidence generation. Where price is negotiated separately from the HTA recommendation, presenting decision uncertainty and expected loss across a range of prices can inform negotiations over price, evidence requirements, or outcome-based agreements. Clinicians require information on the implications for eligibility, treatment sequencing, monitoring, and implementation. Patients and caregivers need to understand what uncertainty means for access, outcomes, safety, and any conditions attached to a recommendation. They may also contribute evidence and perspectives that inform the assessment ([Sections 3.6.3](#), [3.8](#), and [6.7](#)).

This report recommends presenting the same analysis at three levels of detail, with clear links between them and consistent results, terminology, and interpretation (Table 6.1). The evidence producer would typically provide the technical report, the review group or assessment team would prepare material for committee deliberation, and the HTA organization would prepare the public summary and maintain consistency across the three layers. Box 1 introduces a worked example, and [Appendix A](#) presents the same analysis at all three levels.

Layer 1: Public and decision summary. This layer should be written in plain language and published with the recommendation. Where feasible, it should be shared with patient and caregiver contributors before the committee meeting.

Layer 2: Committee decision brief. This layer should provide the information required for deliberation.

Layer 3: Technical report. This layer should provide sufficient detail for independent review and, where feasible, replication. It should include the minimum reporting items shown in [Section 6.6](#), and the further detail shown in Table 6.1.

Table 6.1. Information Reported at Each Layer

Information	Layer 1: Public and decision summary	Layer 2: Committee decision brief	Layer 3: Technical report
Decision and results	The decision under consideration and the options assessed; the recommendation and its practical implications for patients, including eligibility, any conditions, and whether access could change	Expected net health or monetary benefit for each option, with incremental costs and health effects reported separately	Results across relevant thresholds, comparators, and time horizons; comparison with other models addressing the same decision problem where such models are available
Decision uncertainty and its consequences	The possible consequences if further evidence were to indicate a different decision, including consequences for patients receiving the treatment and for others whose care may be affected	Uncertainty across the relevant threshold range and, where price is under negotiation, across a range of prices; the expected loss associated with choosing an option that is not optimal	The number of simulations and evidence that results are stable
Sources of uncertainty	The main uncertainties and whether they could change the recommendation; what the evidence shows about benefits and harms, presented separately from whether the treatment represents good value for money	The uncertainties most likely to affect the recommendation, and whether they are likely to resolve over time, could be reduced through feasible research, or are likely to persist; uncertainty about clinical benefit reported separately from uncertainty about value for money	The full uncertainty assessment (Section 3.5); probability distributions and the ranges used in any deterministic analysis, with their sources; correlations between inputs; data sources; elicitation methods; diagnostic checks

Information	Layer 1: Public and decision summary	Layer 2: Committee decision brief	Layer 3: Technical report
Uncertainty outside the probabilistic analysis	What the reported figures cover and what they do not	A concise summary of the principal assumptions and important uncertainties that have not been parameterized or remain outside the analysis, including the likely direction of effect where this can be judged credibly, and the scenarios that could change the decision	Structural scenarios and any weights applied to them
Further evidence and decision options	Whether further evidence would be useful, what evidence is needed, who could collect it, what participation would require of patients, and what findings would be sufficient to inform reassessment; the circumstances under which the recommendation would be reviewed	The consequences and feasibility of available policy responses, including the further evidence required and whether it can be obtained	Value of information results where relevant (Section 4.5)
Who is affected	How evidence and perspectives from patients, clinicians, and other stakeholders were considered	The implications for patients, groups underrepresented in the evidence, and implementation, with consequences identified for the groups most likely to bear them	Results for relevant subgroups, including groups underrepresented in the evidence where data allow

Box 1. A Worked Example of Layered Reporting

A new oncology medicine for advanced disease (Example C in [Table 5.2](#)) is compared with standard care for about 3,000 patients a year. The evidence is a single randomized trial of 450 patients with a median follow-up of 22 months; overall survival data are about 60% mature and 55% of the control arm crossed over to the new medicine after progression. Eligibility requires a biomarker test that is not yet in routine use. Continued follow-up of the trial is feasible while the medicine is available.

6.4 Standard Outputs and Their Limitations

Common outputs for communicating uncertainty include the cost-effectiveness plane, CEAC, and tornado diagram. Each provides useful information, but none gives a complete account of uncertainty when presented alone. The cost-effectiveness plane shows the joint distribution of incremental costs and health effects from probabilistic analysis. Because it represents pairwise comparisons, it becomes less useful when several comparators are relevant.

The CEAC shows the probability that each option is cost-effective across a range of cost-effectiveness thresholds (64). It does not show the consequences of choosing an option that later proves not to be optimal. Expected loss addresses this limitation by showing the magnitude of the potential loss associated with decision uncertainty and should therefore accompany the CEAC (65). When several options are compared, the cost-effectiveness acceptability frontier (CEAF) should also be presented because the option with the highest probability of being optimal is not necessarily the option with the highest expected net benefit. Where useful, the distribution of net benefit rankings can provide additional information about the range of possible outcomes.

Figure 6.1 shows the two outputs together for the example in Box 1. At the threshold used here, the new medicine has the highest net benefit in only 30% of the probabilistic simulations, so the probability that recommending it is wrong is about 70%. Its expected incremental net health benefit (INHB) is nevertheless positive, at 0.10 QALYs per patient, because the distribution is right-skewed, as is typical when overall survival is extrapolated from immature data: in most simulations the medicine is slightly worse than standard care, but in a minority the survival gain is large. The expected loss from choosing the new medicine is 0.14 QALYs per patient, versus 0.24 QALYs per patient from choosing standard care; the difference is the expected INHB. The probability alone would therefore point to the option with the lower expected net benefit. Both figures are conditional on the uncertainties included in the probabilistic analysis; survival extrapolation and crossover adjustment are examined separately as scenarios.

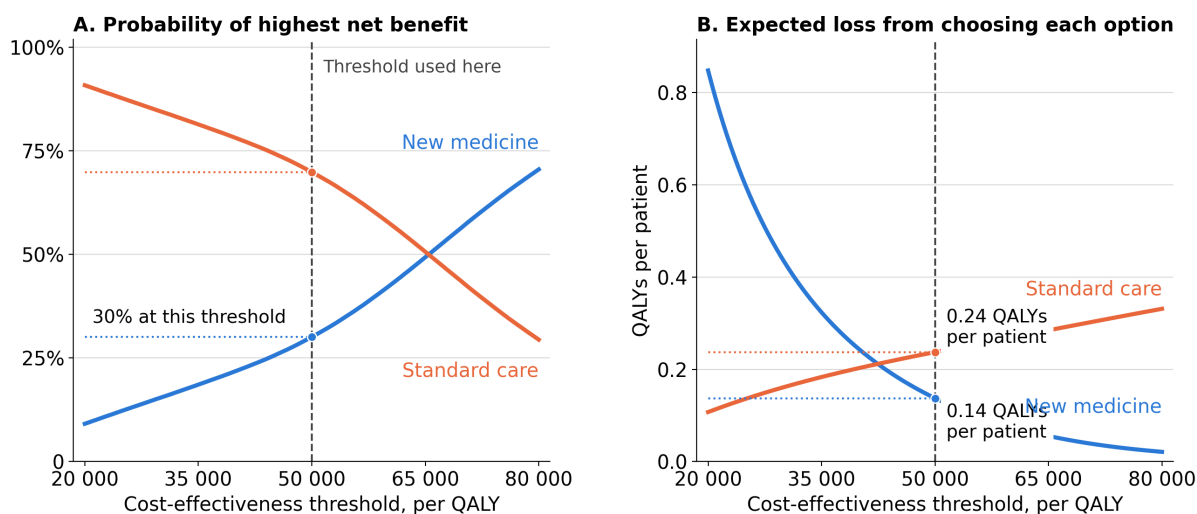


Figure 6.1. Probability that each option has the highest net benefit, and expected loss from choosing each option, across a range of cost-effectiveness thresholds.

The vertical dashed line marks the threshold used in the example; the horizontal guides show the values read at that threshold. The figures are illustrative and are not drawn from an appraisal.

The relevant threshold or threshold range should be stated explicitly because these results depend on it (66). At the same time, presenting a CEAC prominently may draw attention away from uncertainties that remain outside the probabilistic analysis.

Tornado diagrams based on one-way sensitivity analysis are familiar and easy to interpret but should be used cautiously. They examine inputs individually, may rely on poorly justified ranges, and generally do not preserve correlations between inputs ([Section 4.3](#)). They can still be useful for identifying inputs that warrant closer examination or for showing that results are insensitive to a plausible concern. Their findings should, however, be interpreted alongside analyses that account for joint uncertainty, including probabilistic analysis and, where appropriate, EVPI or EVPPI.

No single output should determine how uncertainty is interpreted. Reports should combine outputs that describe the probability of an incorrect decision, the consequences of that decision, and the uncertainties most likely to affect the recommendation, while also making clear what remains outside the quantitative analysis.

6.5 Emerging Tools and Presentation Formats

The tools described in [Section 3.5](#) can also support communication by making model assumptions, structural choices, and uncertainties outside the probabilistic analysis more visible. Newer approaches increasingly combine the probability and consequences of decision uncertainty. The HTA Risk Analysis Chart (67), for example, distinguishes the expected loss associated with uncertainty in the option with the highest expected net benefit from the additional loss incurred when another option is selected. This can help link

uncertainty to policy responses such as price agreements or further evidence generation. The Decision Uncertainty Toolkit similarly combines measures of decision risk with visual displays of downside risk and may have applications in technology appraisal (68). Interactive tools can also allow users to explore scenarios, subgroups, and value of information results, although their use in HTA remains limited.

Presentation format can affect how uncertainty is understood. Requiring readers to translate between formats, such as probabilities and frequencies, can increase cognitive burden and error, particularly for non-technical audiences (69). Icon arrays may improve understanding among people with low numeracy (70), while verbal expressions of confidence are best presented alongside quantitative information (71). Available evidence suggests that clearly communicating numerical uncertainty does not substantially reduce trust in the estimate or its source (72, 73). Presenting results as net benefit may also avoid some of the difficulties associated with interpreting ratios (74). Presentation formats should therefore be tested with their intended users and used consistently.

How best to communicate the value of further research and the consequences of uncertainty to committee members without technical training remains an important research question ([Section 7](#)).

6.6 Minimum Reporting Expectations and Committee Deliberation

CHEERS (Consolidated Health Economic Evaluation Reporting Standards) 2022 statement specifies what should be reported in an economic evaluation but does not prescribe how uncertainty should be analyzed (42). CHEERS-VOI provides additional reporting guidance for VOI, including the methods, evidence base, parameters of interest, proposed study designs, and, where relevant, the costs and expected benefits of further research (75). These guidelines improve transparency, but compliance alone does not ensure that decision makers can judge how much confidence to place in the results.

Reporting practice also remains inconsistent. An audit of NICE technology appraisals found that probability distributions were often specified without justification, correlations between parameters were rarely included, and ranges reported in guidance frequently appeared to come from deterministic analyses (76). This report therefore proposes the following minimum reporting expectations for evidence producers:

1. an uncertainty assessment table as described in [Section 3.5](#);
2. expected incremental costs, health effects, and net health benefit from the probabilistic analysis, and the deterministic result and an explanation of any material difference;
3. the probability that each option has the highest net benefit, and the expected loss across the relevant threshold range and, where relevant, across a range of prices; where more than two options are compared, the CEAF should also be reported;

4. disaggregated cost components where uncertainty in costs materially affects the result, with the main contributors identified;
5. scenarios used to examine structural or other non-parameterized uncertainties, with the rationale for their selection, and an assessment of their credibility;
6. the number of simulations and evidence that the probabilistic results are stable; and
7. a clear statement of important uncertainties omitted from the analysis, including their likely direction of effect where this can be judged credibly.

Where an analysis departs from an HTA organization's methods guidance, the report should explain the departure and its implications. Before making any further adjustment for uncertainty, committees should consider whether the same concern has already been accounted for elsewhere in the assessment, as recorded in the uncertainty assessment ([Section 3.5](#)).

These reporting expectations can support more consistent deliberation, but reporting standards alone are insufficient. General principles for deliberative HTA processes are available (58, 59), but guidance on how committees should interpret and respond to uncertainty remains limited and is identified as a priority for further work ([Section 7.3](#)).

6.7 Communication and Engagement with Patients and the Public

Patients, caregivers, and members of the public are directly affected by HTA decisions, and in many systems, contribute through submissions, consultation, advisory roles, and committee membership. Engagement should be planned early and focused on points where lived experience or broader public perspectives can inform the assessment, interpretation of uncertainty, or communication of findings ([Sections 3.6.3](#) and [3.8](#)). Patient-facing materials should, where feasible, be developed with patients and caregivers. Guidance from the HTAi Patient and Citizen Involvement Interest Group provides practical approaches to plain-language communication and emphasizes making information available early enough for patient organizations to respond (77). Guidance is also emerging on engagement in economic evaluation, including identifying care pathways, costs, consequences, assumptions, and uncertainties that may otherwise be overlooked (78).

Plain-language summaries should begin with the recommendation and its implications for patients, then explain the evidence and uncertainty supporting it. Numerical information should be retained but expressed in familiar terms. Participants in the patient workshop conducted for this report ([Appendix A](#)) found “about 7 times out of 10” easier to understand than “70% probability,” and visual displays such as shaded proportions or icon arrays may further aid understanding (70, 79). Terms such as “opportunity cost,” “decision uncertainty,” and “expected loss” should be explained or replaced, as should terms whose everyday meaning differs from their technical use. For example, “preferred option” may be interpreted

as the option preferred by patients or clinicians, when it may instead mean the option expected to generate the greatest health benefit for the resources used.

Communication should distinguish uncertainty about clinical effects from uncertainty about cost-effectiveness. Uncertain evidence should not be described as evidence that a treatment is ineffective; the uncertainty may concern the size or duration of benefit, or confidence in the estimate. Language implying blame or error should also be avoided. For example, “what could happen if further evidence supports a different decision” may be clearer and less judgmental than “the cost of being wrong.” The underlying trade-off should also be explained plainly. Resources used for one intervention are unavailable for other care, so HTA decisions consider the health gained from an intervention alongside the health that could have been generated elsewhere with the same resources.

Table 6.2 applies the guidance in this section to five situations, using the stylized examples in Table 5.2 ([Section 5.7](#)). Each row assumes one of the options indicated for that example in Table 5.2. The wording is illustrative; each summary should be developed and tested with patients.

Table 6.2. Communicating Uncertainty to Patients and the Public in Five Situations

Situation	Key Message for Patients	Example Wording	What to Avoid
A. Medicine for obesity in a large population	The medicine is available now. The trial shows short-term weight loss, but the longer-term health benefits depend on how long the effect lasts. More information will be collected during use.	“This medicine will be available now. The trial shows that it reduces weight. The analysis suggests that, if the effect lasts, it may also reduce future heart disease and diabetes. We do not yet know how long the effect will last. Information will be collected over the next three to five years, and the decision will then be reviewed. We will explain what information is collected and what, if anything, this asks of patients.”	Presenting long-term benefit as certain. Technical terms such as “durability,” “sufficient evidence,” or “confidence” without explanation.
B. Medicine for a rare complication of surgery; small population	The condition is rare and urgent. The study had no comparison group, and another comparative trial could take years. The decision must therefore consider the remaining uncertainty, the consequences of waiting, and price.	“Because this condition is rare and urgent, the study did not compare the medicine directly with current care. We therefore do not yet know how much better it is. Small studies may be unavoidable in rare conditions and do not necessarily mean that the research was poor. A new comparative trial could take many years. The remaining uncertainty therefore has to be considered alongside the consequences of waiting and the price of the medicine.”	Treating a small study as poor evidence simply because it is small. Implying that patients are responsible for the cost of treatment.

Situation	Key Message for Patients	Example Wording	What to Avoid
C. Oncology medicine for advanced disease with a moderate population	The medicine is expected to be good value on average, even though it is not the better option in most of the analysis calculations. The main uncertainty—how long the benefit lasts—is not captured by the 3-in-10 figure.	“In about 3 out of 10 of the analysis calculations, this medicine is the better use of the money available. In the others, current care does slightly better. However, when the medicine is better, the health gains can be much larger, so overall the analysis expects it to be good value. This is about value for money, not your chance of benefiting from treatment. The main uncertainty is how long the benefit lasts, and that was examined separately. The trial is continuing, so the medicine can be funded now and the decision reviewed in two to three years.”	“30% probability” without explaining what the probability refers to. Presenting 3 in 10 as the chance that the medicine works. Using “preferred option” without explaining it.
D. Advanced therapy medicinal product for a moderately rare cancer	At the current price, the analysis does not support funding. Patients need to know what could change the decision and what options are available while it is being resolved.	“This treatment may help, but at the current price, the health gained is expected to be less than the health the same resources could produce elsewhere in the health system. A new trial would be difficult and could take a long time. The treatment is therefore not recommended at the current price. We will explain what could change that decision, what other treatment options are available, and when the decision could be reviewed.”	“The cost of being wrong,” which may sound as though patients are being blamed. Language suggesting that the decision reflects the value of patients’ lives.
E. New biologic therapy for a common skin condition in young people, licensed in moderate to severe disease	The treatment improves the condition, but the quality-of-life measure may miss effects that matter to patients. A larger, longer study could address this, and patients can help decide what should be measured.	“This treatment improves the condition, but the way we measure quality of life may not capture everything that matters to you, such as sleep, confidence, school, or work. This does not mean these effects are absent; it means the current measure may miss them. A larger and longer study could provide better information. At the current price, the treatment would be available only within that study. Patients should help decide which outcomes the study measures.”	Treating an effect that was not measured as though it does not exist. Describing uncertainty as evidence that the treatment does not work.

Patient-facing summaries should describe who is represented in the evidence and which groups were poorly represented or absent from the main studies. They should explain how this affects confidence in the findings and whether eligibility criteria, access restrictions, or requirements for further evidence may affect some groups differently. Sample size should be reported but not treated as a proxy for relevance or quality. Small studies are more affected by sampling uncertainty, whereas large studies may still have limited applicability if

their participants differ from those who will receive the intervention. In rare diseases, small studies may be unavoidable and should not automatically be interpreted as inadequate research.

Where funding is conditional on further evidence generation, patients should be informed about what evidence will be collected and who is paying for it, what participation may require, what findings would count as success, what happens if the uncertainty remains at the end of the period, how reassessment could affect future access, and whether continuation of treatment for existing patients may differ from access for future patients. Additional burdens, including visits, travel, data collection, or privacy concerns, should also be made clear.

Reports should also recognize consequences for patients that the analysis may not capture. Waiting can carry an emotional cost and a loss of time that cannot be recovered, and some people may deteriorate or become too unwell to receive treatment. A negative or restricted recommendation may lead people to seek treatment through private payment, clinical trials, or compassionate use, each with its own cost and delay, and access may come to depend on ability to pay. Where such consequences are relevant, they should be described in the committee brief and the public summary, even when they cannot be quantified.

Complex documents may be especially difficult for people experiencing fatigue, cognitive effects, or other consequences of illness or treatment. HTA organizations may also provide training for patient and public contributors who wish to engage more deeply with economic analyses, while recognizing the time and support this requires.

Members of the public can contribute a broader perspective when questions concern societal values, distribution, or the allocation of collectively funded resources. Citizens' juries and other deliberative approaches can help explore broader societal trade-offs, particularly where decisions involve contested values or distributional consequences (80). Their usefulness depends on a clear link between deliberation and decision making, and they can be resource intensive. They should complement, not replace, engagement with patients and caregivers directly affected by the decision.

6.8 Summary

This section recommends presenting the same uncertainty analysis at three levels of detail, with consistent results, terminology, and conclusions across audiences ([Appendix A](#)). No single output provides a complete account of uncertainty. CEACs should be accompanied by expected loss and, when more than two options are compared, by the CEAF. Reports should also identify important uncertainties that remain outside the quantitative analysis and meet the minimum reporting expectations set out in [Section 6.6](#). Patient-facing materials should lead with the recommendation and its implications, explain who is represented in the evidence, describe the main uncertainties and what further evidence could address them,

and avoid language that implies blame or error. Where feasible, they should be shared with patient and caregiver contributors before the committee meeting.

7. Discussion and Recommendations

In brief. Uncertainty should be considered in terms of how it could affect an HTA decision and what the consequences might be. The response depends on where uncertainty comes from, whether it can be reduced, whether more research is practical and worthwhile, and what could happen if a decision is made now or delayed. Options may include providing access now, changing the price, providing access while more evidence is collected, or not funding the treatment. The analysis should clearly explain what was considered, what remains uncertain, and how that uncertainty affected the decision.

Improving practice will require better methods, clearer reporting, and enough staff, expertise and resources within HTA systems to analyze uncertainty and revisit decisions when new evidence becomes available. Patients and caregivers can help identify important uncertainties, plan further evidence collection, and explain what decisions could mean in people's lives. Reports should also make clear when uncertainty, or the consequences of a decision, may affect groups of population differently.

7.1 Key Messages

Economic evaluation for HTA asks whether the expected lifetime health gained by patients receiving a new technology exceeds the expected health forgone elsewhere in the health system to fund it. Uncertainty does not change this objective, but it changes what an assessment must provide. It is not enough to report whether a technology is cost-effective on average. Decision makers also need to know how likely a recommendation based on current evidence is incorrect, what the consequences would be, and whether further research could reduce the uncertainty at a cost worth paying. These are decision questions, not about statistical inference, which focuses on whether evidence is sufficiently compelling to reject a null hypothesis (9). Two quantities are therefore central: *decision uncertainty*, the probability that the recommendation based on expected incremental net health benefit is incorrect — and *the cost of uncertainty* — the population health expected to be forgone because the decision must be made using current evidence ([Section 2](#)). The importance of uncertainty depends on whether it could affect the recommendation, when considered with other sources (9, 26). These may include sampling uncertainty, bias, limited relevance or transferability, extrapolation beyond observed data, evidence gaps that require assumptions, and uncertainty about how the model represents the decision problem. Standard errors describe sampling uncertainty in the study from which an estimate was derived, but say little about bias, applicability, or model structure. Methodological and reporting shortcomings should be corrected where possible and should not be treated as uncertainty. A structured uncertainty assessment can document these different sources and how each has been addressed ([Section 3](#)).

PSA is the main approach for combining characterized uncertainties (52). It samples uncertain inputs jointly to generate a distribution of expected incremental net health benefit, from which decision uncertainty and its expected consequences can be estimated. DSA serves a more limited diagnostic purpose. The strength of a PSA depends on how well its distributions, correlations, and assumptions reflect the underlying evidence. Some important uncertainties, including bias, limited relevance, and structural choices, are still rarely incorporated directly. In practice, PSA therefore often provides only a partial representation of uncertainty. Reports should state clearly what has been included and what remains outside the analysis so that the results are not interpreted as more precise than the evidence allows ([Section 4](#)).

The response to uncertainty should reflect the consequences of an incorrect decision and the extent to which the uncertainty can be reduced. Funding a technology that later proves not to be cost-effective may displace health elsewhere in the system. Delaying or denying a technology that later proves cost-effective may also result in health loss for patients who would otherwise have benefited. Available responses include recommending the submitted price, price adjustment, managed access, an only-in-research arrangement, and recommending that the technology not be funded. Which option is appropriate depends on the source of uncertainty, whether it can be resolved, and the consequences of acting on the current evidence ([Section 5](#)). These considerations should be reported consistently across audiences, with the probability of an incorrect decision presented alongside its consequences ([Section 6](#)) (61, 67).

7.2 Further Considerations

Uncertainty analysis should help HTA organizations judge which of the decision options is most appropriate. When further research is one of those options, economic evaluation can help assess whether the expected value of additional research justifies its cost. This requires consideration of whether the research is feasible with or without routine access to the intervention, how long it would take, and how much of the uncertainty a particular study could resolve.

The approaches for analyzing and communicating uncertainty described in this report assume a defined decision problem. However, HTA increasingly deals with technologies that develop rapidly. Software and digital technologies may continue to evolve after an assessment has been completed. Devices may undergo repeated modifications, so the version supported by the available evidence may differ from the one used in practice by the time further evidence is available. Regenerative medicines and gene therapies can involve substantial upfront costs while evidence on durability remains immature, making the timing of payment and the consequences of early adoption particularly important. The same principles apply, but the timing of evidence and reassessment becomes more important. Suitable responses may include time-limited recommendations, planned reassessment, registry-based evidence generation, or agreements tied to a defined version of the

technology. The feasibility and burden of repeated data collection should also be considered.

A negotiated lower price is a legitimate response to uncertainty because it improves expected cost-effectiveness and may reduce both decision uncertainty and the value of obtaining further information. It does not, however, generate evidence or resolve uncertainty about treatment effects. Recommending immediate access without research, even where further evidence generation is feasible and timely, leaves the underlying uncertainty unresolved and carries it into subsequent assessments of other interventions for the same indication.

Institutional capacity places practical limits on these approaches. HTA organizations and external assessment groups have finite resources for additional analysis and reassessment. This matters for any decision that depends on returning to an appraisal when new evidence becomes available. Commissioned research is not always completed, generated evidence is not always reported, and the organizations responsible for assessing uncertainty, negotiating price, and making funding decisions are not always the same. These constraints should inform how decision options are designed and implemented, but they should not limit the framework to what every system can currently support.

Patients, caregivers, and members of the public contribute to uncertainty analysis and receive its findings. Their experience can identify outcomes or care pathways missing from a model, inform assumptions where evidence is limited, and help determine whether proposed evidence generation is feasible and what participation would require. Their input is particularly useful early in the assessment, while the decision problem, model structure, and outcomes can still change, and it should continue through the communication of the recommendation and any conditions attached to it. Patient-facing materials should be tested with intended users and made available early enough for patient organizations to respond. When further evidence is requested, it should also be clear which uncertainties the research is intended to address and how the resulting evidence could inform the decision.

The report does not resolve how different levels of tolerance for uncertainty should influence HTA decisions. A patient with a serious condition and few alternatives may be willing to accept greater uncertainty than an organization making decisions for a population. The position taken here is that economic evaluation should continue to focus on expected population net health benefit, without adjusting the analysis for the risk attitudes of decision makers or individual patients. The consequences for different groups should nevertheless be made explicit so that differences in tolerance for uncertainty can be considered openly during deliberation.

Uncertainty and the consequences of HTA decisions may also differ across population groups. Evidence may be less direct for groups that are underrepresented in clinical studies, including older adults, people with multiple conditions, and some racial and ethnic

groups, leaving greater uncertainty about outcomes in those populations ([Section 3.2](#)). Eligibility criteria, access restrictions, and requirements attached to further evidence generation may also affect groups differently. Requiring additional evidence can itself raise equity concerns if the populations enrolled in that research do not reflect those who will ultimately receive the intervention. Formal distributional methods, including aggregate distributional cost-effectiveness analysis (81), are beyond the scope of this report. At a minimum, assessments should describe who is represented in the evidence, identify important uncertainty for groups for whom the evidence is limited, and make clear which groups are most likely to experience the consequences of the decision.

7.3 Suggestions for Further Work

Several methodological and practical questions fall outside the scope of this report. Some reflect limitations of the evidence base and methods; others arise because approaches reported in this report are not yet used routinely in HTA. In particular, full integration of uncertainty into probabilistic analysis remains difficult in practice. Where this is not possible, the report provides a structured way to identify and reason about what remains outside the analysis, but this does not replace fuller integration where credible methods are available.

There is no agreed standard for judging whether a probabilistic analysis adequately represents the uncertainty relevant to the decision. Further guidance is needed on which uncertainties should be included, how distributions and correlations should be specified, and how important omissions should be identified and reported. A practical checklist or good practice standard could support more consistent review.

Methods are available to incorporate sources of uncertainty, such as bias, limited relevance, extrapolation, and alternative model structures into the PSA, including approaches that assign weights to alternative assumptions. These methods are rarely used in HTA. Further work is needed to establish when they are appropriate, how they should be implemented, and what additional analytical and review burden they create.

Guidance remains limited on how HTA committees should interpret uncertainty and use it in deliberation, including how to avoid accounting for the same uncertainty more than once. Empirical research on how committees currently use uncertainty analyses, and which forms of presentation influence deliberation, could inform more consistent practice.

More evidence is needed on how best to communicate decision uncertainty, expected loss, and the value of further research to audiences without technical training. Evaluation should focus on whether different formats improve understanding and support meaningful participation in the decision process.

Methods for incorporating lived-experience expertise into uncertainty analysis remain underdeveloped. Further work should consider how such input should be elicited, how it should be considered alongside other evidence, how it should be reported, and whether it changes the uncertainties identified or the way they are represented in the analysis.

Further work is needed on whether differences in tolerance for uncertainty should be represented formally in economic evaluation and, if so, how. There is also a need for clearer approaches to explaining these differences during HTA deliberation and in public reporting.

Methods are needed to show how decision uncertainty and the consequences of an incorrect decision differ across population groups. Further work should also consider how these differences can be communicated to committees and the public, and whether distributional approaches can be linked with decision uncertainty and value-of-information methods.

The appropriate division of analytical responsibility between evidence producers and HTA review teams remains unclear and will vary across jurisdictions. Further work is needed to define minimum expectations for each, including the level of completeness required in uncertainty analysis and the resources needed to support reassessment, further evidence generation, and other decision options described in [Section 7.2](#).

Technologies that change rapidly pose additional challenges for uncertainty analysis and reassessment. Approaches, such as valuing learning over time, linking recommendations to defined technology versions, and scheduling reassessment, require further methodological development and evaluation before they can be used routinely.

7.4 Recommendations

The recommendations below assume that the full range of decision options is available. Where the ability to require or support further research is limited, decision makers have fewer options for managing uncertainty, which may reduce the opportunity to improve population health. The report makes the following recommendations.

Report expected net health benefit alongside the incremental cost-effectiveness ratio. Expected incremental net health benefit should be presented at both individual and population levels to show the expected health impact of adopting the technology and to support comparison with the population expected value of perfect information.

Keep the focus on decision uncertainty and its drivers. Uncertainty analysis should address the decision uncertainty facing the HTA organizations and its consequences, and the implications for the available decision options.

Identify and report important sources of uncertainty. Sampling uncertainty, bias, transferability, extrapolation, model structure, and patient-relevant outcomes missing from the model should be identified and characterized using methods appropriate to each source. For each uncertainty, reports should state why it arises, what part of the analysis it affects, how it has been represented, and whether further evidence could reduce it.

Correct what can be corrected. Coding errors, inadequate reporting and avoidable methodological shortcomings should be addressed during the assessment and should not be carried forward as sources of uncertainty.

Use PSA to quantify joint decision uncertainty. PSA should incorporate as many characterized sources of uncertainty as can be represented credibly, including correlations between model inputs. Reports should state clearly which sources are included, and which remain outside the probabilistic analysis.

Communicate PSA results consistently. Report cost-effectiveness acceptability curves and, when necessary, the cost-effectiveness acceptability frontier, together with measures of the consequences of uncertainty. At a minimum, the latter should include the population EVPI. When scenario analyses are used to examine non-parameterized uncertainty outside the PSA, the PSA results should be reported for each scenario.

Support greater integration of uncertainty into PSA. HTA organizations should consider greater use of methods that allow sources of uncertainty commonly examined through scenario analyses to be incorporated into probabilistic analysis. These may include model averaging, bias adjustment, and structured expert elicitation, while recognizing their methodological and practical limits.

Consider both the probability and the consequences of an incorrect decision. The consequences of funding a technology that may prove not to be cost-effective and of delaying or declining a technology that may prove cost-effective should both be considered. This includes the health consequences of acting on current evidence, the consequences of delay, and the expected value, feasibility, and timing of further research.

Let the uncertainty analysis inform the choice of decision option. Decisions about immediate access, price adjustment, or access linked to further research should reflect the sources and consequences of uncertainty. Where further research is considered, the assessment should address whether it could resolve important sources of uncertainty, whether it is feasible, and whether its expected value justifies the cost, delay, and burden involved.

Present a consistent account of uncertainty across audiences. Technical reports, committee materials, and public summaries may differ in depth and language, but the underlying results and interpretation should remain consistent. Each should explain the recommendation, the main uncertainties that could affect it, the consequences if another decision option proves better, and any important uncertainty that remains outside the probabilistic analysis.

Involve patients and caregivers in plans for further evidence generation. When further evidence is proposed, patients and caregivers should help shape what information is collected and how it is collected. They should be told what participation involves, which uncertainties the research is intended to address, how the resulting evidence could affect the decision, and what may change when the decision is reassessed.

Invest in the capacity to undertake this work. HTA organizations should consider investing in capacity building, training and resources to support use of the methods and considerations recommended in this report.

These recommendations emphasize transparency and traceability while allowing for differences in HTA processes and institutional capacity. They are intended to help HTA organizations understand what is uncertain, why it matters to the decision, how the remaining uncertainty should inform deliberation and the choice of decision option.

References

1. McQueen R, Wailoo A, Tsuchiya A, Cipriano L, Guertin J, Sanders G, et al. Defining appropriate benefits for economic evaluation of health care technologies. Health Economics Methods Advisory; 2026.
2. Rome BN, Kesselheim AS. Raising Medicaid Rebates For Drugs With Accelerated Approval. Health affairs (Project Hope). 2021;40(12):1935-42.
3. National Institute for Health and Care Excellence. NICE technology appraisal and highly specialised technologies guidance: the manual. London: NICE; 2022 2026 Mar 31.
4. National Institute for Health and Care Excellence. Pembrolizumab with axitinib for untreated advanced renal cell carcinoma: appraisal consultation document. London: NICE; 2020.
5. 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.
6. 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.
7. Drummond MF, Sculpher MJ, Claxton K, Stoddart GL, Torrance GW. Methods for the economic evaluation for health care programmes. 4th ed. Oxford: University Press; 2015.
8. Johannesson M, Weinstein MC. On the decision rules of cost-effectiveness analysis. Journal of health economics. 1993;12(4):459-67.
9. Claxton K. The irrelevance of inference: a decision-making approach to the stochastic evaluation of health care technologies. Journal of health economics. 1999;18(3):341-64.
10. Weinstein M, Fineberg H, Elstein A, Frazier H, Neuhauser D, Neutra R, McNeil B. Clinical decision analysis. Philadelphia: WB Saunders; 1980.
11. Claxton K. Exploring uncertainty in cost-effectiveness analysis. Pharmacoeconomics. 2008;26(9):781-98.
12. Jiao B, Hsieh YL, Li M, Verguet S. Value-Based Pricing for Drugs With Uncertain Clinical Benefits. Health economics. 2025;34(4):780-90.
13. Zivin JG. Cost-effectiveness analysis with risk aversion. Health economics. 2001;10(6):499-508.
14. Arrow K, Lind R. Uncertainty and the evaluation of public investment decisions. In: Diamond P, Rothschild M, editors. Uncertainty in Economics: Academic Press; 1978. p. 405-21.
15. Sandmo A. Discount rates for public investment under uncertainty. International Economic Review. 1972;13(2).

16. Cherbonnier F, Gollier C. Risk-adjusted social discount rates. *The Energy Journal*. 2022;43(4):1-24.
17. Brouwer WB, Culyer AJ, van Exel NJ, Rutten FF. Welfarism vs. extra-welfarism. *Journal of health economics*. 2008;27(2):325-38.
18. 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.
19. 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.
20. Kennedy-Martin T, Curtis S, Faries D, Robinson S, Johnston J. A literature review on the representativeness of randomized controlled trial samples and implications for the external validity of trial results. *Trials*. 2015;16:495.
21. Rothwell PM. External validity of randomised controlled trials: "to whom do the results of this trial apply?". *Lancet*. 2005;365(9453):82-93.
22. Loree JM, Anand S, Dasari A, Unger JM, Gothwal A, Ellis LM, et al. Disparity of Race Reporting and Representation in Clinical Trials Leading to Cancer Drug Approvals From 2008 to 2018. *JAMA oncology*. 2019;5(10):e191870.
23. Gianfrancesco MA, Tamang S, Yazdany J, Schmajuk G. Potential Biases in Machine Learning Algorithms Using Electronic Health Record Data. *JAMA Intern Med*. 2018;178(11):1544-7.
24. Obermeyer Z, Powers B, Vogeli C, Mullainathan S. Dissecting racial bias in an algorithm used to manage the health of populations. *Science*. 2019;366(6464):447-53.
25. All-Party Parliamentary Group on Sickle Cell and Thalassaemia. No one's listening: an inquiry into the avoidable deaths and failures of care for sickle cell patients in secondary care. London: Sickle Cell Society; 2021.
26. Grimm SE, Pouwels X, Ramaekers BLT, Wijnen B, Knies S, Grutters J, Joore MA. Development and Validation of the TRansparent Uncertainty ASsessmentT (TRUST) Tool for Assessing Uncertainties in Health Economic Decision Models. *Pharmacoeconomics*. 2020;38(2):205-16.
27. Bojke L, Claxton K, Sculpher M, Palmer S. Characterizing structural uncertainty in decision analytic models: a review and application of methods. *Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research*. 2009;12(5):739-49.
28. Strong M, Oakley J, Chilcott J. Managing structural uncertainty in health economic decision models: a discrepancy approach. *Journal of the Royal Statistical Society: Series C*. 2011;61(1):25-45.
29. Lavelle TA, Kent DM, Lundquist CM, Thorat T, Cohen JT, Wong JB, et al. Patient Variability Seldom Assessed in Cost-effectiveness Studies. *Medical decision making : an international journal of the Society for Medical Decision Making*. 2018;38(4):487-94.

30. Grimm SE, Pouwels X, Ramaekers BLT, Wijnen B, Otten T, Grutters J, Joore MA. State of the ART? Two New Tools for Risk Communication in Health Technology Assessments. *PharmacoEconomics*. 2021;39(10):1185-96.
31. Haber NA, Wood ME, Wieten S, Breskin A. DAG With Omitted Objects Displayed (DAGWOOD): a framework for revealing causal assumptions in DAGs. *Annals of epidemiology*. 2022;68:64-71.
32. Round J, Kirwin E, van Katwyk S, McCabe C. Improving Transparency of Decision Models Through the Application of Decision Analytic Models with Omitted Objects Displayed (DAMWOOD). *PharmacoEconomics*. 2024;42(11):1197-208.
33. Hatswell AJ, Baio G, Berlin JA, Irs A, Freemantle N. Regulatory approval of pharmaceuticals without a randomised controlled study: analysis of EMA and FDA approvals 1999-2014. *BMJ Open*. 2016;6(6):e011666.
34. Downing NS, Aminawung JA, Shah ND, Krumholz HM, Ross JS. Clinical trial evidence supporting FDA approval of novel therapeutic agents, 2005-2012. *JAMA : the journal of the American Medical Association*. 2014;311(4):368-77.
35. Phillippo DM, Ades AE, Dias S, Palmer S, Abrams KR, Welton NJ. Methods for Population-Adjusted Indirect Comparisons in Health Technology Appraisal. *Medical decision making : an international journal of the Society for Medical Decision Making*. 2018;38(2):200-11.
36. Hernán MA, Robins JM. Using Big Data to Emulate a Target Trial When a Randomized Trial Is Not Available. *Am J Epidemiol*. 2016;183(8):758-64.
37. National Institute for Health and Care Excellence. NICE real-world evidence framework. London: NICE; 2022 23 June 2022.
38. Canada's Drug and Health Technology Agency. Guidance for reporting real-world evidence. Ottawa: CADTH; 2023.
39. Haute Autorité de Santé. Real-world studies for the assessment of medicinal products and medical devices. Saint-Denis La Plaine: HAS; 2021.
40. Drummond M, Barbieri M, Cook J, Glick HA, Lis J, Malik F, et al. Transferability of economic evaluations across jurisdictions: ISPOR Good Research Practices Task Force report. *Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research*. 2009;12(4):409-18.
41. Parkin D, Briggs A, Abangma G, Lloyd A, Devlin N. Uncertainty around Health State Values Used in Cost-Effectiveness Analysis: How It Arises and How to Deal with It. *Medical decision making : an international journal of the Society for Medical Decision Making*. 2026;46(2):130-43.
42. Husereau D, Drummond M, Augustovski F, de Bekker-Grob E, Briggs AH, Carswell C, et al. Consolidated Health Economic Evaluation Reporting Standards 2022 (CHEERS 2022) Statement: Updated Reporting Guidance for Health Economic Evaluations. *Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research*. 2022;25(1):3-9.
43. Doble B, Tan M, Harris A, Lorgelly P. Modeling companion diagnostics in economic evaluations of targeted oncology therapies: systematic review and methodological checklist. *Expert Rev Mol Diagn*. 2015;15(2):235-54.

44. Guyatt GH, Oxman AD, Schünemann HJ, Tugwell P, Knottnerus A. GRADE guidelines: a new series of articles in the Journal of Clinical Epidemiology. *Journal of clinical epidemiology*. 2011;64(4):380-2.
45. Karam SG, Zhang Y, Pardo-Hernandez H, Siebert U, Koopman L, Noyes J, et al. ROBVALU: a tool for assessing risk of bias in studies about people's values, utilities, or importance of health outcomes. *BMJ (Clinical research ed)*. 2024;385:e079890.
46. Lash TL, Fox MP, MacLehose RF, Maldonado G, McCandless LC, Greenland S. Good practices for quantitative bias analysis. *International journal of epidemiology*. 2014;43(6):1969-85.
47. Bloem LT, Vreman RA, Peeters NWL, Hoekman J, van der Elst ME, Leufkens HGM, et al. Associations between uncertainties identified by the European Medicines Agency and national decision making on reimbursement by HTA agencies. *Clin Transl Sci*. 2021;14(4):1566-77.
48. Bojke L, Soares M, Claxton K, Colson A, Fox A, Jackson C, et al. Developing a reference protocol for structured expert elicitation in health-care decision-making: a mixed-methods study. *Health technology assessment (Winchester, England)*. 2021;25(37):1-124.
49. O'Hagan T. SHELF: the Sheffield Elicitation Framework version 2.0. Sheffield, UK: University of Sheffield; 2010.
50. Grigore B, Peters J, Hyde C, Stein K. EXPLICIT: a feasibility study of remote expert elicitation in health technology assessment. *BMC Med Inform Decis Mak*. 2017;17(1):131.
51. Oakley J, Ren S, Forsyth J, Gosling J, Wilson K, Latimer N, et al. NICE DSU Technical Support Document 26: expert elicitation for long-term survival outcomes. London: NICE; 2025.
52. Briggs AH. Handling uncertainty in cost-effectiveness models. *PharmacoEconomics*. 2000;17(5):479-500.
53. Briggs A, Sculpher M, Claxton K. Decision modelling for health economic evaluation. Oxford: Oxford University Press; 2006.
54. Bilcke J, Beutels P, Brisson M, Jit M. Accounting for methodological, structural, and parameter uncertainty in decision-analytic models: a practical guide. *Medical decision making : an international journal of the Society for Medical Decision Making*. 2011;31(4):675-92.
55. Trentham-Dietz A, Alagoz O, Chapman C, Huang X, Jayasekera J, van Ravestein NT, et al. Reflecting on 20 years of breast cancer modeling in CISNET: Recommendations for future cancer systems modeling efforts. *PLoS Comput Biol*. 2021;17(6):e1009020.
56. Strong M, Oakley JE, Brennan A. Estimating multiparameter partial expected value of perfect information from a probabilistic sensitivity analysis sample: a nonparametric regression approach. *Medical decision making : an international journal of the Society for Medical Decision Making*. 2014;34(3):311-26.
57. Rothery C, Strong M, Koffijberg HE, Basu A, Ghabri S, Knies S, et al. Value of Information Analytical Methods: Report 2 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(3):277-86.

58. Thorn J, Coast J, Andronis L. Interpretation of the Expected Value of Perfect Information and Research Recommendations: A Systematic Review and Empirical Investigation. *Medical decision making : an international journal of the Society for Medical Decision Making*. 2016;36(3):285-95.
59. Farmer C, Barnish MS, Trigg LA, Hayward S, Shaw N, Crathorne L, et al. An evaluation of managed access agreements in England based on stakeholder experience. *International journal of technology assessment in health care*. 2023;39(1):e55.
60. McElwee F, Cole A, Kaliappan G, Masters A, Steuten L. Alternative Payment Models for Innovative Medicines: A Framework for Effective Implementation. *Applied health economics and health policy*. 2025;23(4):535-49.
61. Trowman R, Powers A, Ollendorf DA. Considering and communicating uncertainty in health technology assessment. *International journal of technology assessment in health care*. 2021;37(1):e74.
62. Versteeg J, Goettsch W, Mantel-Teeuwisse A, Ollendorf D, Leopold C. A multi-stakeholder survey on communicating cost-effectiveness uncertainties to stakeholders: a case study of ICER. *Health Affairs Scholar*. 2026;4(3):qxag009.
63. Botwright S, Munira SL, Megiddo I, El Desouky N, Ngwira LG, Turner HC, Teerawattananon Y. Uncertainty in Economic Evaluation: A Pragmatic Guide for Health Technology Assessment (HTA) Agencies in Resource-Constrained Settings. *PharmacoEconomics*. 2026.
64. Fenwick E, Claxton K, Sculpher M. Representing uncertainty: the role of cost-effectiveness acceptability curves. *Health economics*. 2001;10(8):779-87.
65. Alarid-Escudero F, Enns EA, Kuntz KM, Michaud TL, Jalal H. "Time Traveling Is Just Too Dangerous" but Some Methods Are Worth Revisiting: The Advantages of Expected Loss Curves Over Cost-Effectiveness Acceptability Curves and Frontier. *Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research*. 2019;22(5):611-8.
66. Siverskog J, Henriksson M. On the role of cost-effectiveness thresholds in healthcare priority setting. *International journal of technology assessment in health care*. 2021;37:e23.
67. 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.
68. Wiggins M, Varughese M, Rafferty E, van Katwyk S, McCabe C, Round J, Kirwin E. The decision uncertainty toolkit: Risk measures and visual outputs to support decision making during public health crises. *PLoS One*. 2025;20(10):e0332522.
69. Kale A, Kay M, Hullman J. Visual Reasoning Strategies for Effect Size Judgments and Decisions. *IEEE Trans Vis Comput Graph*. 2021;27(2):272-82.
70. Galesic M, Garcia-Retamero R, Gigerenzer G. Using icon arrays to communicate medical risks: overcoming low numeracy. *Health Psychol*. 2009;28(2):210-6.
71. Padilla LMK, Powell M, Kay M, Hullman J. Uncertain About Uncertainty: How Qualitative Expressions of Forecaster Confidence Impact Decision-Making With Uncertainty Visualizations. *Front Psychol*. 2021;11:579267.

72. van der Bles AM, van der Linden S, Freeman ALJ, Mitchell J, Galvao AB, Zaval L, Spiegelhalter DJ. Communicating uncertainty about facts, numbers and science. *R Soc Open Sci*. 2019;6(5):181870.
73. van der Bles AM, van der Linden S, Freeman ALJ, Spiegelhalter DJ. The effects of communicating uncertainty on public trust in facts and numbers. *Proc Natl Acad Sci U S A*. 2020;117(14):7672-83.
74. Andrews B. Economic Evaluation Under Ambiguity and Structural Uncertainties. *Journal of Benefit-Cost Analysis*. 2024;15(3):456-76.
75. Kunst N, Siu A, Drummond M, Grimm SE, Grutters J, Husereau D, et al. Consolidated Health Economic Evaluation Reporting Standards - Value of Information (CHEERS-VOI): Explanation and Elaboration. *Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research*. 2023;26(10):1461-73.
76. Andronis L, Barton P, Bryan S. Sensitivity analysis in economic evaluation: an audit of NICE current practice and a review of its use and value in decision-making. *Health technology assessment (Winchester, England)*. 2009;13(29):iii, ix-xi, 1-61.
77. Coombes MA, Morgan K, Diaz J, Oehrlein EM, Andersen H, Gilardino R, et al. Plain language summaries supporting patient involvement: lessons and guidance from HTAI Patient and Citizen Involvement Interest Group (PCIG). *International journal of technology assessment in health care*. 2025;41(1):e59.
78. Thompson ER, Wilson M, Nicholls SG, Fergusson DA, Lalu M, Smith M, et al. Making room for patients in economic evaluation: a commentary on current capacity and future directions. *Res Involv Engagem*. 2026;12(1).
79. Zikmund-Fisher BJ, Thorpe A, Fagerlin A. How to Communicate Medical Numbers. *JAMA : the journal of the American Medical Association*. 2025;334(16):1474-5.
80. Street J, Duszynski K, Krawczyk S, Braunack-Mayer A. The use of citizens' juries in health policy decision-making: a systematic review. *Social science & medicine (1982)*. 2014;109:1-9.
81. Love-Koh J, Cookson R, Gutacker N, Patton T, Griffin S. Aggregate Distributional Cost-Effectiveness Analysis of Health Technologies. *Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research*. 2019;22(5):518-26.
82. Staniszewska S, Brett J, Simera I, Seers K, Mockford C, Goodlad S, et al. GRIPP2 reporting checklists: tools to improve reporting of patient and public involvement in research. *BMJ (Clinical research ed)*. 2017;358:j3453.

Appendix A. A Worked Example of Layered Reporting

This appendix applies the three reporting layers described in [Section 6.3](#) to the case introduced in Box 1. It illustrates reporting, not analysis: the case is constructed, and the results are internally consistent but do not come from an actual appraisal. The case corresponds to Example C among the stylized examples in Table 5.2 ([Section 5.7](#)).

A.1 Layer 3: Technical Report (excerpt from the uncertainty assessment; Section 3.5)

Table A.1. Uncertainty Assessment for the Worked Example

Source of Uncertainty	How it is Represented	Direction, whether it could change the recommendation, and whether it may resolve
Extrapolation of overall survival beyond trial follow-up	Not parameterized in the probabilistic analysis; four survival scenarios, of which one is judged most plausible	If the benefit fades sooner, the incremental cost-effectiveness ratio rises above the threshold and expected INHB becomes negative. Continued trial follow-up over 2 to 3 years should resolve much of this uncertainty
Adjustment for crossover after progression	Two crossover-adjustment methods, treated as scenarios; the probabilistic analysis is rerun under each	Under the less favorable adjustment, the incremental cost-effectiveness ratio rises above the threshold and expected INHB becomes negative; the probability that not recommending is incorrect is up to about 0.40. Longer follow-up is less likely to resolve this uncertainty than the extrapolation question, but independent reanalysis of patient-level trial data may help
Effectiveness estimate (sampling uncertainty)	Distributions for the effectiveness estimate and related inputs, sampled jointly with cost and time on treatment	Symmetric; narrows as more patients are observed or more outcome events accrue during follow-up
Cost and uptake of biomarker testing	Distributions for test cost, uptake, and eligibility, sampled together	Modest effect on the incremental cost-effectiveness ratio; likely to narrow as routine-use data on testing accumulate
Quality of life after progression	Distribution around the mapped utility estimate; alternative plausible utility values examined in scenario analysis	Modest effect; mapping uncertainty may persist unless direct utility data are collected

The first two sources are the uncertainties that could change the recommendation, and neither is included in the probabilistic analysis under the most plausible assumptions. The probability and expected-loss estimates reported below therefore do not reflect them. Where feasible, the probabilistic analysis is rerun within the key scenarios so that decision uncertainty is reported conditional on each plausible structural assumption.

A.2 Layer 2: The Committee Decision Debrief

Table A.2. Layer 2 Reporting for the Worked Example

What The Committee Is Told	For This Case
Expected incremental cost and health effect per patient	28,000 and 0.66 QALYs (probabilistic means); 27,400 and 0.58 QALYs (deterministic point estimate)
Expected incremental net health benefit (INHB) at the threshold	0.10 QALYs per patient; about 1,500 QALYs over five years
Probability that the new medicine has the highest net benefit	0.30 under the most plausible assumptions; therefore, the probability that recommending is incorrect is about 0.70
Expected loss from choosing the new medicine	0.14 QALYs per patient; about 2,100 QALYs over five years
Expected loss from choosing standard care	About 0.24 QALYs per patient
Population expected value of perfect information	About 2,100 QALYs over five years
Uncertainties outside the probabilistic analysis	Survival extrapolation and crossover adjustment, with the direction of effect as set out in A.1
Decision options and what each turns on	Recommend at the submitted price: whether the extrapolation and crossover assumptions judged most plausible are accepted, despite a probability of about 0.70 that recommending is wrong. Recommend with managed access/CED, time-limited to overall survival maturity (2 to 3 years): whether the irrecoverable costs of biomarker testing and training are acceptable, with pre-specified review and price adjustment in either direction. Recommend at a reduced price: whether a lower price keeps expected INHB non-negative across credible scenarios and reflects the irrecoverable infrastructure costs. Do not recommend: the expected INHB of 0.10 QALYs per patient is forgone while access is deferred. Only in research is not indicated, because the evidence comes from continued follow-up of an existing trial

Monetary values are in generic currency units, and the threshold is 50,000 per QALY. Population figures assume 3,000 patients per year for five years and are undiscounted. The difference between the two expected losses equals the expected incremental net health benefit of 0.10 QALYs per patient. The probabilistic mean health effect exceeds the deterministic estimate because long-term survival gains are right-skewed: gains are modest in most simulations but large in a minority. The population expected value of perfect information equals the expected loss from choosing the new medicine because it is the option with the highest expected net benefit under the most plausible assumptions.

Research is feasible alongside access in this case, so declining to fund is not required to obtain the evidence. The main irrecoverable implementation costs are those of establishing biomarker testing and training, partly offset if the infrastructure serves other therapies; future medicine expenditure can stop if access changes. Figure 6.1 presents the probability and expected loss across the threshold range.

A.3 Layer 1: The Public and Decision Summary

What is being recommended

The recommendation is to make the new medicine available to adults with advanced cancer whose tumor has a particular genetic marker, while follow-up of the trial continues, with a review in two to three years. About 3,000 people a year would be eligible. The alternative is the current treatment.

What evidence shows about the medicine

One trial of 450 people found that the medicine delayed the cancer from getting worse, and that people taking it lived longer over the time they were followed. Patients have been followed for about two years so far. Side effects were similar to those of current treatment. The trial has not yet run long enough to show how much longer people live overall, and just over half of the people in the comparison group later switched to the new medicine, which makes the size of the benefit harder to estimate.

What the evidence shows about value for money

This is a separate question from whether the medicine works. In most health systems, the budget for care is set in advance, so money spent on one treatment is not available for other care. The analysis asks whether the health gained by people taking this medicine is greater than the health that could be gained by spending the same resources on other care.

The analysis ran many calculations, each using a different plausible value for the things we are unsure about. In about 3 out of 10, the new medicine was the better use of the money available. In the others it was not, usually by only a small margin. In the calculations where the medicine performed well, the health gains were much larger because some people may live substantially longer. Across all calculations, the medicine is expected to be good value on average, which is why it is recommended while more evidence is collected. These figures describe value for money, not the chance that the medicine will help an individual patient. They include only uncertainties that could be put into numbers. The biggest remaining question, how long the benefit lasts, was examined separately and is described below.

What is uncertain, and what it would mean

The main uncertainty is how long the benefit lasts after the end of the trial. If it lasts as long as the analysis assumes, the medicine is good value. If it fades sooner, it is not. A second

question is how much the comparison is affected by people in the trial switching to the new medicine.

If the medicine is funded and the benefit turns out to be smaller, the money spent would have produced more health elsewhere in the service, and other patients would lose out.

If it is not funded and the benefit turns out to be as expected, people with this cancer would have missed a treatment that helped them, and that time cannot be recovered.

Most people in the trial were younger and in better health than many of the patients who would be treated. Less is known about how well the medicine works for older people and for people with other conditions.

What could settle the question

The trial is continuing, and following the same patients for another two to three years should provide better evidence on how long the benefit lasts. Reanalysis of information already collected in the trial may also help address the question about switching. This requires no additional visits beyond those already planned and can occur while the medicine remains available. The decision would be reviewed when the results are available. Patients already receiving the medicine would not have treatment stopped solely because of the review.

Words Used Here

Term	What it means here
Value for money	Whether the health gained is greater than the health the same money would produce if spent on other care.
Uncertain	The evidence does not yet show the size or length of the benefit clearly. It does not mean that the medicine does not work.

Appendix B. How Patient Representatives Informed This Report

This appendix describes how patient representatives contributed to the report, what we heard, and how the report changed in response. It also records what did not change and why. Its structure follows the short-form checklist for reporting patient and public involvement (GRIPP2-SF) (82).

B.1 Purpose

The working group held a workshop with patient representatives while the report was still open to revision. The aim was to test whether patients and members of the public could understand the report's account of uncertainty and the language recommended for patient-facing communication. The workshop was advisory. It was not intended to elicit representative preferences, validate the methods, or ask a small group to speak for all patients.

B.2 Who Took Part and The Session

Seven patient representatives from Canada, the United Kingdom, and the United States took part in a 75-minute online session on 8 September 2026. Several served on patient advisory or expert committees associated with CDA-AMC, ICER, or NICE; others worked for patient organizations. The session was facilitated by a working group member who was not part of the writing group. Members of the writing group attended to listen and answer clarifying questions.

Participants received a short briefing note and worksheet. Four exercises addressed the consequences of uncertainty for patients; the balance between earlier access and stronger evidence; managed access, evidence generation, and reassessment; and the wording of patient-facing summaries, including alternative ways of presenting a probability. The wording tested was illustrative: the aim was to learn how best to communicate. Several participants submitted written responses after the session. We returned a summary of what we heard and the proposed changes to participants and invited them to correct it.

B.3 What We Heard

Understanding. Participants did not think the people they represent would understand the tested passages as written. The main difficulty was the word “probability”, not the number itself. Participants preferred a natural frequency stated one way (for example, “about 7 times out of 10”) and found versions that stated both sides of the fraction confusing. Several terms did not work, including “opportunity cost”, “durability”, “sufficient”, “technology”, and

“preferred option”. “The cost of being wrong” was read as blaming patients for the cost of their treatment. The passage on groups under-represented in research was judged the least successful; participants suggested first stating who was included in the studies informing the decision.

Trust. Most participants did not accept the statement that a wrong yes and a wrong no cost patients the same thing. Several also questioned the statement that funding a treatment of lower value displaces health elsewhere. They viewed this as an assumption about how budgets are managed, not a description of what happens in their health systems. Some interpreted the report’s example wording as a rationale for delaying access.

Consequences the draft had missed. Participants added the following:

- the emotional cost of waiting and its effect on mental health;
- the loss of time that cannot be recovered, and deterioration that may leave people too unwell to be treated;
- other routes to access when the answer is no, such as private payment, clinical trials, or compassionate use, each with its own cost and delay;
- recommendations that are neither yes nor no, such as restrictions that require another treatment to be tried first;
- unequal access according to ability to pay; and
- the difference between stopping treatment for people who are benefiting and not starting it for new patients.

Tolerance for uncertainty. A patient with a serious condition and no alternatives may reasonably accept much more uncertainty than the organization making the decision. Participants felt that the draft did not acknowledge this.

Requests for further evidence. When further evidence is requested, participants wanted to know exactly what evidence is needed, from whom, and which gap it would address. They also asked what happens to the decision once that evidence is available.

B.4 What Changed in the Report

Table B.1. Changes Made in Response to the Workshop

What Participants Told Us	What Changed	Where
A wrong yes and a wrong no do not cost patients the same thing	The recommendation now asks that the consequences of both decisions be considered, rather than assigning them equal weight. The report distinguishes the reversibility of a decision from the recoverability of its consequences.	Executive Summary; Sections 5.1 and 7.4
The claim that funding displaces health elsewhere was presented as fact	Opportunity cost is presented as the reasoning behind a decision, not a prediction of where money goes in a particular system.	Sections 6.7 and 7.1

What Participants Told Us	What Changed	Where
“Probability” is not understood; the number itself should be kept	Patient-facing numbers are expressed as natural frequencies stated one way, with visual displays such as icon arrays.	Sections 6.5 and 6.7 ; Table 6.2; Appendix A.3
Several terms do not work, and “the cost of being wrong” reads as blame	Guidance on terms to explain or replace in patient-facing materials, including “preferred option”; an everyday-language glossary.	Glossary ; Section 6.7 ; Table 6.2
The passage on under-represented groups was the least successful	The revised passage begins by stating who is represented in the evidence. Reports are asked to identify groups that are poorly represented and state what has been assumed about them.	Sections 3.2 and 6.7
Stopping treatment for people who are benefiting differs from not starting it	Continuity and transition arrangements are part of a workable managed access agreement. Patient-facing summaries should state what evidence will be collected and who pays for it, what participation involves, what would count as success, what happens if uncertainty remains, and whether people already on treatment will continue.	Sections 5.5 and 6.7 ; Appendix A.3
The draft’s account of consequences for patients was too narrow	Reports should describe consequences the analysis may not capture, including the emotional cost of waiting, deterioration, other routes to access, and dependence on ability to pay. The account of irrecoverable consequences covers health lost while access is withheld.	Sections 5.1 and 6.7
Patients may reasonably tolerate more uncertainty than the organization deciding	This is acknowledged, the report’s position is explained, and the question is included in the research agenda.	Sections 7.2 and 7.3
What evidence is needed, from whom, and what exactly is the gap?	An analysis supporting a request for further evidence should state which uncertainties affect the decision, what evidence could resolve them, over what period, and how each result would inform the decision. The public summary should explain what evidence collection would require from patients.	Section 5.5 ; Table 6.1
Simpler writing and a worked example	A plain-language summary at the front of the report and at the start of each section; a glossary; one worked example followed through all three layers of reporting	Plain Language Summary ; each section; Glossary ; Box 1 ; Appendix A

B.5 What Did Not Change, and Why

The report is written primarily for HTA organizations, analysts, and manufacturers. Some points raised in the workshop were therefore addressed through the report's guidance on communication rather than through changes to its technical content.

- Technical terms such as “opportunity cost” and “technology” remain in the technical sections, where they have precise meanings for the intended readers. They are defined in the glossary, and [Section 6.7](#) asks that they be explained or replaced in material written for patients and the public.
- The principle that resources used for one treatment are unavailable for other care is retained because it underpins economic evaluation for HTA. The report presents this as the analytical basis for a decision, not as a description of how any particular health-system budget is managed.
- The analysis continues to focus on expected population net health benefit without adjusting for the risk attitudes of decision makers or patients, because the consequences of accepting greater uncertainty may fall on other patients. The report asks instead that consequences for affected groups be made explicit in deliberation ([Section 7.2](#)).
- Whether an HTA organization acts on evidence once it has been produced, whether it is fair to require further evidence at all, and how access should be arranged or prioritized are questions of HTA process and policy rather than methods. They have been recorded and passed to CDA-AMC, ICER, and NICE.
- Some disagreements among participants were recorded but not resolved. These included whether a wrong yes and a wrong no can carry equal weight in principle, and whether an unresolved question about long-term safety belongs in an HTA decision or in the conversation between a patient and their clinician.

B.6 Limitations

The findings should be interpreted in light of two limitations. First, the briefing materials did not reach all participants in advance, so the briefing note was not tested as pre-read. Second, the group was small, drawn from three countries, and familiar with HTA. Difficulties in understanding therefore provide strong evidence that the wording needs revision, while apparent ease of understanding offers limited evidence that it would work for a general audience.