



2020-2023 Value Assessment Framework

January 31, 2020

About ICER

The Institute for Clinical and Economic Review (ICER) is an independent non-profit research organization that evaluates medical evidence and convenes public deliberative bodies to help stakeholders interpret and apply evidence to improve patient outcomes and control costs. Through all its work, ICER seeks to help create a future in which collaborative efforts to move evidence into action provide the foundation for a more effective, efficient, and just health care system. More information about ICER is available at <http://www.icer-review.org>.

Funding for ICER's review activities comes from government grants and non-profit foundations, with the largest single funder being the Laura and John Arnold Foundation. No funding for these activities comes from health insurers, pharmacy benefit managers, or life science companies. ICER receives approximately 21% of its overall revenue from these health industry organizations to run a separate Policy Summit program, with funding approximately equally split between insurers/PBMs and life science companies. For a complete list of funders and for more information on ICER's support, please visit <http://www.icer-review.org/about/support/>.

About this Document

This paper presents final updates to the ICER Value Assessment Framework, including refinements of its conceptual structure and modifications to the specific methods used to gather and assess evidence of different types. Separate documents describing adaptations to this framework for treatments for ultra-rare diseases and single- or short-term transformative therapies can be found on ICER's [website](#).

This update to the ICER value assessment framework builds upon ICER's experience using the 2017-2019 framework in the evaluation of drugs, devices, tests, and delivery system innovations, as well as earlier iterations of the framework. During that time ICER has actively sought the input of all stakeholders and made iterative changes to its methods and overall procedures to enhance their transparency and to improve the ability of all parties to participate meaningfully in the process. ICER has also benefitted from public comment opportunities during each framework revision cycle, including two comment periods for the 2020-2023 framework; the first being a call for open public input to propose changes to the framework, the second providing an opportunity for stakeholders to comment on proposed changes. ICER received feedback from 60 stakeholder organizations during an open call for suggested revisions, and from 41 organizations after ICER posted proposed revisions for comment in August 2019. Organizations who provided input included patient advocacy organizations, clinical societies, drug manufacturers, and payers, as well as several individual commenters. Their comments can be found [here](#), along with ICER's summary response to comments [here](#). ICER wishes to thank all of these commenters for the time and effort they put into these comments, and the many thoughtful contributions they have made.

This paper reflects this combined experience, public input, and many additional discussions with stakeholders in various settings. This finalized update to the ICER value framework and associated methods will be in place to guide reports launched during the three-year period of January 2020 through December 2023, with the next formal update cycle scheduled to begin in 2023.

Contents

1. Introduction	1
1.1. Overview	1
1.2. Overarching Purpose and Principles of the ICER Value Assessment Framework.....	1
1.3. The Population Perspective and Intended Uses of the ICER Value Assessment Framework.....	2
1.4. Conceptual Structure of the ICER Value Assessment Framework.....	3
2. Comparative Clinical Effectiveness	10
2.1. Overview	10
2.2. Scope of Clinical Effectiveness Evaluation	12
2.3. Sources of Evidence	13
2.4. Appraisal and Synthesis of Evidence.....	15
2.5. Judgment of Level of Certainty and Magnitude of Net Health Benefit: the ICER Evidence Rating Matrix™	17
3. Incremental Cost Effectiveness.....	19
3.1. Overview	19
3.2. Model Structure and Data Sources.....	20
3.3. Measures of Health Gain	22
3.4. Impact on Distribution of Health Gains	23
3.5. Perspective.....	24
3.6. Discounting	25
3.7. Patient Populations.....	25
3.8. Costs.....	26
3.9. Base-Case and Cost-Effectiveness Thresholds.....	27
3.10. Sensitivity and Scenario Analyses	27
3.11. Validation and Calibration	29
3.12. Uncertainty and Controversies	29
3.13. Health-Benefit Price Benchmarks	30
4. Potential Other Benefits or Disadvantages and Contextual Considerations.....	32
4.1 Overview	32
4.2. Specific Categories	35

5. Potential Budget Impact Analysis	39
5.1. Overview	39
5.2. Methods	40
5.3 Access and Affordability Alert.....	42
6. Procedures for Report Generation, Stakeholder Engagement, and Public Processes	43
6.1 Topic Selection	43
6.2 Timelines	44
6.3 Stakeholder Engagement.....	46
6.4 Public Meetings.....	53
References	60
Appendices.....	63
A. Glossary.....	63
B. List of Major Revisions to 2017-2019 Framework.....	67
C. Review Timelines	68
D. Cost-Effectiveness Threshold Ranges.....	71

1. Introduction

1.1. Overview

This document contains an overview and discussion of the concepts that underpin ICER's Value Assessment Framework, as well as an overview of the procedures used to conduct assessments. The Framework describes ICER's philosophy and approach assessing the value of a medical intervention at the population level, as well as the implications of its findings for practice and policy. Detailed descriptions of the technical methods ICER uses to conduct its assessments (e.g., the ICER Evidence-Based Medicine Rating Matrix, reference case for economic modeling, and stakeholder engagement guides) may be found on ICER's website and links to these materials are provided in related sections of this document.

1.2. Overarching Purpose and Principles of the ICER Value Assessment Framework

For more than 10 years ICER has been active in developing methods for evidence assessment. Evidence assessment, however, is only one component of ICER's broader effort to provide mechanisms through which all stakeholders and the general public can engage in discussions on how best to use evidence as the foundation for a more effective and sustainable health care system. A formal effort was undertaken between 2014-2015 to gain input through a multi-stakeholder advisory group on ways to define with greater detail the conceptual and methodological underpinnings of ICER reports – a "value assessment framework." ICER's first formal value assessment framework was posted in 2015, and following two years of further experience, and several rounds of public comment, an update to the framework was posted in early 2017 as the guide to ICER's reviews for 2017-2019. This most recent update has also benefited from extended discussions with stakeholders, experience over the past two years, and formal public comment. This version of the ICER Value Assessment Framework will serve as the standard for our methods and procedures for our reports beginning in 2020 through 2022, with current expectations that we will launch another formal update process in 2022 to be implemented in 2023.

Ultimately, the purpose of our value assessment framework is to form the backbone of rigorous, transparent evidence reports that, within a broader mechanism of stakeholder and public engagement, will help the United States evolve toward a health care system that provides fair pricing, fair access, and a sustainable platform for future innovation. In this effort ICER is guided by several key underlying principles. One is that we act with respect for all, in concordance with a presumption of good will on the part of all participants and stakeholders in the health care system.

ICER does not intend to target any particular interest group or organization. There are many areas in which the US health system fails to serve patients well, in which access to care is suboptimal, waste and inefficiency pose major problems, and costs to patients and the health system fail to align with added value. ICER believes that only through collaborative efforts, built upon a foundation of civil discourse and honest consideration of evidence on effectiveness and value, can lasting progress be made on behalf of patients today and those of the future.

The ethical vision inherent in ICER’s work recognizes that many choices that are made in health care – choices in clinical care, insurance coverage, pricing, payment, and allocation of resources within health systems – must address the basic reality that societal resources for health care are not unlimited, and that there will always be trade-offs and dilemmas over how to organize and pay for the services provided within a health system. Too often, these decisions are made without rigorous evidence and with little transparency. Too often, there is little chance for reflection or public engagement in managing the tensions that can arise between innovation, access, and costs. ICER’s value assessment framework seeks to place scientific methods of evidence analysis at the heart of a clearer and more transparent process. The value framework reflects our strong underlying belief that rigorous thinking about evidence can prevent the kind of waste that strains our ability to provide patient-centered care. The framework also is intended to support discussions about the best way to align prices for health services with their true added value for patients. While considering value and linking it to pricing and insurance coverage cannot solve every dilemma, nor satisfy every need, ICER believes it offers the best hope of avoiding rationing of care by the ability of patients to pay for care, and that it can promote a more dynamic, innovative health care system that will make the best use of available resources in caring for all patients.

1.3. The Population Perspective and Intended Uses of the ICER Value Assessment Framework

The ICER Value Assessment Framework describes the conceptual framework and set of associated methods that guide the development of ICER evidence reports. ICER reports are intended to support deliberation on medical policies related to health services (e.g., tests or treatments) and delivery system interventions (e.g., preventive programs, changes to the organization of medical personnel). To inform these kinds of medical policies the ICER value framework takes a “population” level perspective as opposed to trying to serve as a shared decision-making tool to be used by individual patients and their clinicians. Taking a population perspective implies that the ICER value framework seeks to analyze evidence in a way that supports population-level decisions and policies, such as broad guidelines on appropriate care, pricing, insurance coverage determinations, and payment mechanisms. A value framework intended to support decisions about the care of individual patients requires a structure that invites weighting of benefits, harms,

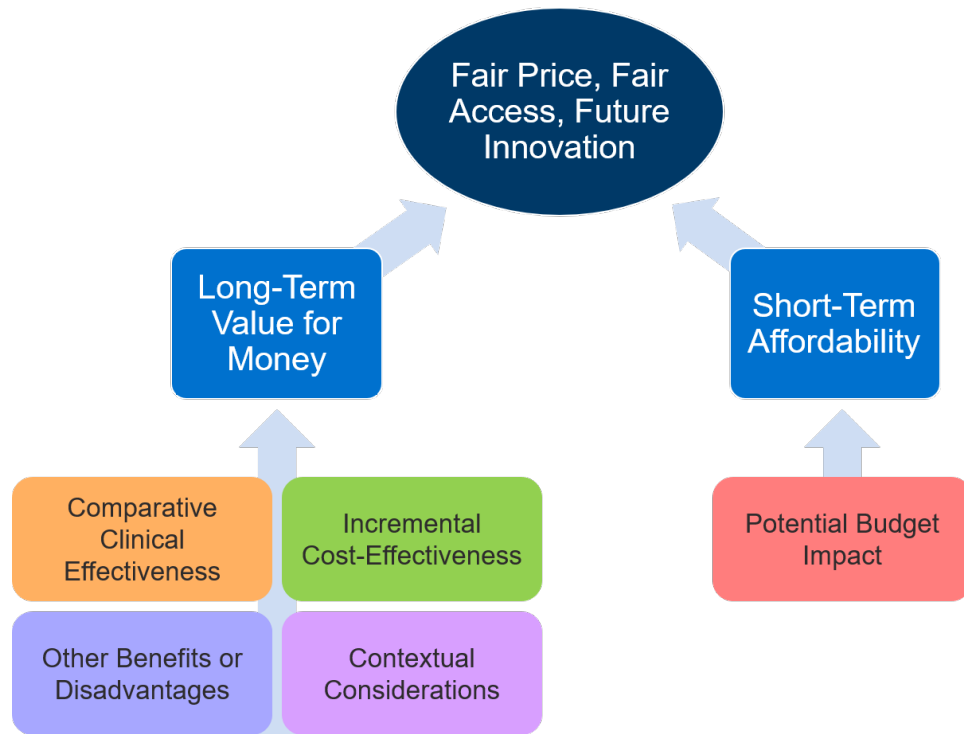
and costs from the individual patient’s perspective. There is an important need for better evidence-based shared decision-making tools for individual patients and clinicians, but this is not the primary intended purpose of the ICER value framework or of ICER reports.

Even with its population-level focus, however, the ICER value framework seeks to encompass and reflect the experiences and values of patients. Representing the diversity of patient outcomes and values in a population-level framework is difficult because there will always be an inherent tension between average findings in clinical studies and the uniqueness of every patient. There will also always be diversity in the way that patients view the balance of risks and benefits of different treatment options. The ICER value framework does not solve these tensions, but neither does it obscure them. Population-level decisions and policies have always been made by life science companies, insurers, and clinical organizations looking at evidence in the same general way. One important goal of the ICER value framework is to provide an evidence report that does a better job of analyzing the strengths and limitations of the available evidence, including what is or is not known about the variation in response to different treatments among patients with different personal and clinical characteristics. The ICER value framework also creates an explicit place and role for consideration of elements of value that are important to individual patients but that fall outside traditional clinical measures.

1.4. Conceptual Structure of the ICER Value Assessment Framework

As shown in the structure of the ICER value framework, it seeks to inform decisions that are aimed at achieving sustainable access to high-value care for all patients (see Figure 1.1 below). This goal requires consideration of two general concepts: long-term value for money and short-term affordability.

Figure 1.1. Conceptual Structure of the ICER Value Assessment Framework



Long-term Value for Money

Long-term value for money serves as the primary anchor of the ICER value framework. It is itself a concept that is comprised of multiple domains: 1) comparative clinical effectiveness, 2) incremental cost-effectiveness, 3) other benefits or disadvantages, and 4) contextual considerations. A description of how these domains are measured and integrated into an ultimate judgment of long-term value for money is described in later sections of this paper. There are several high-level points about this element of the value framework that bear highlighting here:

Long-term perspective

Even though most of the clinical data available on health care services come from studies of relatively short duration, the grounding of any evaluation of value should recognize the long-term perspective on both outcomes for patients and costs. The ICER value framework recognizes this principle by grounding the methods of incremental cost-effectiveness analysis in simulations that estimate outcomes and costs at the longest feasible time horizon, usually the full lifetime of patients. Benefits for patients and potential cost offsets for new treatments that might take many years to be seen are therefore estimated and included as a core element of the value framework.

Foundation in the evaluation of evidence on comparative clinical effectiveness

The ICER value framework is rooted in an objective evaluation of the evidence on the comparative clinical effectiveness of different care or care delivery options. This element of the framework serves as the primary source of information to inform cost-effectiveness analysis and includes a systematic review of available evidence performed according to the highest academic methodological standards. As part of the evaluation of comparative clinical effectiveness, ICER reports include a clear description of the sources of evidence, the strengths and limitations of individual studies, and a summary judgment of the net health benefit of different care options along with a statement explaining the relative certainty that the body of evidence is able to provide. The methods used by ICER in its evaluation of comparative clinical effectiveness are discussed in Section 2 of this paper and described in more detail in documents available on the ICER website (<https://icer-review.org/methodology/icers-methods/>). The ICER rating system for evidence of comparative clinical effectiveness is being updated in this iteration of the value assessment framework but its earlier incarnation was published in a peer-reviewed journal and was endorsed by the AMCP-NPC-ISPOR Comparative Effectiveness Research Collaborative.^{1,2}

Acceptance of multiple forms of evidence

Patients, clinicians, and policymakers are most interested in evidence on the comparative clinical effectiveness of care options, but this does not mean that ICER's value framework limits the type of evidence to be considered to the results of randomized controlled trials (RCTs).

When available, high-quality RCTs and systematic reviews of RCTs provide evidence that is least susceptible to many scientific biases. However, head-to-head RCTs of active comparators are uncommon, especially for interventions near the time of regulatory approval. Without direct head-to-head evidence, insights into comparative clinical effectiveness may require indirect comparisons through formal network meta-analysis. Complementing these sources of information is evidence derived from many different analytic approaches and that are available from a wide range of sources. Although more vulnerable to some important confounding biases, observational methodologies such as cohort studies, case-control studies, and long-term disease and drug registries often provide helpful evidence, particularly on longer-term outcomes. As will be described in greater detail later in this document, ICER also has a commitment to explore how "real-world" observational evidence can contribute to a more comprehensive and accurate view of the risks, benefits, and costs associated with any intervention. This commitment extends not only from using available published sources, but includes the possibility of working with life science companies, patient groups, or data aggregator companies to develop and analyze new sources of real-world evidence in a way that will meet the evidentiary standards relevant to the questions being addressed.

In short, ICER has a flexible and ecumenical approach to sources of evidence and, while stressing the importance of the rigor of clinical trial data in any assessment, the value framework and ICER's methods incorporate multiple sources and types of evidence, seeking the evidence that is most helpful in understanding the long-term net health benefits for patients of different care options.

Recognition that what matters to patients is not limited to measured “clinical” outcomes.

The inclusion of an explicit domain of value labeled “other benefits or disadvantages” demonstrates that the ICER value framework fully acknowledges that all too often what matters most to patients is poorly captured in the available clinical trial data. Sometimes this occurs because the clinical outcomes measured do not reflect what is most important to patients’ day to day quality of life. Even when trials do capture the clinical outcomes that matter most to patients, there are other aspects of the treatment regimen that have a significant impact on the overall value of the treatment. This can be related to the complexity of the treatment regimen or the impact of care options on the ability of patients to return to work, on family and caregivers, on overall public health, or on other aspects of the healthcare system or society.ⁱ The ICER value framework identifies these “potential other benefits or disadvantages” as important elements of any overall judgment on long-term value for money, and all ICER reports have separate sections in which evidence and information pertaining to these elements are presented. We describe in Section 4 of this paper a method for integrating these domains of value and Section 6 includes discussion of how these considerations are incorporated in the public deliberation and voting process at ICER meetings.

Acknowledgment of the role of contextual considerations

Decisions about the value of care options do not happen in a vacuum. How to interpret and apply evidence in clinical care, insurance coverage, and pricing, involves a complex process of integrating information on risks and benefits of treatment within a broader set of contextual considerations. These contextual issues include the severity of the condition, whether other treatments are available or soon will be, and ethical, legal or other societal priorities that are important to acknowledge as part of any discussion on value. The ICER value framework includes these elements and they are explored in a separate section of each ICER report. In addition, contextual considerations often feature prominently in the deliberation on value between independent expert committees and all stakeholders and is a central feature of the public meetings convened by ICER on each report. Linked to the discussion of “other benefits or disadvantages,” we discuss the methods used to integrate contextual considerations into ICER reports in Section 4 and describe how they are incorporated into the ICER meeting and voting process in Section 6.

ⁱ For further insight and examples a useful resource is the FasterCures and Avalere Health work on “Integrating the Patient Perspective into the Development of Value Frameworks” available at <http://www.fastercures.org/assets/Uploads/value-coverage-framework-March-2016.pdf>

Short-term Affordability

With long-term value for money being the dominant element in considerations of value, a complementary perspective is provided by including an evaluation of short-term affordability. The ICER value framework includes an explicit evaluation of the short-term affordability of different care options by analyzing the potential short-term budget impact of changes in health care expenditures with the introduction of a new test, treatment, or delivery system process. Detailed methods used to estimate potential budget impact are presented later in Section 5 of this paper.

Budget impact is a reasonable consideration within a value assessment framework because insurers work in rapid cycles with purchasers and individual subscribers, translating short-term cost projections into planned insurance premiums for the coming year. Rapid cost growth in the short-term, especially when it increases beyond anticipated inflation rates, pushes quickly upstream to purchasers and policymakers who have to make their own short-term decisions about how to find the needed resources. This may lead to decisions to increase deductibles or otherwise reduce health care benefits for employees; for example, state governments might need to consider reducing next year's education budget to find the funds to keep a Medicaid program afloat.

In addition, for provider groups that bear financial risk, budget impact analyses inform very real short-term decisions about how to allocate resources to maximize the quality of health care within a given budget. A rapid increase in costs resulting from the significant budget impact of a new drug might lead to decisions to forgo hiring of needed new staff or delay the introduction of other new services. Quite simply: budget impact, and not long-term cost-effectiveness, determines how affordable health care insurance will be in coming years and shapes what health care can be provided with the resources available.

ICER's value framework represents the conviction that keeping budget impact considerations off the table, to be factored in only *post hoc* by insurers or provider groups in ways unknown, would be a mistake. It would rob our nation of the chance to bring the public directly into the critical discussions about health care and health insurance that we need to have if we are going to achieve sustainable access for all patients to the kind of innovative new tests, treatments, and delivery system interventions that add value to their lives.

Potential budget impact analyses estimate the net budget impact across all elements of the health care system

ICER's methods have never sought to estimate the potential budget impact of treatments within "silos" of a payer budget, such as the expenses only on pharmaceuticals, devices, or hospital costs. It remains a core principle of ICER's value framework that it should evaluate both short and long-term costs across the entire health system, so that care options that might increase spending for

one type of service (e.g. drugs) while reducing other spending (e.g. hospital costs) receive full credit for cost offsets and are not penalized in any way.

At five years, the time frame for considering “short-term” affordability is stretched as far as possible without losing relevance for identifying new care options that may require special measures – in pricing, payment mechanisms, coverage criteria, or budgeting – to maintain patient access without serious financial strain throughout the health care system. Using a five-year time horizon may reduce the utility of the analysis for insurers focused on shorter budget timeframes but helps accommodate some of the important potential clinical benefits and cost offsets that may not occur immediately with the adoption of a new therapeutic option. With the primary anchor of the ICER value framework being the long-term perspective represented by long-term value for money, the time horizon for short-term affordability has been extended as far as it seems possible in order for it to serve the important purpose of informing discussions on whether special efforts need to be taken to manage the introduction of a new therapeutic option so that access and affordability can both be maintained.

Considerations for Assessments of Non-Drug Interventions

Devices

There are many important, unique aspects to the development, early evaluation, regulatory approval, and patterns of use and iterative evidence generation for devices. Therefore, although the conceptual elements of the ICER value framework remain the same for any health care intervention, the specific methods for incorporating and judging evidence will differ for devices. For example, ICER methods acknowledge the practical and ethical considerations that may make it impossible to use RCTs in the early evaluation of clinical effectiveness, while iterative changes to devices, along with the learning curve for practitioners, also raise special considerations about how to judge the available evidence. Evaluations of long-term cost- effectiveness are made challenging because of the potential for evolution of devices and the attendant changes in cost, effectiveness, and the types of patients who will be treated. These complexities are also relevant to estimations of potential budget impact, and, as noted in sections below, it is very difficult to identify the current baseline costs of all device use in the US health care system in order to calculate a growth target for a budget impact threshold. For these reasons the conceptual elements of the ICER value framework remain relevant for devices but within that framework ICER will continue to incorporate specific approaches to evidence evaluation for devices that reflect their unique features.

Tests

Similarly, different approaches to evidence evaluation are required for diagnostic interventions and tests used to monitor patients or provide information on disease prognosis. For example, the general hierarchy in the types and strength of evidence for tests is different than that for

therapeutic interventions.ⁱⁱ As with devices, tests will continue to be evaluated using the overall conceptual approach of the ICER value framework but there will be important modifications based on the distinctive nature of the evidence and the system for the development, evaluation, and use of diagnostic interventions. Further work will be needed to develop a method for estimating a threshold for potential budget impact that should trigger additional policy maker consideration of short-term affordability.

Delivery System Innovations

There are also many distinctive challenges to evaluating the evidence on the effectiveness and value of delivery system interventions. Chief among these is that in most cases a delivery system intervention will be highly variable in its implemented form across different settings, raising great questions about the generalizability of results from studies of one institution or one system of care. RCTs can be difficult to perform, increasing concerns about the internal validity of study findings. ICER will use the same general value assessment framework to guide its reviews of delivery system interventions, but as with devices and tests, some of the specific methods for judging evidence and for determining thresholds for potential budget impact analysis will reflect the unique nature of these kinds of health service innovations.

ⁱⁱSee for example the discussion of the Fryback and Thornbury evidentiary model used as part of the ICER review on cardiac nuclear imaging, coronary computed tomographic angiography, CT colonography, breast cancer screening, and diagnostic tests for Alzheimer’s disease (<https://icer-review.org/topics/>)

2. Comparative Clinical Effectiveness

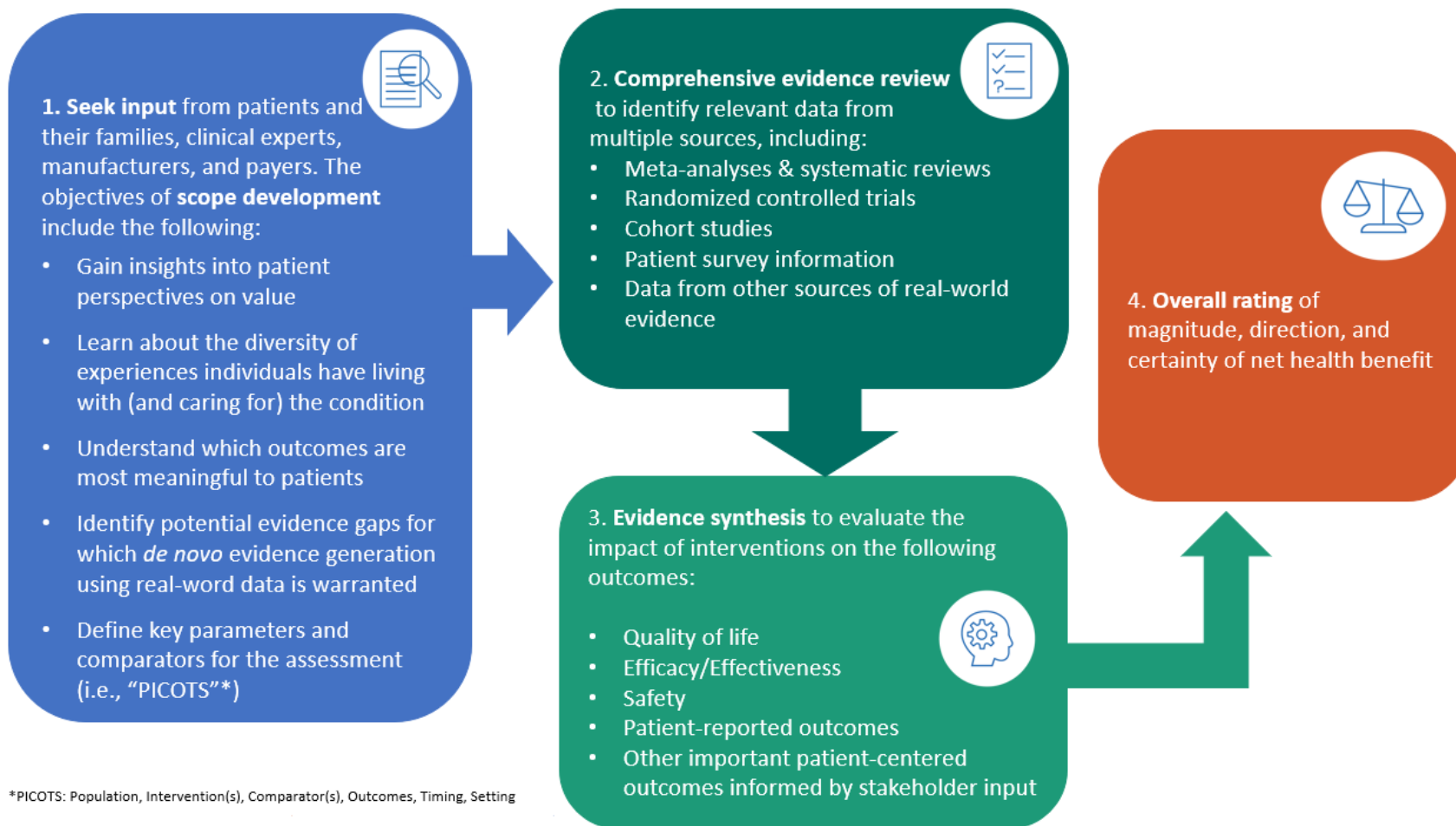
2.1. Overview

A central part of the ICER value framework is an objective evaluation of the evidence on comparative clinical effectiveness. Comparative clinical effectiveness involves weighing the benefits and harms/burdens of one treatment option versus another. The most important benefits and harms are those that are important to patients and their families/caregivers. As such, from the outset of a review ICER solicits input from patients, families, caregivers, and expert clinicians to understand the day-to-day experience of living with a condition and what outcomes it would be most important for a therapy to affect. Information on what has been learned from patient input is presented in the ICER report prior to the discussions of the evidence so that readers can interpret the evidence through the lens of patient experience.

Stakeholder input from clinicians, manufacturers, and payers is used in addition to input from patients and families to frame the questions that an ICER comparative effectiveness review attempts to answer. When evidence on patient-important outcomes is limited or unavailable, ICER will seek evidence on surrogate endpoints that might be associated with outcomes important to patients and families.

Once we have defined the scope of a review, ICER evaluates the available clinical evidence. ICER conducts a systematic review of the existing literature using established best practices for evidence synthesis. The findings of our evidence review are described in a publicly available report, which includes a description of the sources of evidence, the strengths and limitations of individual studies, an assessment of the relevancy and generalizability of the published literature for patient and provider decision making, and a summary rating of the net health benefit of different care options. ICER's approach to evaluating the comparative clinical effectiveness is summarized in Figure 2.1 and discussed in the section that follows; ICER's Methods Guide for Health Technology Assessment, available at http://icer-review.org/wp-content/uploads/2018/08/ICER-HTA-Guide_082018.pdf, describes our methods in greater detail.

Figure 2.1. Summary Process for Assessment of Comparative Clinical Effectiveness



2.2. Scope of Clinical Effectiveness Evaluation

From the inception of the evaluation, ICER examines the contextual landscape of the topic under review. ICER compiles data related to epidemiology, prognosis, standards of care, and natural history, while seeking to understand the lived experiences of patients affected by the condition. Insights from patient groups and other stakeholders, along with reviews of the evidence, inform definitions the population, interventions, comparators, outcomes, timing, and setting (PICOTS) components that anchor ICER’s evaluation of comparative clinical effectiveness. These components are described below.

- **Population:** The population that is eligible to use the intervention(s) under review. For certain topics, such as drug therapies, the population may be defined to align with current or anticipated FDA indications for that therapy. ICER also examines whether there are subpopulations for whom the relative effectiveness or safety of the intervention may vary or whether there are subpopulations for whom variations in baseline risk lead to higher or lower absolute benefits or harms. These different subgroups are highlighted in the ICER report.
- **Interventions:** Interventions may include drug therapies, medical tests, devices, and delivery system innovations, among others. When relevant, ICER may focus its review on specific attributes of an intervention (e.g., mode of administration, line of therapy, etc.). ICER’s process for selecting which interventions to review is described in Section 6 of this document.
- **Comparators:** Appropriate comparators represent alternative therapies used among the populations and settings of focus. Active comparators (i.e., non-placebo interventions) are prioritized when feasible. Relevant comparators are selected through a survey of clinical guidelines from professional societies, consultation with clinical experts and patients, and review of clinical trial designs.
- **Outcomes:** Critical to the evaluation of net health benefit of an intervention are the measures of potential benefit and harm. Health outcomes, i.e., changes in symptoms or conditions that people experience and that affect the quantity or quality of life (e.g., change in pain, quality of life, length of life) are given greater weight than intermediate outcomes (e.g., change in cholesterol). Patient-important outcomes are health outcomes that are central to ICER’s judgements of benefit and harm. When appropriate, ICER also looks for evidence on non-clinical outcomes such as resource utilization or measures of societal benefit.
- **Timing:** The minimum duration of study follow-up considered adequate to capture the outcomes of interest.

- **Setting:** The setting(s) of focus for a review may be specified (e.g., inpatient, emergency department, and/or outpatient) and ICER will state whether these settings will exclude certain study settings from consideration.

2.3. Sources of Evidence

ICER's evaluation of comparative clinical effectiveness is grounded in a systematic review of all available evidence. A systematic review identifies all relevant existing evidence using explicit, replicable methods in a way that minimizes the risk of biased selection of studies. Established best methods of systematic literature reviews are followed in order to foster transparency and facilitate reproduction of results.^{3,4} Reviews are reported in accordance with the [Preferred Reporting Items for Systematic Reviews and Meta-Analyses \(PRISMA\) guidelines](#).⁵

ICER's judgements around comparative clinical effectiveness are informed by evidence arising from multiple sources. When available, high-quality RCTs or their meta-analyses provide evidence that is least susceptible to certain scientific biases. When benefits and harms occur over the course of many years, or when harms are rare but clinically important or even catastrophic, evidence from high quality published peer-reviewed studies using observational data and methodologies such as cohort studies, case-control studies, and long-term disease and drug registries may be used. Furthermore, if important patient reported outcomes have not been collected as part of a manufacturer's clinical development program, ICER will again conduct a comprehensive literature review to identify published, peer-reviewed observational studies providing this information.

Real-World Evidence

RWE may help complement other types of evidence in assessments of comparative clinical effectiveness, in contributing to assessment of the potential other benefits of interventions, and in providing useful information to inform the assumptions of economic models. ICER has consistently sought to incorporate analysis of RWE into our reports whenever it can provide additional perspective on comparative clinical effectiveness or cost-effectiveness. In addition to searching for published RWE and seeking RWE in the grey literature, on several occasions we have collaborated with patient and other stakeholder organizations to obtain new patient and caregiver survey information when it was not available in the medical literature. Findings from this work have been included in our Evidence Reports and helped inform discussions during our Public Advisory meetings and appraisal committee votes.

RWE often has greater vulnerability to known and unknown biases that create limitations in our ability to rely on it when making judgments about relative effectiveness of different care options. Nonetheless, we understand that RCTs have their own limitations and are often inadequate to address all questions relevant to assessments of comparative clinical effectiveness. RWE can be

particularly helpful under certain circumstances such as when long-term safety of a treatment or durability of a medication’s effect is unclear. We have also emphasized how RWE can be helpful in supporting consideration of a treatment’s “potential other benefits” that lie outside traditional clinical trials. Patient-reported outcome studies and studies that capture broader patient and family effects of treatment are especially desired as they can provide evidence usually not included in clinical trials.

ICER’s use of real-world data also may include *de novo* evidence generation under certain circumstances where critical data elements are lacking. Options for generating new RWE may include conducting a patient survey using a validated patient-reported outcome (PRO) instrument or using claims data to better understand adherence and persistence. Such analyses would need to address key gaps in the evidence base and be feasible within the timelines of an ICER review. Any *de novo* analyses would also need to be transparent to all stakeholders so that all participants can engage in deliberation on their validity and relevance.

As with all evidence, ICER will assess the internal and external validity of RWE as part of a larger judgment of whether and how that evidence should be incorporated in an assessment. The process by which ICER will evaluate RWE will follow the general outline presented in our separate framework to guide the optimal development and use of real-world evidence for drug coverage and formulary decisions.⁶ Efforts will be undertaken to assure that the data are curated with input from individuals with knowledge of the nuances of the data source. Methods for adjusting for known and potential unknown confounders will be assessed, and replication of results using different methods within the same data set and/or using different data sources will be pursued. ICER will also apply best practices in real-world data analysis as described in guidelines from ISPOR and other authoritative methods bodies.⁷

Grey Literature

ICER also includes evidence from the “grey literature” as per our criteria available at <https://icer-review.org/methodology/icers-methods/icer-value-assessment-framework-2/grey-literature-policy/>. We supplement our reviews of studies from peer-reviewed publications with data from conference proceedings, regulatory documents, materials from other HTA groups, information submitted by manufacturers, and input gleaned from patients. Consideration of multiple sources of evidence helps us evaluate whether there is a biased representation of study results in the published literature and provides a panoramic understanding of net health benefit.

In summary, ICER has a flexible and inclusive approach to sources of evidence, which stresses the importance of the rigor of clinical trial while augmenting such evidence with data from other real-world or grey-literature sources.

2.4. Appraisal and Synthesis of Evidence

Assessment of Quality of Individual Studies

ICER evaluates the methodological quality of individual studies in part by applying risk of bias tools deemed appropriate for the topic under review. The quality assessment tool developed by the US Preventive Services Task Force (USPSTF) for judging the quality of clinical trials and cohort studies is one of the tools ICER commonly adapts.⁸ However, ICER believes that no single tool exists that is ideal for evaluating all possible studies included across reviews. Thus, for each review, we thoughtfully consider which quality assessment tools are most appropriate for the topic at hand and document the choice in our protocol.

When examining individual study quality, the main focus is on risk of bias and selective reporting rather than other aspects pertaining to study conduct (e.g., obtaining ethical approval or calculating sample size). Of note, ICER's assessment focuses on the internal validity of the study (i.e., how well the study is able to estimate what it set out to measure). Relevant quality issues evaluated in our assessment include selection bias (e.g. was allocation concealed?), performance bias (e.g., were patients blinded?), attrition bias (e.g., was intention to treat analysis used?), detection bias (e.g., was outcome assessment blinded?) and selective reporting (e.g., were the important outcomes measured and analyzed in the study fully reported?). ICER's assessment incorporates how particular aspects of a study may lead to biased results and states the likely direction of such bias.

For each review, the rationale for the assessments is explicitly determined *a priori* in our protocol, and the judgment on each study is provided in the appendix of each evidence report.

Synthesis of Results

ICER employs a transparent approach to evidence synthesis. Evidence is synthesized to help provide single best estimates and ranges of confidence that can help in evaluation of the comparative clinical effectiveness of interventions of interest. Syntheses also assist in understand the limitations and gaps in the evidence base.

Following the identification of studies that meet our PICOTS criteria for a given evidence review, data from the studies are abstracted, and summarized in the text and in evidence tables of the evidence report. This summary is key to understanding the existing evidence base pertaining to the interventions and comparators of interest. Any key differences between the studies in terms of study design, patient characteristics, interventions (including dosing and frequency), outcomes (including definitions and methods of assessments), patient subgroups, and study quality are evaluated and described.

ICER examines the clinical and methodological characteristics of the set of studies reporting data for each outcome of interest and for each subpopulation with the goal of aggregating the results from

the studies. When there is insufficient data or studies are judged to be too dissimilar for quantitative meta-analysis, we describe the results qualitatively in our evidence report and provide the key considerations for interpreting the results from the studies within the context of the evidence base.

When studies are sufficiently similar and report data that are appropriate for analysis, we conduct quantitative synthesis of the results across studies. Quantitative synthesis (e.g., meta-analysis) involves the use of a statistical method to pool results across multiple studies to generate the best estimate of the effect of the intervention on the outcome. In the absence of head-to-head studies comparing two interventions of interest, ICER often derives comparative evidence through quantitative synthesis methods that uses indirect comparisons (e.g., network meta-analysis, matching adjusted indirect comparisons), which may rely on common comparators or common predictors to link data from trials of the various treatments of interest.

The choice of the synthesis method ICER uses on a given topic depends on the question and the available evidence. In all reports, we provide the rationale for the choice of the synthesis method used and explicitly describe our methods.

Heterogeneity and Subgroups

ICER's reviews are not intended to guide individual shared decision making between clinicians and patients and are not able to focus on the sorts of individual patient characteristics, values, and preferences that a skilled clinician would assess in making recommendations for a specific patient.

At the population level, data often show a range of responses to therapy with various distributions, including smooth normal distributions and sharply dichotomous outcomes. Heterogeneity of this sort may be unpredictable for individual patients but will still be highlighted in ICER reports as it can affect assessment of therapies. For instance, a treatment that leads to a six-month increase in survival for all patients with no heterogeneity has different implications from a treatment that leads to a two-week increase in survival for 90% of patients and long-term cure for 10% of patients. This is true even if it is currently impossible to know which patients will achieve each outcome.

In other cases, there is heterogeneity that is knowable *a priori*, based on patient characteristics prior to treatment. This may come in the form of characteristics that are effect modifiers, such that patients respond differently to treatment based on these knowable characteristics, even when treatment has net benefits for all patients. In many other cases, differences in baseline risk leads to groups that will achieve larger or smaller absolute benefits from therapy, even though the relative effect of therapy is the same across risk groups.

In cases where there are knowable effect modifiers or knowable substantial differences in baseline risk (whether continuous or discrete), ICER highlights these differences in its discussion of the evidence. Depending on the nature of the evidence, the treatments, and the structure of the

report, subgroups may be discussed in greatest detail in individual outcomes sections of an ICER report or in a subsection called “Heterogeneity and Subgroups.” Subgroup differences may, on occasion, carry through to different evidence ratings for different subgroups.

In all reports, the subsection “Heterogeneity and Subgroups” will be included, either to present the primary discussion of subgroup effects, or to highlight the other sections of the evidence review that discuss subgroup effects.

As described in the sections below, when there are substantial knowable subgroup effects, economic analyses in ICER reports will also include these subgroup results.

2.5. Judgment of Level of Certainty and Magnitude of Net Health Benefit: the ICER Evidence Rating Matrix™

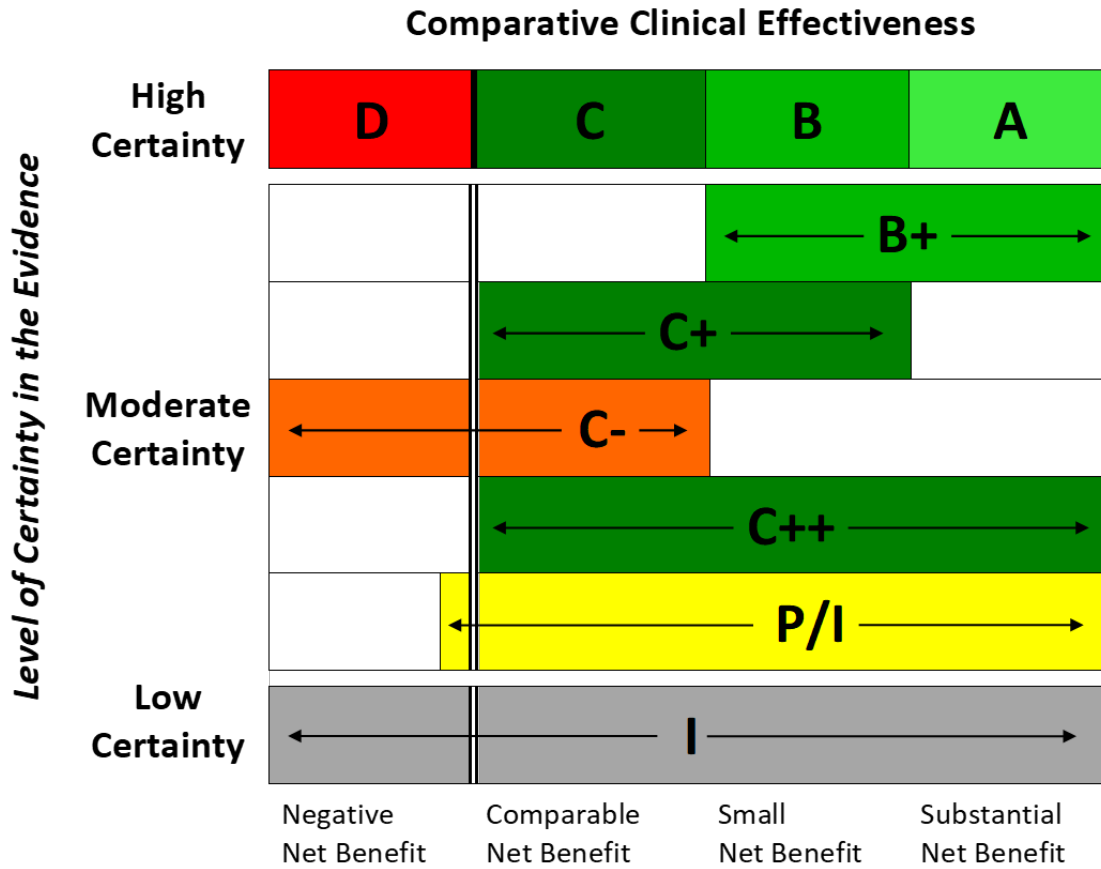
Following synthesis of the evidence by quantitative and qualitative techniques, ICER assigns overall evidence ratings to each of the interventions evaluated in its appraisal. A single intervention may be given more than one evidence rating if there are multiple comparators or if, as discussed above, there are substantial differences in the evidence ratings for a particular comparison across different populations or subgroups. Ratings reflect a judgment made at a moment in time and may be updated as new or additional evidence becomes available.

ICER developed the [ICER Evidence Rating Matrix™](#) (see Figure 2.1) to evaluate the overall strength of evidence for a variety of outcomes. The evidence rating reflects a joint judgment of two critical components:

- a) The **magnitude** of the difference between a therapeutic agent and its comparator in “net health benefit” – the balance between benefits and risks and/or adverse effects AND
- b) The level of **certainty** in the best point estimate of net health benefit.^{1,9}

The design of ICER’S Evidence Rating Matrix was informed by the approaches developed by the United States Preventive Services Task Force (USPSTF);¹⁰ the international Grading of Recommendations Assessment, Development and Evaluation (GRADE) group;¹¹ and the Effective Healthcare Program of the Agency for Healthcare Research and Quality (AHRQ).^{1,9,12} While each organization has developed unique criteria to rate the strength of evidence, each approach evaluates the entire body of evidence along a series of domains. The most important domains common to the four approaches include risk of bias, generalizability to “real-world” populations, consistency of findings across studies, directness (i.e., how closely the evidence measures the populations, interventions, and outcomes of interest), and precision.

Figure 2.2. ICER Evidence Rating Matrix



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

3. Incremental Cost Effectiveness

To ensure consistency in analytic approaches across all of its reviews, ICER has defined a detailed [Reference Case](#) specifying the approach that ICER and its collaborators follow for cost-effectiveness analyses. The reference case details all the methods that ICER and its modeling collaborators follow when conducting the base-case cost-effectiveness analysis. These methods generally follow the recommendations of the Second Panel on Cost-Effectiveness in Health and Medicine for the health care reference case,¹³ and is also generally consistent with published guidance from international HTA organizations.^{14,15} Following the reference case enables consistency in analytical approaches, but in specific cases, reasons may exist for deviating from the reference case. In such cases, the rationale for not fully applying the reference case methods will be clearly specified in the model analysis plan and Evidence Report.

Note that the description below provides general guidance on ICER's value assessment framework for health technology assessments (HTAs) in general. ICER's modifications to its methods for reviews of certain treatments for serious, ultra-rare disorders can be found [here](#), and modifications for reviews of high-impact single or short-term therapies (SSTs) can be found [here](#).

3.1. Overview

Cost-effectiveness analysis (also known as “economic modeling” or “decision analysis”) helps to assess whether a technology is a good value for money in the long run by considering cost in relation to the clinical benefits provided and comparing one treatment and its associated care pathway to another. These comparisons are done through a simulated computer model of patient and cost outcomes of different care pathways.

The objective of the economic evaluation is to determine the incremental cost-effectiveness ratio, or the cost per unit of health benefit gained of one treatment over another. The unit of health gained can be a specific clinical outcome, such as an additional stroke prevented or a case of cancer diagnosed, or a more generalizable unit such as an additional year of life or an additional year of life adjusted for any changes in quality of life.

As a summary measure of cost-effectiveness, ICER follows common academic and health technology assessment standards by using the cost per quality-adjusted life year (QALY) gained as the primary measure of cost-effectiveness,¹³ but also presents cost per life year gained and cost per [equal value life year gained \(evLYG\)](#).¹⁶ Lower incremental cost-effectiveness ratios represent better value for money. When the price of an intervention is known, the incremental cost-effectiveness ratio can be calculated. When the price is unknown (e.g. for an emerging treatment that has not

yet received FDA approval), ICER will often use an estimated price gained from analyst or other sources. ICER will also calculate the prices at which an intervention would hit certain cost-effectiveness threshold targets. For example, ICER calculates the prices needed to achieve \$100,000 per additional QALY and \$150,000 per additional QALY and uses these prices as the bookends of our “health-benefit price benchmark.” ICER also calculates these same price points using the evLYG as the measure of health gain in order to provide a complementary view of cost-effectiveness for stakeholders.

All cost-effectiveness models must make some assumptions about how evidence on the short-term effects of care plays out in clinical and economic effects that happen many years in the future. ICER evaluates this uncertainty by varying the inputs to the model, first one at a time, and then systematically across all model inputs, to assess how robust the results are with different inputs.

3.2. Model Structure and Data Sources

ICER is committed to open and transparent engagement with stakeholders in the development of our economic models. To fulfill this commitment and explain the model approach in detail, ICER develops a model analysis plan following the publication of a final scoping document. The model analysis plan outlines the methods the economic modeling team intends to employ, including information on the model structure and processes, all major inputs and sources of data, and key assumptions. In addition, the plan specifies whether the model is an adaptation of an existing model (with references as appropriate) or is being developed *de novo* for that HTA. The model analysis plan is published on the Open Science Framework (<https://osf.io/7awvd/>). The plan may be updated following review of additional data sources, discussions with stakeholders, and other activities.

In the model analysis plan and evidence report, the specific decision to be addressed by the analysis is specified in terms of the overall objective, the interventions and comparators, the relevant population groups and subgroups being considered, and the outcomes. Any differences in the population, intervention, or outcomes from the aims and structure of the clinical evidence review are documented with justifications. The analytic perspective (typically health care system) and time horizon (typically lifetime) used in primary analyses are also specified.

Following discussions with stakeholders and review of any additional data sources, the model analysis plan may be updated. The final version of the model used in conducting analyses is outlined in the Evidence Report, which is intended to provide enough information for an experienced researcher to be able to replicate the economic model and analyses.

Model Parameters and Data Sources

Model inputs, or “parameters,” include those pertaining to intervention effectiveness, transition rates between health states, measurement and valuation of health states, resource use, and costs. Results from the evidence review, including the results from any meta-analysis, are used to inform input parameters when possible. All model parameters are described in the model analysis plan and evidence report, including risk equations as appropriate. ICER aims to use data from published or publicly available sources, including peer-reviewed journals, supplementary appendices, briefing documents used by regulatory authorities, and conference proceedings. In specific instances, valid analyses may require the use of unpublished information, such as manufacturers’ data on file.

Acceptance of Multiple Forms of Evidence

For comparative cost effectiveness, ICER’s value framework does not limit the type of evidence to be considered to the results of randomized controlled trials (RCTs). When available, high-quality RCTs typically provide evidence on short to mid-term clinical benefits and more commonly occurring harms. When head-to-head trials have not been performed, indirect comparisons through formal network meta-analysis may be used as inputs for economic modeling. When benefits and harms occur over the course of many years, or when harms are rare but clinically important or even catastrophic, evidence from high quality published peer-reviewed studies using observational data and methodologies such as cohort studies, case-control studies, and long-term disease and drug registries may be used. Furthermore, if important patient reported outcomes have not been collected as part of a manufacturer’s clinical development program, ICER will again conduct a comprehensive literature review to identify published, peer-reviewed observational studies providing this information.

Real-World Evidence

Because inputs to economic models are often not included as outcomes in RCTS, the use and integration of evidence, based on observational or real-world data, has been an important source of model inputs and incorporated when appropriate in ICER cost-effectiveness analyses. RWE can be especially useful as a source of model inputs on transitional health states, adherence and persistence, costs, and health utilities among others. The use of real-world data includes *de novo* evidence generation under certain circumstances where critical data elements for comparative cost effectiveness are lacking. This may include analyses of insurance claims data to better understand health states, resource utilization, and costs, or the analysis of new data from patient surveys to provide more direct information on health utilities.

Clinical Expert and Patient Input

For some economic models there will remain gaps in the available evidence despite review of published data and attempts to analyze or generate RWE. In these cases, ICER uses input from clinical experts and/or patient groups to supply best estimates for the elements of a clinical care pathway, the likelihood of specific patient outcomes, and other inputs required to compare two or more treatments.

Data in Confidence

Because life science companies may have relevant information that is currently held in confidence, ICER has structured a process to accept and use such data. We allow manufacturers to submit data that is not yet in the public domain if the use of the information will be of help to the economic evaluation. (ICER has specific protections in place for this confidential data, which are outlined at: <https://icer-review.org/use-of-in-confidence-data/>.)

3.3. Measures of Health Gain

The sources and methods used for health preferences measurement are provided in the model analysis plan. These methods usually involve mapping health states in patients with a condition into a classification system with associated utility weights, such as the EQ-5D.^{17,18} Generic classification systems such as the EQ-5D include measures of health state preferences that reflect those of the general US population, considered to be relevant to inform decisions at the population level (e.g., payer or health system formulary decisions) that involve individuals both with and without the condition of focus. Where general population estimates are not available or appropriate, utility estimates from different populations may be used, such as patients with the specific condition under study, those affected by similar symptoms, proxy respondents, or mixed samples. When there are challenges in translating outcome measures used in clinical trials or available patient-reported data into health states, the report discusses the rationale for choosing specific mapping algorithms.

Health effects are expressed in terms of total and incremental quality-adjusted life-years (QALYs), equal value life years gained (evLYG), life-years, and a condition-specific outcome achieved (e.g., treatment response, event avoided). ICER uses the QALY as part of assessments that compare therapies on their ability to improve quality of life and lengthen life. The QALY is the gold standard for measuring how well a medical treatment improves and lengthens patients' lives, and therefore has served as a fundamental component of cost-effectiveness analyses in the US and around the world for more than 30 years. Economic analyses using the QALY make treatments that alleviate serious illness look especially valuable. Because the QALY records the degree to which a treatment improves patients' lives, treatments for people with serious disability or illness have the greatest

opportunity to demonstrate more QALYs gained and justify a higher price. In addition, a common measure of improved outcomes for patients is needed for cost-effectiveness analyses to support broader efforts to make more transparent, evidence-based coverage policies and pricing decisions.

To provide additional context to the cost per QALY estimates, ICER reports include analyses of cost per evLYG, cost per life-year gained, and cost per some condition-specific consequence as a core part of every report, seeking input from patients, clinical experts, payers, and manufacturers on what outcome(s) will be most important for this comparison. The evLYG analysis counts any gains in length of life equally, regardless of the treatment's ability to improve patients' quality of life. For all additional years of life gained, this analysis awards full health (i.e., the quality of life of the general population), irrespective of the health state patients are in during these additional years of life gained. In other words, if a treatment adds a year of life to a population with a severe condition or disability, that treatment receives the same evLYG as a different treatment that adds a year of life for healthier members of the community.

ICER uses the evLYG in economic analyses whenever relevant and feasible given model structure, as a supplement and not a replacement to the cost per QALY analysis, which reflects the true benefits a treatment may have on the quality of life on the population in which it is used. In certain situations, model structure may make the calculation of evLYG intractable, in which case we will report life years gained rather than evLYG.

Using both the cost per QALY and the cost per evLYG results will enable policy makers to gain a broad overview of the cost-effectiveness of treatments while ensuring that results are available to demonstrate whether there is any impact of extended life at a low quality of life. If ICER's analysis finds a major difference in these two measures, reports include specific language describing the underlying characteristics of the treatment and the condition that lead to the difference. More information on the evLYG analysis is available [here](#). ICER participates in the global dialogue around the best methods for evaluating the value of health services and is always attuned to new developments that might provide a better and fairer system of measuring benefits across different kinds of interventions and patients.

3.4. Impact on Distribution of Health Gains

Health inequality is an important concern for patients and policy makers in health systems across the globe. ICER has explored options for measuring the degree to which treatments may result in greater or lesser inequality across racial or socio-economic groups in the US. Data to support application of available methods are lacking in the US, and none of these methods have been adopted as standards within other HTA agencies. Nonetheless, where judged feasible, ICER may

explore through scenario analyses methods to capture the impact of new technologies on disparities in life expectancy across different subpopulations in the US health care system.

3.5. Perspective

ICER calculates incremental cost-effectiveness from the health care system perspective as its base case, but also performs a modified societal perspective analysis in a scenario including work productivity and other effects that may occur outside the health system. ICER uses the health care system perspective as its primary base case for several reasons. First, ICER's reports are primarily intended to inform population-based medical policy and pricing decisions within the US health care system. Employers, other plan sponsors, insurers, and risk-bearing provider groups in both private and public health insurance systems are not responsible for making trade-off decisions that involve broader societal resources, so the health care system perspective is the most directly relevant for decision-making. This is not to imply that plan sponsors, insurers, and others do not care about effects of health care interventions outside the health system. But their primary responsibility and the framework for the trade-offs they must manage rest within the health system.

Another reason that the health system perspective is favored is that full consideration of the societal perspective often requires inclusion of broad and uncertain assumptions regarding the impact of health care not only productivity, but on income tax generation, educational outcomes, the criminal justice system, and disability and social security benefits. Seeking to capture the full scope of these effects is practically almost impossible, and also raises the potential for unintended consequences, such as potentially favoring a selection of health care interventions that minimize the amount of time individuals spend receiving public financial support. A societal perspective raises several important ethical concerns of this nature, including whether interventions that support the health and productivity of younger – and healthier -- individuals should be favored over interventions for those whose contributions to society cannot be equally measured through salaries, taxes paid, or independence from public services. ICER is sensitive to provide a framework for analyses that does not conflict with important ethical goals of US society.

The Second Panel on Cost-Effectiveness in Health and Medicine recommends reporting results from both the health care system perspective and the societal perspective, with an “impact inventory” used to make transparent which elements of a full societal perspective are included.¹³ ICER follows this approach. To emphasize the important distinctions between health care system and societal perspectives, ICER includes in an Appendix the template from the Second Panel on Cost-Effectiveness to describe the elements of health care system and societal perspectives that are included in each cost-effectiveness analysis. To the extent feasible, the relative impacts of different care options on work productivity and other indirect impacts are estimated in the ICER report and

are considered by ICER independent public appraisal committees as part of their weighing of “other potential benefits and disadvantages,” as described later in this paper.

Modified Societal Perspective as Co-Base Case

To try to strike a balance between the ethical and other risks of a societal analysis and the potential interest of decision-makers in the results of analyses done with modified societal perspective, ICER presents a modified societal perspective as a co-base case for certain topics. When ICER judges that the societal costs of care for any disease are large relative to the direct health care costs, and that the impact of treatment on these costs is substantial (i.e., there are substantial differences in the cost-effectiveness findings between the two perspectives), the societal perspective is included as a co-base case, presented directly alongside the health care sector perspective analysis. This will most often occur in cases where the incremental cost-effectiveness ratio changes by greater than 20%, greater than \$200,000 per QALY, and/or when the result crosses thresholds of \$100,000-\$150,000 per QALY.

3.6. Discounting

To account for time value and ensure comparability across studies, all economic models use constant-rate discounting of both costs and outcomes, at the rate of 3% per year. Discounting is a standard method in economic modeling, and in the US, the standard approach has been confirmed by the Second Panel on Cost-Effectiveness in Health and Medicine as a uniform discount rate of 3% applied to both costs and benefits.¹⁹ The use of a 3% discount rate in the US as standard for both costs and outcomes is based on estimates of the real consumption rate of interest and data on real economic growth, which are thought to reflect the social rate of time preference. The use of a single, uniform discount rate for all assessments allows for consistent comparisons across different or prior evaluations.

3.7. Patient Populations

To the extent possible, the patient populations included in ICER’s economic evaluations are generally those for which the evaluated interventions are indicated. However, at the time of evaluation, the only available data on efficacy may come from trials that do not reflect the (likely) indicated population. In such cases, the discrepancy between the indicated population and the trial populations will be pointed out, along with discussion of the relevance of trial results to the larger population. While cohort models tend to reflect homogeneity in patient populations for whom health technologies are assessed, when relevant, ICER’s evaluations include scenarios with different

patient subgroups to account for the heterogeneity within patient groups within a specific disease area.

Heterogeneity and Subgroups

Evidence Reports include a sub-section on “Heterogeneity and Subgroups” in order to broaden discussion of heterogeneity and subgroups within the patient population. ICER’s Reference Case calls for the inclusion of different patient subgroups when analyzing the cost-effectiveness of health technologies, to the extent possible. Data permitting, subgroup analyses will be considered for patient groups that could be of interest either clinically or economically. ICER’s economic evaluations include analysis of patient subgroups when robust data and relevant inputs from clinical trials and/or real-world evidence are available to do so. Such subgroup analyses have been and will continue to be undertaken when ICER believes that health technologies are likely to be approved or have been used extensively within these subgroups of interest, and as mentioned earlier, pending data availability.

3.8. Costs

Costs are reported in terms of total and incremental costs. When possible, ICER uses estimates of prices net of discounts, rebates, and other price concessions as the base-case input for prices used in cost-effectiveness and potential budget impact analyses. Analyses using wholesale acquisition cost (WAC) prices are also included for context. To provide pricing that can reliably and with relative transparency provide an estimate for net prices in the US market, ICER collaborates with SSR Health LLC, a consultancy which combines data on net US dollar sales with information on unit sales to derive net pricing at the unit level across all payer types. Further details on the mechanism used to estimate net prices are available in ICER’s Reference Case.

ICER’s cost-effectiveness analyses will not routinely make estimates of price changes across comparator treatments linked to patent and exclusivity time horizons. However, when high likelihood of a major change to pricing can be anticipated within 12-24 months, a scenario analysis may be developed to explore the impact of price changes on long-term cost-effectiveness ratios, if there are consistent historical findings of price changes that can be applied to the topic under review.

3.9. Base-Case and Cost-Effectiveness Thresholds

In the presentation of the results of incremental cost-effectiveness analyses, health benefits and costs are summarized as incremental cost per QALY gained, cost per evLYG, cost per life-year gained, and cost per condition-specific measure of clinical benefit. ICER will provide cost-per-QALY results at \$50,000, \$100,000, \$150,000 and \$200,000 per QALY and per evLYG for all assessments, including those for treatments of ultra-rare disorders. The range for health-benefit price benchmarks remains \$100,000-\$150,000 per QALY and evLYG, reflecting ICER’s judgment that the most recent research confirms that both “opportunity cost” and “willingness to pay” paradigms produce estimates of an operational cost-effectiveness threshold at approximately \$100,000 per QALY. Because ICER’s suggested health-benefit price benchmarks are most often used as ceiling prices, we continue to use \$150,000 as a more flexible and liberal upper bound. For more information regarding ICER’s rationale for its cost-effectiveness threshold range, please see [Appendix D](#).

ICER’s Evidence Reports present a broader range of results symmetrically around this range, from \$50,000-\$200,000 per QALY/evLYG. This range is meant to accommodate the needs of decision-makers in the US to think about their own desired interpretation of cost-effectiveness thresholds while considering uncertainty, other benefits or disadvantages, and contextual considerations.

3.10. Sensitivity and Scenario Analyses

As a method to evaluate uncertainty in the economic evaluation, the Evidence Report also includes one-way sensitivity analyses, presenting the results in “tornado diagrams” that display the findings across a feasible range for each input parameter estimate and a table containing the ranges and distributional assumptions around the input parameters varied. Expected values of costs and outcomes for each intervention are also estimated through probabilistic sensitivity analyses, which characterizes some of the uncertainty in the input parameter estimates. This type of analysis takes repeated samples, typically 1,000 or more, from the (joint) distribution of all key model input parameters simultaneously; results are presented tabularly in terms of the percentage of simulations that achieve \$50,000, \$100,000, \$150,000, and \$200,000 per QALY thresholds, and graphically using scatter plots or cost-effectiveness acceptability curves (CEAC) which reflects the percentage of simulations that result in incremental cost-effectiveness ratios that fall at or under various cost-effectiveness thresholds.

Scenario Analyses

Specific scenario analyses (including one using a modified societal perspective that incorporates estimates such as productivity losses, caregiver burden, and other indirect costs) and subgroup

analyses are conducted when appropriate. In addition, the report presents results from threshold analyses which estimate the intervention prices that correspond to cost-effectiveness thresholds extending from \$50,000 per QALY gained to \$200,000 per QALY gained.

Hypothetical Shared Savings Scenarios

Two hypothetical scenarios are generated for all high-impact SSTs under review, as well as other, non-SST treatments with relevant and substantial potential cost-offsets. In most cases this will be situations in which potential cost offsets are greater than \$1 million over a lifetime. To stimulate further consideration of how the cost offsets generated by new treatments should be incorporated in calculations of the health benefit-related value for a new treatment, Evidence Reports include two hypothetical economic analysis scenarios that evaluate cost-effectiveness outcomes with different approaches to the cost offsets from a new treatment. Threshold analyses for treatment price will be presented but will not be suggested as normative guides to pricing. Further description of the rationale for these scenarios can be found in our [SST methods document](#).

The two hypothetical scenarios to be generated for all high-impact SSTs under review, as well as other, non-SST treatments with relevant and substantial potential cost-offsets are:

1. A 50/50 shared savings model in which 50% of the lifetime health system cost offsets from a new treatment are “assigned” to the health system instead of being assigned entirely to the new treatment; and
2. A cost-offset cap model in which the health system cost offsets generated by a new treatment are capped at \$150,000 per year but are otherwise assigned entirely to the new treatment.

Outcome-Based Payment Arrangements.

When relevant, Evidence Reports include information from manufacturers and payers to model a scenario analysis including a limited number of outcome-based payment arrangements for the intervention under review. In some cases, these payment arrangements can be a useful tool in managing uncertainty and increasing the ultimate cost-effectiveness of treatment. ICER actively seeks information from manufacturers and payers about the potential outline of outcomes-based contracts for scenario analyses in our reports. In cases where the list price of the treatment is known but there is no guidance from stakeholders, an exploratory scenario analysis using outcomes and levels of financial risk-sharing that could meet cost-effectiveness thresholds may be performed.

Exclusion of Unrelated Costs

In cases where an intervention that increases QALYs would not be found to be cost effective, even with a zero-dollar price, a separate scenario analysis excluding unrelated (non-drug) health care costs will be presented. We have encountered specific situations in assessments where the cost-

effectiveness analysis is not able to produce a non-negative threshold price that would make a given treatment cost-effective. In such cases, there are no positive prices for an intervention that will reach specific cost-effectiveness thresholds. This may occur in situations where a new treatment is added on to existing treatment that is already near or beyond the cost-effectiveness threshold. Another example where this may occur is when a new treatment results in more time spent in health states that have very high costs and/or a low utility value, making it impossible for the incremental cost-effectiveness ratio to reach specific thresholds even at zero price.²⁰ In such cases a scenario analysis excluding health state costs that are not related to the intervention *per se*, may be informative.

3.11. Validation and Calibration

All economic models are validated prior to conducting analyses, as well as during the production of the Evidence Report. Validation entails assessing whether a model has been implemented correctly (internal validation) and if its assumptions and results are in line with the current evidence and expectations (face validity, external validation). All models are internally validated by an independent modeler. The specific approach to internally validate the model during development is detailed in the Model Analysis Plan and follows ICER’s validation checklist (publication forthcoming). After the posting of the Model analysis Plan and a presentation of model structure, assumptions, and inputs, key stakeholders also provide feedback on the model assumptions, parameters, structure, and overall face validity. In addition, ICER releases economic model files and code to manufacturer stakeholders willing to agree to confidentiality and privacy restrictions, allowing participating stakeholders to include detailed critique of the model in public comments submitted on the Draft Evidence Report.

Calibration entails assessing if the model inputs and outputs are consistent with known scenarios. Any calibration procedures used during model development are proposed in the Model Analysis Plan, including the calibration target (and source), the goodness-of-fit metric, and criteria for judging fit. Results from the calibration procedure are presented in the Evidence Report.

3.12. Uncertainty and Controversies

Evidence Reports include a sub-section on “Uncertainty and Controversies” in order to broaden discussion of alternative model structures and assumptions suggested by manufacturers or other stakeholders. One important goal of this section is to provide further elaboration of the rationale behind methodological decisions that underpin the base case. This sub-section also serves as an avenue to discuss how different assumptions or scenarios might affect model results and as a useful tool for decision-makers to understand the issues and uncertainties that may remain controversial.

To accomplish this goal the sub-section provides discussion of different model variations that could be viewed as more conservative or optimistic. In particular, this sub-section addresses alternative model structures or inputs suggested by manufacturers or other stakeholders that differ importantly from the base case. This sub-section also consolidates and expands discussion of factors related to uncertainty, including lack of information on natural history, limitations of the data on patient outcomes, difficulties translating existing data into measures of quality of life, and disagreements over the plausibility of certain inputs or assumptions.

Summaries of relevant published cost-effectiveness analyses are also included in this sub-section, pointing out differences in model structure, inputs and assumptions, and the impact of these differences on model results. We review and compare the current model to published models that included the same interventions or comparators of interest, were developed in the last 10 years, and were similar to the current model from a setting and population perspective.

3.13. Health-Benefit Price Benchmarks

For all assessments, an ICER “health-benefit price benchmark” (formerly called a “value-based price benchmark”) is developed for the new intervention, which reflects prices aligned with commonly-cited long-term cost-effectiveness thresholds ranging from \$100,000 to \$150,000 per QALY gained and from \$100,000 to \$150,000 per evLYG. The prices represent discounts or price premiums from wholesale acquisition cost (WAC) that would be required to reach these cost-effectiveness thresholds. For more information regarding ICER’s rationale for the cost-effectiveness threshold range used for the health-benefit price benchmarks, please see [Appendix D](#).

ICER’s health-benefit price benchmarks suggest a price range, net of any discounts and rebates, that aligns fairly with a treatment’s added benefits for patients over their lifetime. Prices at or below these thresholds help ensure that the health benefits gained by patients using new treatments are not outweighed by health losses due to long-term cost pressures that lead individuals to delay care, abandon care, or lose health insurance.

As noted earlier in Section 3.3, ICER believes that there is a confluence of results between research exploring opportunity cost thresholds and willingness to pay thresholds in the US setting. For conceptual reasons ICER favors a view of thresholds based in an opportunity cost paradigm. Claxton has presented data analyses supporting the adoption of cost-effectiveness thresholds in the UK, US, and other countries that are far lower than traditional thresholds, given the marginal productivity of the respective health care systems.²¹⁻²³ For the US, Claxton estimates an opportunity cost threshold of approximately \$30,000-\$40,000 per QALY.²³ More recently, Vanness has estimated health opportunity costs for private plans in the US,²⁴ and produced an estimate of \$84,000 per QALY as the threshold. Working within this paradigm, this means that any new

intervention introduced at a price that leads to an incremental cost-effectiveness ratio greater than \$84,000 per QALY produces a net loss of health due to its impact on premium increases and thereby loss of insurance and the attendant negative health effects, especially among poorer members of the insurance pool.

In the US market-based system with multiple payers, there is a case for multiple thresholds based on willingness-to-pay which may differ by payer type (e.g., government vs. commercial insurance).²⁵ However, there are broad requirements across the US health care system to fund all “medically necessary” care. There is also a widely accepted ethical goal in the US to have a common standard of care available for all patients, albeit with acknowledged differences in access due to network constraints, out-of-pocket payment, and other benefit design features. That the US does not yet achieve the goal of a common standard of care available for all patients does not imply, in our view, that ICER should abstain from framing a range of cost effectiveness that should apply broadly across many, if not all, health insurance systems in the US.

Despite the lack of an explicit overall budget for health care in the US, the current environment of the US health care system is one in which policy-makers sense that the opportunity cost for current spending is already substantial, and that real harm is being done as health care costs continue to rise. We believe that anecdotal evidence and testimony from these policymakers further supports ICER’s decision to apply an opportunity cost approach to a threshold range, the goal being to ensure that the prices paid for health gains from effective new treatments are aligned with the magnitude of those health gains such that greater health is not lost through the effects of rising health costs at the system and societal level.

Reflecting on the most recent conceptual and empirical research, a case could be made for reducing our health-benefit price benchmark range to \$50,000-\$100,000 per QALY. However, the top end of the price benchmark range is usually interpreted as a “ceiling” price beyond which a treatment will be viewed as not cost-effective. There is also value in retaining a consistent threshold range as a level playing field for all stakeholders. Therefore, ICER continues to use the cost-effectiveness range of \$100,000 to \$150,000 to support health-benefit price benchmark recommendations. ICER recognizes that single cost-effectiveness thresholds should not be used as a blunt decision rule, and that decision-makers may want to consider different thresholds given their own view of their opportunity costs and their interpretation of a treatment’s potential other benefits and contextual considerations.

ICER will continue to evaluate which cost-effectiveness thresholds should be used to generate health-benefit price benchmarks to reflect ongoing academic work that may support a different threshold range and may update these thresholds prior to 2023.

4. Potential Other Benefits or Disadvantages and Contextual Considerations

4.1 Overview

The inclusion of explicit domains of value labeled “potential other benefits or disadvantages” and “contextual considerations” demonstrates that the ICER Value Assessment Framework fully acknowledges that all too often what matters most to patients is poorly captured in the available clinical trial data. Sometimes this occurs because surrogate outcome measures do not reflect true patient-centered outcomes; but even when trials do capture the clinical outcomes that matter most to patients, there are other aspects of value related to the complexity of the treatment regimen or the impact of care options on the ability to return to work, on the negative impact of the condition on family and caregivers, on public health, or on other aspects of the health system or society. The ICER value framework identifies these “potential other benefits or disadvantages” as important elements of any overall judgment on long-term value for money, and all ICER reports have separate sections in which evidence and information pertaining to these elements are presented.

Similarly, decisions about the value of care options do not happen in a vacuum. There may be broader contextual issues related to the severity of the condition, whether other treatments are available or soon will be, and ethical, legal or other societal priorities that are important to acknowledge as part of any discussion on value. The ICER value framework includes this element and it is explored in a separate section of each ICER report.

Many researchers and policymakers have explored different ways to elicit potential other benefits and contextual considerations and apply them to weight QALYs or adjust cost-effectiveness thresholds. However, all proposals involve potential risks, such as the risk that considerations of productivity gains will adversely affect the relative value of treatments for the elderly and disabled. Therefore, there are no widely accepted protocols for how best to weight factors outside traditional cost-effectiveness analysis, and most health technology assessment groups around the world do not attempt to quantify these domains of value, believing that their relative weight in any overall judgment of value should be left qualitative and subject to public discussion.

ICER has considered several methodological options that could enhance the transparency and explicit integration of these considerations. Formal multi-criteria decision analysis (MCDA) has been considered but rejected because we do not believe that the methods for weighting individual elements are robust enough to add to reliability of value judgments. ICER has attempted formal MCDA with its independent committees on several occasions in the past and found the technique too complicated for reliable use. Based on discussions with stakeholders, benchmarking other value frameworks around the world, and the input of public comment ICER reports will continue to

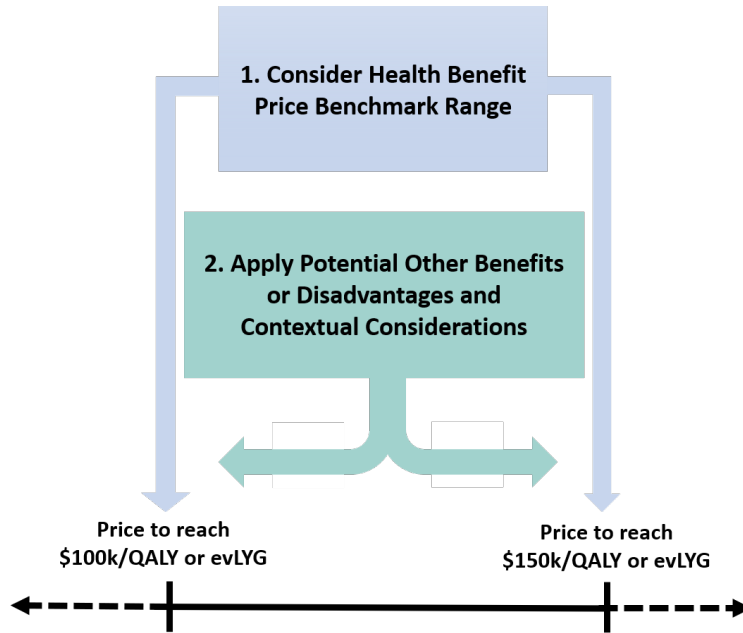
use a variation on MCDA that makes other benefits or disadvantages and contextual considerations explicit and gives clear guidance on their relevance to judgments of value, but that does not attempt an overly facile quantification. Decision-makers will be given guidance, however, that consideration of these factors should guide part of their thinking about how to use the cost-effectiveness threshold range, with higher ends of the range more applicable when there are important positive contributions of other benefits and contextual considerations, and lower ends of the range reflecting relatively less consequential added value considerations.

In ICER public deliberation meetings, independent appraisal committees will take votes on each specific category of potential other benefit or disadvantage and contextual consideration. After several iterations of voting approaches over the past four years, ICER now uses a three-item Likert scale voting format. The goal of this voting method is to provide the appraisal committees with a clearer understanding of the ends of the spectrum within which they are expected to vote. It is also intended to produce a more transparent record of how the appraisal committee feels that these considerations should be applied when integrated with the cost-effectiveness results in making decisions about pricing. Individual votes will be noted in the Final Report, but our methods do not include calculation of an overall average vote across all categories. An average will not reflect the great likelihood that certain individual other benefits/disadvantages or contextual considerations are and should be of far greater consequence in judging the overall value of an individual intervention.

While the specific implications of each numerical vote may vary due to the language used in the consideration, they are broadly meant to indicate the guidance for decision-makers making judgments of the overall long-term value for money of interventions. A schematic view of this approach is shown in Figure 4.1 below, and general interpretation of the votes on potential other benefits and contextual considerations is as follows:

- **1:** Within the general range suggested by clinical evidence and findings of cost-effectiveness analysis, consideration of potential other benefits or disadvantages and/or contextual considerations points toward relatively lower longer-term value for money
- **2:** Within the general range suggested by clinical evidence and findings of cost-effectiveness analysis, consideration of potential other benefits or disadvantages and/or contextual considerations points toward relatively intermediate longer-term value for money
- **3:** Within the general range suggested by clinical evidence and findings of cost-effectiveness analysis, consideration of potential other benefits or disadvantages and/or contextual considerations points toward relatively higher longer-term value for money

Figure 4.1. Conceptual Guide to Application of “Potential Other Benefits or Disadvantages” and “Contextual Considerations” to Judgements of Value



evLYG: equal value life-years gained, QALY: quality-adjusted life years

4.2. Specific Categories

The following specific categories of potential other benefits and contextual considerations will be noted in ICER reports and subject to deliberation and voting by the independent appraisal committee at each public meeting (Table 4.1). Minor adaptations to the below list can be found in the framework adaptations for [single- and short-term therapies](#) and [treatments for ultra-rare diseases](#).

Table 4.1. Potential Other Benefits or Disadvantages and Contextual Considerations

1 (Suggests Lower Value)	2 (Intermediate)	3 (Suggests Higher Value)
Uncertainty or overly favorable model assumptions creates significant risk that base-case cost-effectiveness estimates are too optimistic		Uncertainty or overly unfavorable model assumptions creates significant risk that base-case cost-effectiveness estimates are too pessimistic
Very similar mechanism of action to that of other active treatments		New mechanism of action compared to that of other active treatments
Delivery mechanism or relative complexity of regimen likely to lead to much lower real-world adherence and worse outcomes relative to an active comparator than estimated from clinical trials		Delivery mechanism or relative simplicity of regimen likely to result in much higher real-world adherence and better outcomes relative to an active comparator than estimated from clinical trials
The intervention offers no special advantages to patients by virtue of presenting an option with a notably different balance or timing of risks and benefits		The intervention offers special advantages to patients by virtue of presenting an option with a notably different balance or timing of risks and benefits
This intervention will not differentially benefit a historically disadvantaged or underserved community		This intervention will differentially benefit a historically disadvantaged or underserved community
Small health loss without this treatment as measured by absolute QALY shortfall		Substantial health loss without this treatment as measured by absolute QALY shortfall
Small health loss without this treatment as measured by proportional QALY shortfall		Substantial health loss without this treatment as measured by proportional QALY shortfall
Will not significantly reduce the negative impact of the condition on family and caregivers vs. the comparator		Will significantly reduce the negative impact of the condition on family and caregivers vs. the comparator
Will not have a significant impact on improving return to work and/or overall productivity vs. the comparator		Will have a significant impact on improving return to work and/or overall productivity vs. the comparator
Other		Other

Model Uncertainty and Assumptions

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

Mechanism of Action

Although the benefits of a new mechanism of action should be captured within the clinical outcomes demonstrated in clinical trials and other evidence, some decision-makers may wish to give some prioritization or greater value weighting to interventions that represent a new mechanism of action either because they believe it will have positive spillover effects in other clinical areas and/or because it may allow for the treatment of many patients for whom treatments using other available mechanisms of action have not been effective.

Complexity of Regimen and Real-World Adherence and Outcomes

Clinical trials are designed and administered such that patient and clinician adherence to treatment protocols are maximized, usually significantly above levels seen when treatments are used in widespread practice. Comparing the evidence from clinical trials of two active comparators may therefore overlook the potential for some methods of administration to perform relatively better in real-world practice than others. One common example is the comparison of clinical outcomes between infused treatments and oral options. The relative simplicity of taking a pill once a day is often found to lead to superior patient adherence and outcomes compared to more complex methods of administration.

Balance or Timing of Risks and Benefits

We believe that the concept of “value of hope” is poorly named to convey the advantages that some treatments may offer if they have a distinctly different timing or balance of risks and benefits compared to other available treatments. The classic example is a treatment for cancer that may have, overall, the same total QALYs gained as existing options, but achieves that equivalent overall gain through a distinctively higher risk of short-term death matched by a higher chance of longer-term survival. For risk-taking patients this treatment option, although its QALYs might be identical to other options, offers a special advantage, and this category of potential other benefit seeks to capture situations in which a treatment offers a distinctive balance or timing of risks and benefits that will be valued by patients and clinicians.

Disadvantaged or Underserved Communities

Many societies have explicit goals for health care that include the reduction of historical disparities in access or outcomes of care. Interventions which proportionately target patients in

disadvantaged communities may therefore merit special consideration. Relevant communities are those that have been historically disadvantaged through discrimination, neglect, reduced research funding, or other factors. We do not believe there are robust methods for quantifying this social value (i.e., we do not believe there are reliable methods to weight QALYs gained by patients from disadvantaged or underserved communities). However, through voting on this potential other benefit we hope to highlight for policymakers when an intervention may affect patient outcomes in a way that would help address historical inequities and lead to less disparity in important health outcomes.

Absolute QALY Shortfall

Another important social value is that which gives some preference to treatments for patients with more severe conditions.^{26,27} Countries and health technology assessment groups have conceptualized this idea somewhat differently. Some have seen that giving some priority to treatments according to “lifetime burden of illness” or “need” best represents the ethical instincts of a society or other decision-makers.^{28,29}

To inform this contextual consideration ICER will provide in its reports empirical results for the absolute QALY shortfall and proportional QALY shortfall. The absolute QALY shortfall is defined as the total absolute amount of future health patients with a condition are expected to lose without the treatment that is being assessed.³⁰ The ethical consequences of using absolute QALY shortfall to prioritize treatments is that conditions that cause early death or that have very serious lifelong effects on quality of life receive the greatest prioritization. Thus, certain kinds of treatments, such as treatments for rapidly fatal conditions of children, or for lifelong disabling conditions, score highest on the scale of absolute QALY shortfall. The Norwegian health technology assessment program is perhaps the most notable organization currently using measures of absolute QALY shortfall as a component in their appraisal process.³¹

Proportional QALY Shortfall

ICER reports will also include empirical calculations of the proportional QALY shortfall.³² The proportional QALY shortfall is measured by calculating the proportion of the total QALYs of remaining life expectancy that would be lost due to untreated illness.³³ The proportional QALY shortfall reflects the ethical instinct to prioritize treatments for patients whose illness would rob them of a large percentage of their expected remaining lifetime. As with absolute QALY shortfall, rapidly fatal conditions of childhood have high proportional QALY shortfalls, but the highest numbers can also often arise from severe conditions among the elderly who may have only a few years left of average life expectancy but would lose much of that to the illness without treatment.

In order to provide some anchoring to the deliberation at ICER public meetings, the results of absolute and proportional QALY calculations will be accompanied by league tables of absolute and

proportional QALY shortfalls for a variety of interventions from the academic literature.³⁴ We will also explore real-time use during meetings of a burden of disease calculator developed by Dutch investigators (see <https://imta.shinyapps.io/iDBC/>) that allows for calculation of absolute and proportional QALY shortfalls under different assumptions.

Caregiver and Family Impacts

When empirical data are available, ICER will include caregiver and family productivity outcomes in scenario analyses of its modified societal perspective analysis. When no data are available, assumptions may be made based on patient/family input and/or clinical expert input. In all cases, effects on caregiver and families will be included in the voting of the appraisal committees to signal whether they believe the impact of the intervention has important effects that should be considered by policymakers.

Caregiver and family utilities are difficult to incorporate into economic modeling because there is no established way to determine how to aggregate QALYs across multiple family members in a way that is consistent and can be applied to opportunity cost thresholds. In addition, incorporation of caregiver utilities into lifetime economic models can produce counter-intuitive findings. For example, a treatment that extends the life of a disabled child may extend caregiver “disutilities” and would appear to produce lower lifetime quality of life for caregivers than would allowing a child to die quickly. This finding could be technically “correct” but is in total opposition to caregivers’ deepest beliefs and hopes for extended time with a family member. We therefore have discussion and voting on impact on caregivers and family through a more qualitative approach.

Return to Work and/or Productivity

Data or best estimates on the impact of treatment on patients’ ability to return to work and their level of productivity will be included in every analysis done through our modified societal perspective. We will ask the appraisal committee to vote on the relative impact of treatment on this potential other benefit without providing them with specific guidance on what a “significant” or “substantial” impact is.

5. Potential Budget Impact Analysis

5.1. Overview

While it is important to understand how expensive a new technology is for a given unit of benefit, it is also important to look at the technology in terms of short-term financial impact to the overall health care system. ICER analyzes the short-term potential budget impact of changes in health expenditures with the introduction of a new test, treatment, or delivery system process. The potential budget impact is an estimate of the projected cumulative resource expenditure across all elements of the health care system for a specific intervention in a specific population over a period of time. ICER uses a five-year timeframe for its potential budget impact analysis to capture important potential clinical benefits and cost offsets provided by newer care options. Potential budget impact models aim to quantify the net cost over a short period of time for all eligible patients to receive the new technology.

For pharmaceuticals, the results of the budget impact analysis are compared to a national annual threshold for a new drug that is tied to growth in the overall US economy. This threshold, calculated by ICER, is updated each calendar year using the most recent inputs available to reflect changes in US gross domestic product, medical and pharmaceutical spending, and the average annual number of drugs approved by the FDA over the last five years. The current potential budget impact threshold calculations are detailed at <https://icer-review.org/methodology/icers-methods/icer-value-assessment-framework-2/>.

This comparison is intended to signal to stakeholders and policy-makers when a new treatment, even one priced at a level commensurate with good long-term value, may add short-term health care costs that are so substantial that they would be difficult for the health care system to absorb over the short term without displacing other needed services or contributing to rapid growth in insurance costs that could threaten sustainable access to high-value care for all patients. ICER seeks to include information for estimating short-term potential budget impact but also to use clinical expert testimony to identify when intended clinical use of a new treatment may be at a scale that would trigger access and affordability concerns. In such cases, the goal is to trigger discussions of possible policy steps to alleviate potential access restrictions or sudden sharp increases in health insurance premiums. The role of the potential budget impact analysis is not to suggest a cap on spending, but to signal to the health care system that special arrangements, such as lower prices, enhanced efforts to eliminate waste, or prioritizing treatment for the sickest, may be needed to ensure availability of the new drug without short-term adverse effects on patients and families seeking to pay for affordable health insurance.

5.2. Methods

The cost-effectiveness model in each economic evaluation is used to estimate the potential total budgetary impact of new treatments in the US, assuming different prices, including the treatment's list and net prices, and the three threshold prices to achieve cost effectiveness at \$50,000, \$100,000, and \$150,000 per QALY. Potential budget impact is defined as the total differential cost of using each new therapy rather than relevant existing therapy for the treated population, calculated as differential health care costs (including drug costs) minus any offsets in these costs from averted health care events or other aspects of treatment. The potential health care system budgetary impact of the intervention is explored over a five-year time horizon.

Potential budget impact analyses are based on net cost per patient across all sectors of health care spending, not just drugs. ICER uses epidemiologic and other data to estimate the size of the potential candidate population for each new treatment. For each threshold price, we then assume that an equal proportion of patients (20%) would be treated with the new treatment each year over five years, arriving at a cumulative 100% uptake at five years.

The analysis indicates when the potential budget impact threshold is reached at each combination of price and percent uptake among eligible patients at five years. This analysis does not attempt to estimate the uptake of a new intervention. Rather than try to estimate real-world uptake, the analysis presents information on a national level that allow stakeholders to ascertain the potential budget impact of a new service given a range of prices. The goal of ICER's potential budget impact analysis is to estimate the net cost per patient treated with new interventions so that decision-makers can use their own assumptions about uptake and pricing to determine their own estimate of potential budget impact.

Evidence Reports note the percent uptake of a new intervention, at its net price level, that would produce a potential budget impact that exceeds this threshold, or that a new intervention will not exceed the threshold regardless of uptake level. Results of the analysis are presented as a cumulative per-patient potential budget impact for each year over the five-year time horizon, with results being presented graphically for each intervention assessed, and numerical data presented in tabular format in an appendix of the report. The graph allows readers to see the average potential budget impact for a single patient over various time horizons from one to five years, and the estimated average net cost of treating a patient with an intervention relative to comparator(s) over the five years of the potential budget impact analysis. We also seek to produce calculations that will help policy makers identify situations in which the potential uptake of a new treatment, at various pricing levels, might exceed a budget impact threshold that signifies that the budget impact in the near term (over five years) would contribute to overall health care cost growth at a higher rate than growth in the national economy (plus 1%).

To accomplish these goals, ICER's potential budget impact analyses must evaluate whether a new drug would be likely to take market share from one or more drugs. The analysis uses clinical expert opinion regarding the treatments likely to be displaced by use of a new treatment within the eligible population. The procedures used in the analysis vary depending on whether and how many existing treatments are being displaced, with more details provided in ICER's [Reference Case](#) document. These are explicitly not meant to represent our assumptions of the budget impact of new interventions that are most likely in the real world. Our methods are intended to provide the calculations that can underpin a graphic figure that allows decision-makers and policy makers to make their own assumptions.

This analysis presents information allowing stakeholders to know when the combination of price and uptake of new drugs at the national level would lead to a potential budget impact that would exceed a threshold linked to a growth target for the overall health care system that would not significantly outpace growth in the overall US economy. Evidence Reports note the percent uptake of a new intervention, at its net price level, that would produce a potential budget impact that exceeds this threshold, or that a new intervention will not exceed the threshold regardless of uptake level. Results of the analysis are presented as a cumulative per-patient potential budget impact for each year over the five-year time horizon, with results being presented graphically for each intervention assessed, and numerical data presented in tabular format in an appendix of the report. The graph allows readers to see the average potential budget impact for a single patient over various time horizons from one to five years, and the estimated average net cost of treating a patient with an intervention relative to comparator(s) over the five years of the potential budget impact analysis.

The potential budget impact threshold for new drugs is calculated as double the average net budget impact for new drugs that would contribute to overall health care cost growth beyond the anticipated growth in national GDP plus an additional 1%. See Table 5.1. for the template for deriving the annual potential budget impact threshold. For services other than new drugs, potential budget impact is estimated but not compared to a potential budget impact threshold.

Table 5.1. Template for Annual Potential Budget Impact Threshold Calculation

Item	Parameter	Source
1	Growth in US GDP + 1%	World Bank
2	Total personal medical care spending	CMS National Health Expenditures
3	Contribution of drug spending to total health care spending (%) (Row 4 ÷ Row 2)	Calculation
4	Contribution of drug spending to total health care spending	CMS National Health Expenditures, Altarum Institute
5	Annual threshold for net health care cost growth for ALL drugs (Row 1 x Row 4)	Calculation
6	Average annual number of new molecular entity approvals over 5 years	FDA
7	Annual threshold for average cost growth per individual new molecular entity (Row 5 ÷ Row 6)	Calculation
8	Annual threshold for estimated potential budget impact for each individual new molecular entity (Doubling of Row 7)	Calculation

5.3 Access and Affordability Alert

Within the potential budget impact analysis section of each final report, ICER will include an “affordability and access alert” if discussion among clinical experts at the public meeting of ICER’s independent appraisal committees suggests that full, “clinically optimal” utilization at estimated net pricing (or at the \$150,000 per QALY threshold price if estimated net price is not available) would exceed the ICER annual potential budget impact threshold, without active intervention by insurers and others to manage access to the treatment. The affordability and access alert signals that the additional health care costs with a new intervention may be difficult for the health care system to absorb over the short term. In this situation, other needed services may be displaced, or health care insurance costs may rapidly rise, which would threaten sustainable access to high-value care for all patients.

6. Procedures for Report Generation, Stakeholder Engagement, and Public Processes

6.1 Topic Selection

ICER seeks to evaluate new pharmaceutical treatments and other health care interventions that may significantly improve patient outcomes, raise questions about their comparative effectiveness in relation to other treatment options, or that may have a significant financial impact on patients or the broader health system; several additional criteria are described on ICER's [website](#). When evaluating emerging drug therapies, ICER aims to issue reports and hold public meetings at or near the time of approval by the US Food and Drug Administration (FDA), so as to provide stakeholders with an independent evaluation of the benefits, risks, and economic considerations surrounding a new treatment at the time it is approved.

We identify potential topics through a process we call “horizon scanning,” which involves research led by ICER staff as well as discussions with stakeholders, including members of ICER's [advisory boards](#) and [independent voting committees](#) as well as other clinical societies, patient organizations, and industry sources including IPD Analytics, an independent organization that performs analyses of the emerging drug pipeline for a diverse group of industry stakeholders, including payers, pharmaceutical manufacturers, providers, and wholesalers. ICER also accepts topic [suggestions](#) at any time from members of the public. To maintain the independence of our evaluations, ICER does not accept funding to review a specific intervention or intervention(s) and the final selection of which topics to pursue is ICER's alone.

Once each year, ICER publicly issues a list of potential drug topics so as to provide stakeholders with an early view of ICER's current research priorities. Due to the dynamic nature of the drug development pipeline, it is possible that not every topic included in the list will be subject to an ICER review, just as it is possible that ICER may review other topics not included on the preliminary list. The final decision on the topic for each ICER review will be formally announced at the outset of the review process which begins approximately eight months (or ten months, if the review is of large class of treatments) prior to a corresponding public meeting held under the auspices of one of ICER's independent voting committees.

ICER sends the potential topic list to stakeholders who may be impacted by the review, including patient advocacy organizations, clinical societies, and the drug manufacturers whose products may be evaluated. ICER also schedules annual discussions with patient groups from major disease areas in which it has performed a review (i.e., rheumatoid arthritis, psoriasis, multiple sclerosis) to gather their perspective on the treatment landscape and any considerations that ICER should be aware of

prior to topic announcement. ICER provides early notifications to patient groups from other disease areas when it has high certainty that it will pursue an assessment pertaining to their focus.

6.2 Timelines

Standard and Class Reports

As described above, ICER assessments typically span eight to ten months, depending on the number of interventions included in a given review. Throughout this process, ICER engages with relevant stakeholders (i.e., patients and patient advocacy organizations, clinicians and specialty societies, drug makers, and payers) to ensure that ICER's report addresses questions relevant to decisionmakers and reflects the best available evidence at the time the report is released (stakeholder input opportunities are described in Section 6.3). ICER's process can be broadly divided into several phases, described in the list below, each of which builds on the work of previous phases. [Appendix Figures C1-C2](#) provide a week-by-week overview of milestones and stakeholder input opportunities within the phases described below.

1. **Open Input:** ICER notifies relevant stakeholders of its review and accepts input from stakeholders to inform its initial approach to the review.
2. **Scoping:** ICER seeks input from all stakeholders to inform its understanding of the disease area and available treatments, as well as what evidence it should seek.
3. **Draft Evidence Report:** ICER conducts the formal literature search and analysis of the clinical and economic evidence and issues a Draft Evidence Report for public comment.
4. **Revised Evidence Report:** ICER reviews public comments from stakeholders and revises the draft report as needed before issuing a revised Evidence Report, which serves as the foundation for the subsequent public meeting.
5. **Public Meeting:** ICER hosts a public meeting to present the findings of its revised report. One of ICER's independent appraisal committees deliberates and votes on key questions raised by the report. A policy roundtable of experts from the stakeholder community discusses how best to apply the evidence and votes to real-world practice and policy.
6. **Final Report:** ICER summarizes the public meeting proceedings (i.e., the votes and policy roundtable discussion) and issues its final report, which includes recommendations to inform policymaking and practice considerations.

Report Updates

ICER recognizes that new clinical or economic evidence may emerge following the posting of a final report that could change its conclusions. For example, new evidence could emerge demonstrating additional clinical benefits of therapy not captured in the studies available at the time of the original

review, or the introduction of a novel therapy may raise new questions about the relative benefits and risks of the therapeutic options for a condition.

ICER has developed two approaches to consider new evidence that may emerge shortly after the approval of a new therapy, described below. In addition to these two approaches, ICER may determine that an ad hoc New Evidence Update may be needed at any time after the release of a final report (i.e., if new evidence emerges before or after the 12-month report check-up process).

12-Month Report Check-Up

One year after it issues a Final Report and Meeting Summary, ICER will begin a three-month process to determine whether the findings of the initial report remain current (See [Appendix Figure C3](#) for a visual representation of this process).ⁱⁱⁱ During the first month, ICER will solicit input from manufacturers and participants on the policy roundtable of the initial public meeting about whether new information or treatments have emerged that warrant consideration as part of the update process. For example, this may be information that could lead to revision of a clinical evidence rating or a substantial shift in incremental cost-effectiveness results. ICER will review how new information would impact model results by comparing it with the range of inputs tested through one-way sensitivity analyses and probabilistic sensitivity analyses in the initial report's model. If no new evidence has been identified, ICER will issue a statement describing why we believe the original report is still current and will mark the report to reflect this judgment.

If new information that potentially warrants an update is identified by stakeholders, ICER will update the literature search from its initial report to more systematically determine if any additional evidence should be considered. It may be the case that while new evidence is identified, either in the published or grey literature, it is not likely to meaningfully change the conclusions of ICER's initial report. For example, this could occur if preliminary outputs from a long-term follow-up trial confirm the initial benefits described in the pivotal trials, but do not offer any new evidence on additional outcomes. In such cases, ICER will issue a statement describing the evaluated evidence and the rationale for why the original report does not require an update; the initial report will then be marked as still current.

If, alternatively, the new information identified by stakeholders and/or ICER is likely to substantially impact the findings of the original report, we will issue a statement describing how the evidence may change our findings and will add language to the original report to indicate it is no longer current. The statement will also include a recommendation on whether the report would require a brief New Evidence Update or a full update. Note, however, that ICER must balance the need to revisit prior reports with its goal of evaluating important emerging therapies. As a result, ICER may not have the resources to begin an update immediately when issuing a report check-up statement

ⁱⁱⁱ The first review eligible for 12-Month Report Check-Up will be the assessment started in January 2020. Earlier assessments may be updated on an ad hoc basis.

and, in some cases, may not update the prior report once it has been marked as no longer current. We believe that clearly marking our prior reports to indicate whether new developments may impact the findings is the best way to signal to stakeholders whether the information as presented within the report remains actionable.

A New Evidence Update would typically be required when there is new data on a small number of key outcomes for a limited subset of the interventions included in the original review and is a standalone document that evaluates the impact of this evidence on the prior report conclusions. This form of update will not typically be presented at a public meeting but will instead be posted to ICER's website and disseminated to stakeholders. A full update, in contrast, would be recommended when new evidence is available for many or most of the originally assessed interventions such that revising the entirety of the original report is necessary. Full updates will follow the standard or class review timelines described under the "Standard and Class Reports" heading, including presentation at a public meeting.

24-Month Real-World Evidence Update Pilot

As part of the 2020-2023 Framework update, ICER will begin a pilot program to generate new RWE to inform updated assessments of therapies approved by the FDA under accelerated approval pathways. The goal of the pilot will be to supplement the comparatively limited evidence base that often accompanies accelerated approvals with real-world evidence to provide stakeholders with a more comprehensive understanding of the early impacts of these therapies. The process will begin at the 24-month anniversary of the Final Report posting, and will span a period of several months, depending on the nature of the evidence to be generated. For the pilot effort, ICER will identify a drug that it has already reviewed and that has been on the market for at least 24 months; we will provide additional details on the process before undertaking the first of these updates.

6.3 Stakeholder Engagement

ICER's belief is that collaborative efforts among these stakeholders, grounded in a civil and honest discussion of evidence on effectiveness and value, is essential to drive lasting improvements to the health care system on behalf of current and future patients. From the outset of each review, ICER actively engages and seeks input from patients, caregivers and patient advocacy organizations; clinical experts; drug manufacturers; and payers (i.e., public/private insurers, pharmacy benefit managers, and purchasers). Each of these stakeholders brings important expertise to ICER's process and is, ultimately, affected by any policy action that is catalyzed by ICER reviews.

Table 6.1 provides broad overview of the formal stakeholder input opportunities by report phase. Additional details about these opportunities can be found on ICER's [website](#), which includes links to

dedicated engagement guides for drug manufacturers and patients/patient advocacy organizations, as well as information on logistical considerations such as formatting requirements and page limits for individual comment opportunities. Subsequent sections of this chapter provide broad information on engagement opportunities by stakeholder type.

Table 6.1. Overview of Formal Stakeholder Input Opportunities

Review Phase	ICER Public Documents and Events	Stakeholder Input Opportunities	Potential Impact of Stakeholder Input
Pre-Announcement	<ul style="list-style-type: none"> None 	<ul style="list-style-type: none"> Early engagement with patient groups 	<ul style="list-style-type: none"> Initial conversations about potential patient survey
Open Input	<ul style="list-style-type: none"> Topic Announcement 	<ul style="list-style-type: none"> Open input submissions (written) Patient input survey (written) Key informant interviews (discussion) 	<ul style="list-style-type: none"> Informs initial research plan, background knowledge of condition, understanding of patient experience
Scoping	<ul style="list-style-type: none"> Draft Scoping Document Revised Scoping Document Public comments received on Revised Scoping Document 	<ul style="list-style-type: none"> Public comment on draft scope (written)* Key informant interviews (discussion) 	<ul style="list-style-type: none"> Continuation of above impact Refinement of specific research plan described in draft scope
Draft Evidence Report	<ul style="list-style-type: none"> Research Protocol Modeling Analysis Plan Draft Evidence Report Draft Voting Questions 	<ul style="list-style-type: none"> Feedback on preliminary model (written and discussions; invited stakeholders†) Public comment on draft report (written)* Formal patient survey‡ 	<ul style="list-style-type: none"> Continuation of above impact Informs interpretation of evidence Provides insights into considerations not represented in evidence base Refinement of analyses
Evidence Report	<ul style="list-style-type: none"> Evidence Report Revised voting questions Public comments received on draft report ICER response to public comments 	<ul style="list-style-type: none"> See above public comment opportunity on draft report 	<ul style="list-style-type: none"> Continuation of above impact Revisions to quantitative and qualitative aspects of ICER report
Public Meeting	<ul style="list-style-type: none"> Public meeting (in-person attendance or via webcast) 	<ul style="list-style-type: none"> Oral Public Comments* Oral Public Comment Summary (written)* Participation on Policy Roundtable (invited stakeholders only)* 	<ul style="list-style-type: none"> Informs discussion and CEPAC votes Informs policy recommendations discussed during Policy Roundtable
Final Evidence Report	<ul style="list-style-type: none"> Final Evidence Report and Meeting Summary “Report-at-a-Glance” Summary 	<ul style="list-style-type: none"> Post-meeting debriefs with key patient organizations 	<ul style="list-style-type: none"> Final report reflects the stakeholder input gathered throughout review, including policy guidance

*Denotes stakeholder input that is publicly released in original form (the commenter is publicly identifiable)

†Manufacturers and other stakeholders who are able to provide detailed feedback on the technical aspects of the preliminary model structure, assumptions, and inputs.

‡A formal survey may not be conducted for every review.

Patient Engagement Program

From the outset of each review, ICER seeks input from patients and patient advocacy organizations to gather their perspectives on the patient, family, and caregiver experience of a disease and its treatments. These opportunities include participation in the formal input opportunities described in Table 6.1 as well as ad hoc touchpoints. ICER's approach to patient engagement is intended to provide a flexible system under which individual patients can describe their lived experiences and advocacy organizations can contribute in multiple ways depending on their organizational focus and resources (i.e., through dissemination of ICER's [patient input questionnaire](#), phone discussions, written comments on ICER's research plan and draft findings, participation during a public meeting as a commenter and/or Policy Roundtable participant). Detailed information and guidance for patients and advocacy organizations can be found in ICER's [Patient Participation Guide](#).

With the launch of the 2020-2023 Value Assessment Framework, ICER will continue its efforts to expand and formalize its Patient Engagement Program. This work will involve consolidating and updating ICER's current materials surrounding patient engagement, as well as several new and enhanced initiatives intended to augment patient and patient organization's ability to inform ICER's research, including:

- Outreach to patient groups in advance of topic announcement to provide an overview of ICER procedures and to facilitate connections with other patient leaders who have participated in an ICER review. For major therapeutic classes (i.e., immunomodulators for which ICER has performed a class review, treatments for multiple sclerosis, etc.), ICER will schedule an annual conference call or meeting to discuss the emerging pipeline of new treatments, get patient input on key priorities, and explore opportunities to gather new data on outcomes of care that are important to patients and families. ICER provides early notifications to patient groups from other disease areas when it has high certainty that it will pursue an assessment pertaining to their focus;
- Work with patient groups at the outset of each review to determine whether/how patient groups can contribute empirically to the economic model;
- Identify gaps in key outcomes measures and work with patient groups to identify sources of RWE and potentially develop patient survey to inform the economic model and provide additional information regarding potential other benefits or disadvantages and contextual considerations surrounding the disease and its treatment;
- At the conclusion of each review,
 - Invite participating patient groups for a formal debrief on the experience
 - Offer to co-write and promote a letter to FDA and other stakeholders with proposals for improving the generation of patient-relevant data as part of the drug development process

ICER's early outreach focuses on identifying the outcomes of greatest importance to the patient community, including related evidence, which informs ICER's selection of outcomes measures to include and prioritize in its clinical assessment, as well as the selection of cost-consequence measures used in ICER economic models. Learnings from these and subsequent conversations also inform a dedicated chapter in each ICER report on the patient experience^{iv} that precedes sections describing the clinical and economic evidence. This sequence ensures that readers and appraisal committees are presented with information on patient perspectives in the early pages of each assessment, allowing them to interpret the subsequent evidence and analyses through the lens of the patient experience.

Engagement opportunities during the Evidence Report phases of the report include conducting formal surveys to elicit patient preferences for treatment and potentially to inform model inputs (e.g., health state utilities) when not described in the published literature, as well a public comment period on the draft Evidence Report. ICER seeks to involve a representative from a patient organization as an expert reviewer on a pre-publication version of the draft report to ensure the accuracy and comprehensiveness of sections describing the patient experience.

ICER public include patients and advocacy organization representatives on Policy Roundtables, and also provide an opportunity for oral public comments prior to the appraisal committee vote. There are typically eight overall participants on the roundtable, of which two seats are reserved for patients and representatives from advocacy organizations. Patient and advocacy representatives sit at the main session table throughout the day, providing insight and commentary to the appraisal committee members as they review the evidence presentation and while the committee deliberates and votes. They then participate during the formal roundtable discussion to including on discussions pertaining to insurance coverage policy and future research needs.

Clinical Experts

ICER seeks input from clinical experts throughout its review process. Initial outreach begins around the Open Input period and informs ICER's understanding of how clinicians weigh available treatment options and how emerging therapies may fit into current practice patterns, as nuances contained in the clinical evidence base for approved and investigational agents, and identifies key sources of evidence (i.e., published research, grey literature, conference proceedings) that ICER may consider during its review. During later stages of the review process, ICER seeks input from clinical experts to validate its interpretation of the clinical evidence; these opportunities include ad hoc outreach for advice during ICER's systematic review and analysis of evidence, and through participation as a formal expert reviewer of a pre-posting version of ICER's draft report.

^{iv} This information was part of the background section (Section 1) of older reports. The review associated with the March 2020 public meeting will be the first to feature a standalone chapter and subsequent reviews will follow the same approach.

Clinical experts serve in a role analogous to that of patients and patient advocacy organization representatives during public meetings, providing input and guidance to the appraisal committee throughout the presentation of the evidence and deliberation on voting questions, as well as playing an active role during the subsequent roundtable discussion. In addition to participating in discussions about what may represent clinically-reasonable insurance coverage policy and future research needs, clinical experts may also be asked to comment on how an emerging therapy may change clinical practice and key questions that should be resolved by clinical specialty societies to promote evidence-based medicine.

Manufacturers

ICER's outreach to manufacturers begins at the Open Input period and is focused on companies that produce the interventions of interest, but also includes manufacturers of branded comparator products. ICER typically schedules scoping discussions with manufacturers after the release of a draft scoping document so that they have an opportunity to reflect on ICER's proposed research and modeling approach. At the end of the scoping phase, ICER also issues a request for data to manufacturers, responses to which inform ICER's clinical and economic evaluations. ICER has developed a policy under which it can accept "data in confidence" from drug developers, details of which can be found on ICER's [website](#).

During the draft report phase, ICER offers several additional opportunities for manufacturers to comment on its draft approach, including through ad hoc discussions surrounding release of a research protocol and participation in the "Preliminary Model Structure, Assumptions, and Inputs" presentation, which is followed by subsequent discussion and an additional opportunity to provide data to inform ICER's modeling effort. Manufacturers are also invited to participate in a formal model sharing program, which is described later in this section, and provides an opportunity for evaluation of executable models.

Manufacturers are invited to participate in ICER's public meeting through participation in an oral public comment session and as formal participants on the policy roundtable, where they provide insights into topics such as pricing, perspectives on insurance coverage policy, and clinical trial design and outcomes selection.

Payers

ICER outreach to payers begins during the Open Input period and, as with manufacturers, involves a scoping call after the release of a draft scoping document. These conversations inform selection of treatments of interest, comparators, and key outcomes to help ensure that ICER's research answers questions central to the development of evidence-based coverage policy; as well as ICER's initial understanding of payer approaches. Engagement from this period through the report phase is typically done on an ad hoc basis and may include invitations to provide RWE through claims

analyses to inform aspects of the model such as adherence rates. At public meetings, payers participate on the Policy Roundtable to discuss considerations around coverage policy development and the intersection of pricing, access, and affordability.

Additional Opportunities for Stakeholder Input

Identification of Low-Value Care

In its reports, ICER seeks to include information on wasteful or lower-value services in the same clinical area that could be reduced or eliminated to create “headroom” in health care budgets for higher-value innovative services. These services, including treatments and mechanisms of care, are ones that would *not* be directly affected by the intervention under review, as these would be captured in the economic model (e.g., an effective intervention for acute pain management that would reduce emergency department visits). Rather, these services are those used in the current management of the condition that represent ineffective or overused approaches to care (e.g., use of imaging for uncomplicated headache).³⁵ The goal of this section is to highlight for policymakers the opportunities for reallocating resources from lower value services in order to help make headroom for the added cost of high-value drugs and other high-value services.

ICER request input on these categories from patients, clinicians, manufacturers and payers through requests in draft and revised scoping documents, draft reports, and during discussion calls. Services that meet the criteria described above may be included in ICER reports with attribution to the organization that identified the service, as well as citations provided by the commenting organization.

Economic Model Transparency

ICER’s approach to economic model transparency is based on the Modeling Good Research Practices Task Force report on “Model Transparency and Validation” jointly produced by the International Society for Pharmacoeconomics and Outcomes Research (ISPOR) and the Society for Medical Decision-Making (SMDM).³⁶ We aim to include information in each report that describes model structure and processes, all major inputs and sources for data, and key assumptions used in our economic analyses, so that readers can judge their confidence in the results while preserving the intellectual property rights of those we collaborate with.

ICER’s commitment to open and transparent engagement extends to the sharing of economic models with drug manufacturers to provide an opportunity to comment on executable versions of the model. This is supplemented by an earlier touchpoint during week 15 of an assessment called a “Preliminary Model Structure, Inputs, and Assumptions Presentation,” during which manufacturers and other invited stakeholders with expertise in economic modeling may provide feedback on an

early version of the model before ICER posts a draft report. At the same time, ICER publicly releases Modeling Analysis Plan to the [Open Science Framework](#) website describing its modeling approach.

Detailed information about ICER’s Economic Model Transparency Program can be found on ICER’s [website](#), with additional information included in the manufacturer engagement guide described earlier in this section.

6.4 Public Meetings

Structure and Purpose

ICER hosts a public meeting for each assessment that can be broken down into four broad stages: 1) presentation of the revised Evidence Report findings, 2) testimony and discussion with manufacturer representatives and patient/public commenters; 3) deliberation and vote by the independent appraisal committee on key questions surrounding the clinical and economic evidence; and 4) a policy roundtable discussion with patients, clinical experts, manufacturers, and payers to explore how to apply the evidence and votes to real-world practice and policy. This stepwise process represents ICER’s goal to facilitate decision-making that is grounded in a thorough and public exploration of the evidence. Importantly, ICER does not issue formal policy recommendations prior to a public meeting to reflect the reality that analysis of the evidence does not, in isolation, provide “the” answer to the complex circumstances surrounding pricing, coverage policy, and clinical practice; rather, our analysis serve as the foundation for discussions surrounding these topics.

Stakeholders are involved during each broad meeting phase, details of which can be found in Table 6.2 below. Each meeting is open to the public, with a webcast available for those who cannot attend in person, and a recording is posted to ICER’s website within a few days of the conclusion of the meeting.

Table 6.2. Public Meeting Agenda Overview

Agenda Item	Primary Participants
1. Presentation of the Evidence and Economic Modeling, Q&A/Discussion	ICER staff and consultants, appraisal committee, patient and clinician members of the policy roundtable, manufacturers (as needed), patient advocacy organizations (dependent on topic)
2. Manufacturer Public Comments and Discussion	Manufacturers, ICER staff and consultants, appraisal committee
3. Public Comments from Patients, Clinicians, and Public	Patients, clinicians, payers, researchers, and other stakeholders, appraisal committee
4. Voting on Clinical Effectiveness and Value Questions; Additional Discussion	Moderator; appraisal committee; clinical, patient, and subject-matter experts from the policy roundtable; manufacturers (as needed)
5. Policy Roundtable Discussion	Moderator, appraisal committee, policy roundtable
6. Reflections from Voting Panel	Moderator, appraisal committee
7. Summary and Closing Remarks	Moderator

Independent Appraisal Committees

Each public meeting involves deliberation and voting on key questions related to the Evidence Report by an independent appraisal committee. ICER currently convenes three such committees: the New England and the Midwest Comparative Effectiveness Public Advisory Councils, and the California Technology Assessment Forum. These committees are standing bodies (i.e., they do not change from one meeting to the next), and members are recruited for their clinical and policy expertise technology assessment, including research methods, economic analysis, evidence-based practice, and patient advocacy, among other disciplines. All members meet strict conflict of interest [requirements](#) to limit any bias that may be introduced by the presence of certain personal or financial relationships. One implication of this approach is that, by design, ICER appraisal committees do not necessarily include those affected by the condition under review, whether they are individual patients or practicing clinicians, though this may occur from time to time (i.e., a neurologist may serve on an appraisal committee for a neurology topic, provided he or she does not have any disqualifying conflicts). This approach aligns with that of many other organizations, including the United States Preventive Services Task Force ([USPSTF](#)) and all international HTA organizations.

ICER recognizes how vital the patient and clinical expert perspective is to our review process and public meeting, which is why we seek input from patient and clinical experts throughout the report development process, and by including several such experts as active participants as throughout our public meetings, including in the development of any policy recommendations that emerge from the voting results. This approach provides members of ICER appraisal committees with sufficient insight into the patient experience and clinical practice to inform voting.

Additional information about these independent appraisal committees, including current membership and conflict of interest criteria, can be found on ICER's [website](#).

Categories of Questions

ICER develops voting questions to ensure that the questions are framed to address the issues most important in applying the evidence to practice, price negotiations, and insurance coverage policy. Appraisal committees vote on three categories of questions pertaining to the comparative clinical effectiveness of treatments; any broader potential other benefits, disadvantages, or contextual considerations that may complement the clinical evidence; and, finally, the long-term value for money of an intervention at current prices (if no prices are available at the time of the public meeting, votes will be taken regarding which end of the health-benefit price benchmark range is most appropriate as a guide for price negotiations). Results of the vote, which are accompanied by public deliberation, are intended to provide stakeholders and policymakers with insights into the implications of the evidence for practice and policy. These votes represent the judgments of Appraisal Committee members as individuals and not as representatives of their employers. All votes are taken in public with supporting advice provided by the authors of the report and, as mentioned earlier, invited experts from the patient and clinical provider community.

In the sections below, we present example voting questions and provide commentary on the meaning of voting responses to provide stakeholders with an understanding of their implications for policy and practice.

Comparative Clinical Effectiveness

Comparative clinical effectiveness votes are meant to present a judgment of the *net health benefit* of a treatment versus a comparator for a population or subpopulation of patients. This foundational question is designed to establish whether current evidence is sufficient to demonstrate that one intervention is superior to another for patient populations or subpopulations within a given condition. As part of the deliberation process, appraisal committee members formally weigh the relative magnitude of differences in risks and benefits, as well as the relative confidence that the body of evidence can provide regarding the accuracy of estimates of risks and benefits. These questions are typically framed using the following language:

For [patients with condition X], is the evidence adequate to demonstrate that the net health benefit of [intervention A] is greater than that of [intervention B]?

Yes / No

The implication of a “yes” vote is relatively straightforward, indicating that there is sufficient evidence to demonstrate one intervention is superior to another for the specified patient population. This can imply, among other possibilities, that clinicians may wish to consider using this

therapy preferentially for such patients, and that insurers may wish to develop coverage policies that prioritize access for patients who are likely to benefit for intervention A over intervention B.

By contrast, the implications of a “no” vote may be less readily apparent. Such a vote does not necessarily mean that a treatment is ineffective; rather, it indicates that the current evidence base is insufficient to demonstrate effectiveness in the specified patient population. In this way, a “no” vote can signal to stakeholders that additional research is necessary to determine an intervention’s benefits and risks to patients.

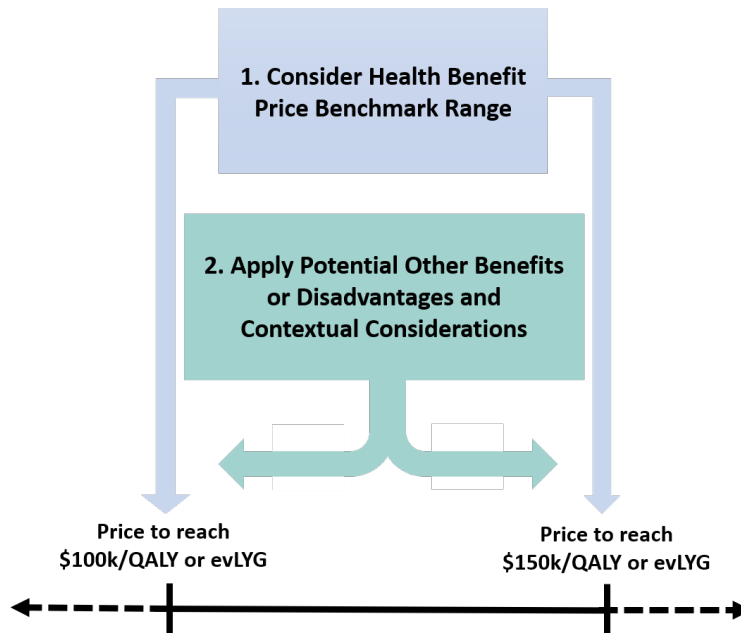
Potential Other Benefits and Contextual Considerations

Next, appraisal committees deliberate on whether an intervention provides any significant “potential other benefits or disadvantages” or impacts any broader “contextual considerations” relative to the comparator of interest. As described in [Section 4](#), these categories are intended to represent benefits or disadvantages offered by the intervention to the individual patient, caregivers, the delivery system, other patients, or the public that may not have been considered or captured in the evidence on comparative clinical or cost effectiveness. These are meant to be considered independent of cost.

Appraisal committee members will vote on a 1-3 Likert scale described below and in Section 4, and votes will be displayed as tallies. As noted earlier, while the specific implications of each numerical vote may vary due to the language used in the consideration, they are broadly meant to indicate the guidance for decision-makers making judgments of the overall long-term value for money of interventions. A schematic view of this approach is shown in Figure 6.1 below, and general interpretation of the votes on potential other benefits and contextual considerations is as follows:

- **1:** Within the general range suggested by clinical evidence and findings of cost-effectiveness analysis, consideration of potential other benefits or disadvantages and/or contextual considerations points toward relatively lower longer-term value for money
- **2:** Within the general range suggested by clinical evidence and findings of cost-effectiveness analysis, consideration of potential other benefits or disadvantages and/or contextual considerations points toward relatively intermediate longer-term value for money
- **3:** Within the general range suggested by clinical evidence and findings of cost-effectiveness analysis, consideration of potential other benefits or disadvantages and/or contextual considerations points toward relatively higher longer-term value for money

Figure 6.1. Conceptual Guide to Application of “Potential Other Benefits or Disadvantages” and “Contextual Considerations” to Judgements of Value

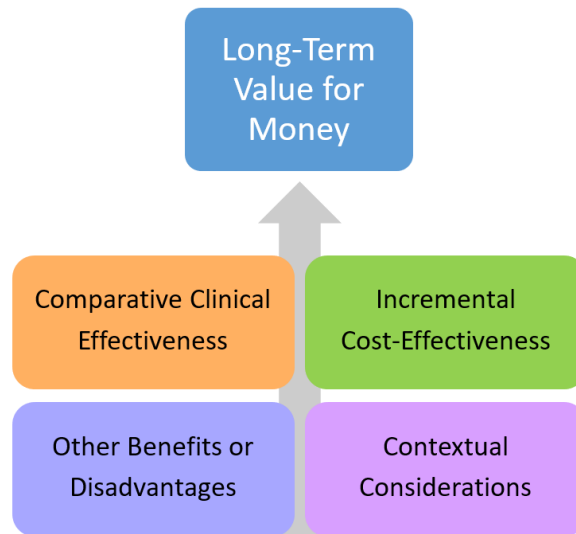


evLYG: equal value life-years gained, QALY: quality-adjusted life years

Long-Term Value for Money

Each Appraisal Committee will be asked to integrate all value dimensions into its vote on long-term value for money. Committees do not vote on cost-effectiveness results in isolation, but rather as a component of a broader concept that integrates comparative clinical effectiveness, potential other benefits or disadvantages, and contextual considerations (Figure 6.1), consistent with recommendations by the Second Panel on Cost Effectiveness.

Figure 6.2. Conceptual Structure of Long-Term Value for Money



Cost-effectiveness analyses will have been done using either established net prices for interventions or estimated net prices (a “placeholder” price) based on analyst forecasts and/or guidance from other sources. When voting on the long-term value for money at current pricing for therapies with known prices, the voting questions are posed as follows:

Given the available evidence on comparative clinical effectiveness and incremental cost effectiveness, and considering other benefits and contextual considerations, what is the long-term value for money of [intervention A] compared to [intervention B] for [patients with condition X]?

High... / Intermediate... / Low Long-Term Value for Money at Current or Estimated Prices

This vote will also be taken in certain circumstances when there is no known net price. If the estimates of net price are felt to be extremely reliable, whether by market history and/or by pledges by manufacturers, the appraisal committees will vote on the intervention’s long-term value for money at current pricing.

Policy Roundtable Discussion

Each public meeting culminates in a discussion of how to apply the evidence and appraisal committee votes to real-world practice and policy. All stakeholders participate in this discussion, and the typical composition of a policy roundtable includes two patients and/or representatives from patient advocacy organizations; two clinical experts in the topic under review; two representatives from the manufacturers of the therapies under review; and two payer/purchaser

representatives (i.e., public and/or private insurers, pharmacy benefit managers, or employers). While the specific topics of discussion may vary from one meeting to the next, the broad themes of these discussions are generally consistent and include discussion of:

- Evidence-based insurance coverage policy (i.e., patient eligibility criteria, special considerations for patient subpopulations, step therapy, provider criteria, etc.)
- Pricing and payment mechanisms (i.e., outcomes-based contracting and other innovative approaches to payment)
- Future research needs (i.e., study of additional outcomes measures, long-term data needs, key questions that can be addressed by real-world evidence, etc.)
- The guidance that clinical specialty societies and patient organizations should provide to their communities

These conversations serve as the foundation for policy recommendations for stakeholders that are included in each final report.

References

1. Ollendorf DA, Pearson SD. An integrated evidence rating to frame comparative effectiveness assessments for decision makers. *Med Care*. 2010;48(6 Suppl):S145-152.
2. CER Collaborative, Academy of Managed Care Pharmacy (AMCP), National Pharmaceutical Council (NPC), International Society for Pharmacoeconomics Research (ISPOR). Comparative Effectiveness Research Tool. <https://www.cercollaborative.org/global/default.aspx?RedirectURL=%2fhome%2fdefault.aspx>. Accessed October 11, 2019.
3. Cook DJ, Mulrow CD, Haynes RB. Systematic reviews: synthesis of best evidence for clinical decisions. *Ann Intern Med*. 1997;126(5):376-380.
4. Moher D, Liberati A, Tetzlaff J, DG A. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Annals of Internal Medicine*. 2010;8(5):336-341.
5. Pearson SD, Dreitlein WB, Towse A, Hampson G, Henshall C. A framework to guide the optimal development and use of real-world evidence for drug coverage and formulary decisions. *Journal of Comparative Effectiveness Research*. 2018;7(12):1145-1152.
6. Garrison LP, Jr., Neumann PJ, Erickson P, Marshall D, Mullins CD. Using real-world data for coverage and payment decisions: the ISPOR Real-World Data Task Force report. *Value Health*. 2007;10(5):326-335.
7. US Preventive Services Task Force. *Procedure Manual*. 2008.
8. Ollendorf DA, Pearson SD. ICER Evidence Rating Matrix: A User's Guide. Institute for Clinical and Economic Review. <https://icer-review.org/methodology/icers-methods/icer-evidence-rating-matrix/>. Published 2020. Updated January 31, 2020. Accessed January 31, 2020.
9. Sawaya GF, Guirguis-Blake J, LeFevre M, Harris R, Petitti D. Update on the methods of the U.S. Preventive Services Task Force: estimating certainty and magnitude of net benefit. *Ann Intern Med*. 2007;147(12):871-875.
10. Guyatt GH, Oxman AD, Vist GE, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *Bmj*. 2008;336(7650):924-926.
11. Agency for Healthcare Research and Quality. Methods Guide for Effectiveness and Comparative Effectiveness Reviews. https://effectivehealthcare.ahrq.gov/sites/default/files/pdf/cer-methods-guide_overview.pdf. Published 2014. Accessed.
12. Sanders GD, Neumann PJ, Basu A, et al. Recommendations for Conduct, Methodological Practices, and Reporting of Cost-effectiveness Analyses: Second Panel on Cost-Effectiveness in Health and Medicine. *JAMA*. 2016;316(10):1093-1103.
13. Guidelines for the economic evaluation of health technologies: Canada. In: 4th ed. Ottawa: CADTH; 2017: https://www.cadth.ca/sites/default/files/pdf/guidelines_for_the_economic_evaluation_of_health_technologies_canada_4th_ed.pdf.
14. National Institute for Health and Care Excellence (NICE). Guide to the methods of technology appraisal 2013. nice.org.uk/process/pmg9. Published 2013. Accessed January 27, 2020.
15. Institute for Clinical and Economic Review. The QALY: Rewarding the Care That Most Improves Patients' Lives. ICER. https://icer-review.org/wp-content/uploads/2018/12/QALY_evLYG_FINAL.pdf. Published 2018. Accessed January 27, 2020.
16. Pickard AS, Law EH, Jiang R, et al. United States Valuation of EQ-5D-5L Health States Using an International Protocol. *Value Health*. 2019;22(8):931-941.

17. Shaw JW, Johnson JA, Coons SJ. US valuation of the EQ-5D health states: development and testing of the D1 valuation model. *Med Care*. 2005;43(3):203-220.
18. Neumann PJ, Sanders GD, Russell LB, Siegel JE, Ganiats TG. *Cost-effectiveness in health and medicine*. New York: Oxford University Press; 2016.
19. Davis S. NICE Decision Support Unit Methods Development. In: *Assessing Technologies That Are Not Cost-Effective at a Zero Price*. London: National Institute for Health and Care Excellence (NICE); 2014.
20. Woods B, Revill P, Sculpher M, Claxton K. Country-Level Cost-Effectiveness Thresholds: Initial Estimates and the Need for Further Research. *Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research*. 2016;19(8):929-935.
21. Lomas JRS, Martin S, Claxton KP. Estimating the marginal productivity of the English National Health Service from 2003/04 to 2012/13. In. York, UK: Centre for Health Economics, University of York; 2018.
22. Claxton K, Lomas J, Martin S. The impact of NHS expenditure on health outcomes in England: Alternative approaches to identification in all-cause and disease specific models of mortality. *Health Econ*. 2018;27(6):1017-1023.
23. Vanness D. Deriving an opportunity cost-based threshold for CEA in the United States. Paper presented at: International Health Economic Association; July, 2017; Boston, MA.
24. Pauly MV. The Questionable Economic Case for Value-Based Drug Pricing in Market Health Systems. *Value Health*. 2017;20(2):278-282.
25. Lakdawalla DN, Doshi JA, Garrison LP, Jr., Phelps CE, Basu A, Danzon PM. Defining Elements of Value in Health Care-A Health Economics Approach: An ISPOR Special Task Force Report. *Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research*. 2018;21(2):131-139.
26. Clark S, Weale A. Social values in health priority setting: a conceptual framework. *Journal of health organization and management*. 2012;26(3):293-316.
27. National Institute for Health and Care Excellence (NICE). Consultation Paper: Value Based Assessment of Health Technologies. In.
28. Reckers-Droog VT, van Exel NJA, Brouwer WBF. Looking back and moving forward: On the application of proportional shortfall in healthcare priority setting in the Netherlands. *Health policy (Amsterdam, Netherlands)*. 2018;122(6):621-629.
29. Versteegh MM, Ramos IC, Buyukkaramikli NC, Ansaripour A, Reckers-Droog VT, Brouwer WBF. Severity-Adjusted Probability of Being Cost Effective. *Pharmacoeconomics*. 2019.
30. Ottersen T, Forde R, Kakad M, et al. A new proposal for priority setting in Norway: Open and fair. *Health Policy*. 2016;120(3):246-251.
31. van de Wetering EJ, Stolk EA, van Exel NJ, Brouwer WB. Balancing equity and efficiency in the Dutch basic benefits package using the principle of proportional shortfall. *Eur J Health Econ*. 2013;14(1):107-115.
32. Stolk EA, van Donselaar G, Brouwer WB, Busschbach JJ. Reconciliation of economic concerns and health policy: illustration of an equity adjustment procedure using proportional shortfall. *Pharmacoeconomics*. 2004;22(17):1097-1107.
33. Claxton K, Sculpher M, Palmer S, Culyer AJ. Causes for concern: is NICE failing to uphold its responsibilities to all NHS patients? *Health Econ*. 2015;24(1):1-7.
34. American College of Radiology. Choosing Wisely - Ten Things Physicians and Patients Should Question. <https://www.choosingwisely.org/societies/american-college-of-radiology/>. Published 2012. Updated 2019. Accessed January 14, 2020.

35. Eddy DM, Hollingworth W, Caro JJ, Tsevat J, McDonald KM, Wong JB. Model transparency and validation: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force-7. *Med Decis Making*. 2012;32(5):733-743.
36. Linley WG, Hughes DA. Societal views on NICE, cancer drugs fund and value-based pricing criteria for prioritising medicines: a cross-sectional survey of 4118 adults in Great Britain. *Health Econ*. 2013;22(8):948-964.
37. Karlsberg Schaffer S, Cubi-Molla P, Devlin N, Towse A. Shaping the research agenda to estimate relevant cost-effectiveness thresholds for health technology assessment decision making: report for ABPI. In: Office of Health Economics Consulting; 2016.
38. Ryen L, Svensson M. The Willingness to Pay for a Quality Adjusted Life Year: A Review of the Empirical Literature. *Health Econ*. 2015;24(10):1289-1301.
39. Hutubessy R, Chisholm D, Edejer TT-T, Who C. Generalized cost-effectiveness analysis for national-level priority-setting in the health sector. *Cost Effectiveness and Resource Allocation*. 2003;1(1):8.
40. World Health Organization. WHO Commission on Macroeconomics and Health & World Health Organization. (2001). Macroeconomics and health : investing in health for economic development : executive summary / report of the Commission on Macroeconomics and Health. <https://apps.who.int/iris/handle/10665/42463>. Published 2001. Accessed August 19, 2019.
41. Marseille E, Larson B, Kazi DS, Kahn JG, Rosen S. Thresholds for the cost-effectiveness of interventions: alternative approaches. *Bull World Health Organ*. 2015;93(2):118-124.
42. Bertram MY, Lauer JA, De Joncheere K, et al. Cost-effectiveness thresholds: pros and cons. *Bull World Health Organ*. 2016;94(12):925-930.
43. Robinson LA, Hammitt JK, O’Keeffe L. Valuing Mortality Risk Reductions in Global Benefit-Cost Analysis. *Journal of Benefit-Cost Analysis*. 2019;10(S1):15-50.
44. Hirth RA, Chernew ME, Miller E, Fendrick AM, Weissert WG. Willingness to Pay for a Quality-adjusted Life Year: In Search of a Standard. *Medical Decision Making*. 2000;20(3):332-342.
45. Phelps CE. A New Method to Determine the Optimal Willingness to Pay in Cost-Effectiveness Analysis. *Value Health*. 2019;22(7):785-791.
46. Garber AM, Phelps CE. Economic foundations of cost-effectiveness analysis. *J Health Econ*. 1997;16(1):1-31.
47. Arrow K. *Essays in the Theory of Risk-Bearing*. 2016/08/17 ed. Amsterdam: North-Holland Publishing Company; 1971.
48. US Bureau of Economic Analysis. Personal Income Per Capita [A792RC0A052NBEA]. In. Vol 2019: Federal Reserve Bank of St. Louis; 2018.
49. Maryland Health Services Cost Review Commission. Maryland's Total Cost of Care Model: Update for Providers. <https://hsrcr.state.md.us/Pages/tcocmodel.aspx>. Accessed August 20, 2019.
50. Massachusetts Health Policy Commission. Health Care Cost Growth Benchmark. <https://www.mass.gov/info-details/health-care-cost-growth-benchmark>. Published 2019. Accessed August 19, 2019.

Appendices

A. Glossary

Adverse event – A harmful, unintended occurrence during or after the administration of a health technology, which may or may not be caused by the use of that technology.

Analytic framework – A graphical representation of the key assumptions the systematic review seeks to evaluate in order to determine the net health benefit of the intervention(s) under appraisal. The framework diagrams linkages between an intervention and its possible benefits and harms in a given population.

Base-case analysis – the analysis using the initial set of assumptions and input parameter values.

Budget impact – an estimate of the projected cumulative resource expenditure for a particular intervention in a specific population over a period of time.

Cost-Effectiveness Acceptability Curve – A graph that plots the percentage of simulations that result in ICERs that fall at or under different cost-effectiveness thresholds.

CEPAC – Comparative Effectiveness Public Advisory Council; ICER convenes public meetings of two regionally-focused appraisal committees based in New England and the Midwest to review objective evidence reports and develop recommendations for how stakeholders can apply evidence to improve the quality and value of health care. The mission, processes, and role of the CEPAC programs are the same as that of CTAF, despite a different naming convention.

Clinical effectiveness – The degree of health benefit produced by an intervention.

Comparator – an alternative health technology against which an intervention is evaluated.

Cost-effectiveness analysis – a type of economic evaluation in which an outcome is measured in incremental costs per incremental health unit, such as life years gained, or clinical event avoided.

Cost-Effectiveness Threshold – the maximum amount of money a decision-maker is willing to pay to ensure that the health benefits gained by patients using new treatments are not outweighed by health losses due to long-term cost pressures.

CTAF – California Technology Assessment Forum; CTAF represents one of ICER’s core appraisal committees that convene three times a year to review objective evidence reports and develop recommendations for how stakeholders can apply evidence to improve the quality and value of health care. The mission, processes, and role of CTAF are the same as the CEPAC programs, despite different naming conventions.

Direct comparison – An evaluation of two interventions that have been assessed head-to-head.

Dominance – In cost-effectiveness analysis, when one intervention is more effective and less costly than its comparator, the comparator is considered to be ‘dominated.’

evLYG analysis – An analysis that counts any gains in length of life equally, regardless of the treatment’s ability to improve patients’ quality of life. For all additional years of life gained, this analysis awards full health (i.e., the quality of life of the general population), irrespective of the health state patients are in during these additional years of life gained.

Forest plot – A graphical depiction of the results of a meta-analysis, which includes the data from each individual study as well as a pooled effect estimate.

Health-benefit price benchmarks– the treatment prices that would achieve incremental cost-effectiveness ratios of \$100,000 and \$150,000 per QALY or evLY gained.

Health technology assessment (HTA) – the systematic evaluation of evidence related to any healthcare intervention that can be used to improve health and prevent and treat disease; HTAs inform policy- and decision-making surrounding the use of such interventions.

Inconsistency – Disagreement between direct and indirect evidence.

Incremental cost-effectiveness ratio (ICER) – The ratio of the difference in costs between two possible interventions, divided by the differences in their effectiveness.

Indirect comparison – An evaluation of two interventions via one or more common comparators.

Meta-analysis – A type of statistical analysis that combines data from multiple studies assessing the same two interventions and generates a pooled, summary estimate of the relative effect of one treatment versus a comparator.

Net health benefit – the balance between benefits and risks and/or adverse effects.

Network meta-analysis – An extension of pairwise meta-analyses to include many interventions and generate a series of pooled, summary estimates of the relative effect of each treatment versus each comparator.

Observational study – a non-experimental study in which investigators draw inferences about what is observed without trying to influence the outcome of the study; types of observational studies include cohort, cross-sectional, case-control and ecological studies.

One-way sensitivity analysis – a method of analysis in which the value of one model input parameter is varied at a time to assess the effect of the parameter on results.

Opportunity cost – the value of something that must be foregone in order to acquire something else.

Parameter – a characteristic that influences the output of a model.

PICOTS – Population, intervention, comparator, outcomes, timing, and setting; ICER uses these items as a framework for defining the scope of its appraisals.

PRISMA – Preferred Reporting Items for Systematic Reviews and Meta-Analyses; PRISMA is a set of criteria that guide the conduct and reporting of systematic reviews and meta-analyses.

Probabilistic sensitivity analysis – a method of analysis used to account for parameter uncertainty in which values for input parameters are sampled based on pre-specified probability distributions.

Quality-adjusted life-year (QALY) – A measure of health benefit that accounts for changes in both quantity (e.g., mortality) and quality of life.

Randomized controlled trial – a type of study design in which participants are allocated at random into intervention and control groups.

Reference case – the framework of methods that ICER follows when conducting the base-case cost-effectiveness analysis.

Scenario analysis – A type of analysis that estimates results using alternative model assumptions.

Sensitivity analysis – a method of analysis in which model inputs are varied in order to determine how such changes affect the results.

Statistical heterogeneity – variation in the magnitude or direction of results between studies.

Subpopulation – A subset of a larger population.

Systematic review – a literature review that identifies and summarizes the results of all empirical studies meeting pre-defined eligibility criteria.

Threshold analysis – a type of sensitivity analysis in which the values of model input parameters are varied in order to determine the value that produces a specific result (e.g., a given cost-effectiveness value).

Time horizon – the period of time over which outcomes are evaluated.

Tornado diagram – a graphical depiction of the results of one-way sensitivity analyses in which the analyses with the greatest impact on model results are displayed with the largest bars and are stacked at the top of the chart.

Transitivity assumption – A key property in a network meta-analysis that preserves the ranking order relations of interventions with respect to a given outcome.

Utility – a measure of preference for a health outcome.

Validity – the assessment of whether a model has been implemented correctly (internal validity) and if its assumptions and results are in line with the current evidence and expectations (face validity).

Value assessment framework – a decision support tool intended to guide stakeholders in making decisions that will promote sustainable access to high-value care for all patients.

B. List of Major Revisions to 2017-2019 Framework

Comparative Clinical Effectiveness
Revise EBM Matrix categories (Separate C+ into two separate categories: C+ / C++)
Create new “Heterogeneity and Subgroups” subsection
Long-Term Cost-Effectiveness
Standardize \$50-\$200k/QALY and /evLYG threshold across all reports
Change value-based price benchmark terminology to “Health-benefit Price Benchmark”
Include shared savings scenario analysis (50/50 offset and cost offset cap model) when there are substantial potential cost offsets
For all reports, societal perspective analysis will be promoted to co-base case when the impact of care outside the health system is substantial in proportion to health effects
Create new “Heterogeneity and Subgroups” subsection for consolidated discussion of these considerations
Create new “Uncertainties and Controversies” subsection that will include discussion of model variations
Potential Other Benefits or Disadvantages and Contextual Considerations
Add category for model uncertainty and/or assumptions that may lead base-case model to be overly optimistic or pessimistic
Add categories for absolute and for proportional QALY shortfall
Add category for balance and timing of benefits/risks
Refine language for other categories
Change “potential other benefits or disadvantages and contextual considerations” votes to 1-3 Likert scale
Report Generation, Stakeholder Engagement, and Public Processes
Consolidate and enhance patient engagement efforts under Patient Engagement Program
Create separate “Patient Perspectives” chapter for each report, including sub-section for “Impact on Caregivers”
Long-term value for money votes will be taken in all circumstances (i.e., even if cost/QALY or /evLYG exceeds \$175,000)
Adopt formal code of conduct for public meetings
Implement process for 12-Month Report Check-Up
Implement pilot effort for 24-Month Real-World Evidence update for select drugs approved under accelerated pathways.

C. Review Timelines

Appendix Figure C1. Standard Review Timeline

ICER Process	Week	Milestones	Comments
Topic Announced	0	Topic Announcement	ICER begins scoping calls with clinical experts and patient groups. Stakeholders may submit information through the open input period.
		Open Input Period Begins	
Draft Scope	1		
	2		
	3	Open Input Period Ends Draft Scoping Document Posted	Stakeholders have 15 business days to comment on the draft scope.
Final Scope	4	Public Comment Period	ICER holds scoping calls with manufacturers and payers to discuss the draft scoping document
	5		
	6		
	7	Revised Scoping Document Posted ICER Sends Request for Data	ICER sends formal requests for data to each manufacturer. Supplemental data requests may be sent on an ad hoc basis.
Draft Evidence Report	8		
	9		
	10		
	11	Mfr. Evidence Submissions Due	
	12	Research Protocol Posting	Posting of evidence review protocol
	13		
	14		
	15	Preliminary Model Presentation Posting of Model Analysis Plan	Individual discussion calls with invited stakeholders 2-3 days after the preliminary model presentation. After reviewing ICER's preliminary model presentation, stakeholders may send supplemental data.
	16		
	17	Supplemental Data Submission Due	Supplemental data sent in response to ICER's preliminary model presentation are due 11 business days after call.
	18		
	19		
	20		
21	Draft Evidence Report Posted		
Evidence Report	22	Public Comment Period	Stakeholders have 20 business days to comment on the Draft Evidence Report. When possible, economic models are available for review by manufacturers.
	23		
	24		
	25		
	26		
	27		
Public Meeting	28	Evidence Report Posted	The relevant voting committee reads this version of the report.
	29		
	30	Public Meeting	
Final Report	31		
	32		
	33	Final Evidence Report Posted	

Legend:

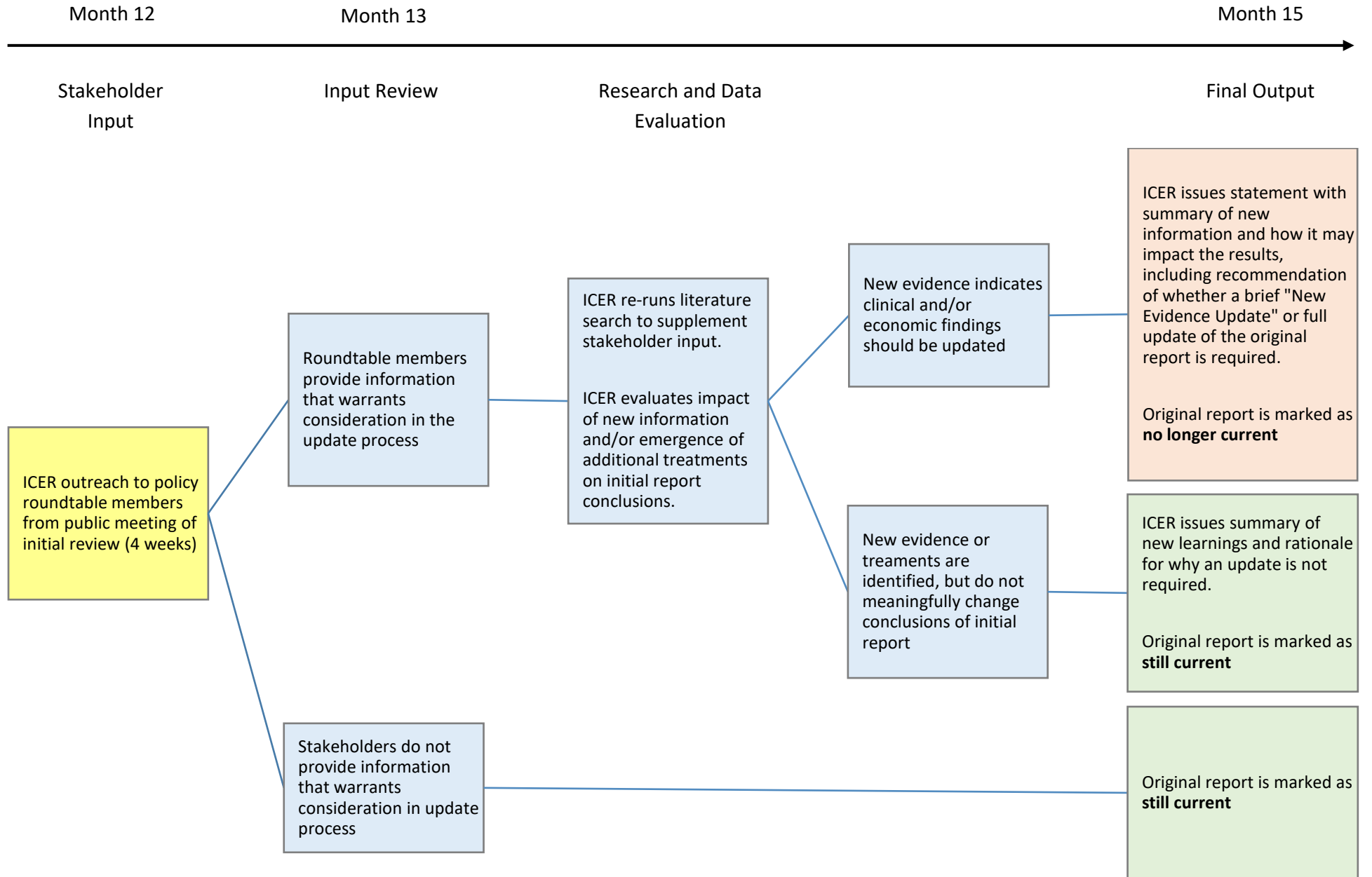
Document Release	Data Request	Input Opportunity
------------------	--------------	-------------------

Appendix Figure C2. Modified Timeline for Large Class Reviews

ICER Process	Week	Milestones	Class Review Adaptation
Topic Announced	0	Topic Announcement	
		Open Input Period Begins	
Draft Scope	1		
	2		
	3	Open Input Period Ends Draft Scoping Document Posted	
Final Scope	4	Public Comment Period	
	5		
	6		
	7		+1 week for additional scoping calls
Draft Evidence Report	8	Revised Scoping Document Posted ICER Sends Request for Data	
	9		
	10		
	11		
	12	Mfr. Evidence Submissions Due	
	13	Research Protocol Posting	
	14		+3 weeks to systematic literature review and model development timelines
	15		
	16		
	17		
	18		
	19	Preliminary Model Presentation Posting of Model Analysis Plan	
	20		
	21	Supplemental Data Submission Due	
	22		
	23		+1 week to address feedback on preliminary model
	24		
	25		
26		+1 week to facilitate revision of longer and more complex report	
27	Draft Evidence Report Posted		
Evidence Report	28	Public Comment Period	
	29		
	30		
	31		+1 week to public comment period to facilitate review of longer report
	32		
	33		
	34		
	35		+1 week to review a higher volume of stakeholder comments
Public Meeting	36	Evidence Report Posted	
	37		+1 week to allow voting committees sufficient time to review complex report
	38		
Final Report	39	Public Meeting	
	40		
	41		
	42	Final Evidence Report Posted	

Legend: Document Release Data Request Input Opportunity

Appendix Figure C3. 12-Month Report Check-Up Timeline



D. Cost-Effectiveness Threshold Ranges

Base-Case and Cost-Effectiveness Thresholds

ICER provides incremental results at \$50,000, \$100,000, \$150,000 and \$200,000 per QALY and per evLYG for all assessments, including those for treatments of ultra-rare disorders. While there are some recommendations to adopt differential cost-effectiveness thresholds for different types of treatments and/or different types of conditions, there continues to be no strong consensus among academic health economists or ethicists on whether or how to quantify and integrate these values into cost-effectiveness analyses, and we believe that it remains premature to seek to create a separate series of cost-effectiveness thresholds related to severity, burden of illness, or “need.”

In part, the challenge in this area is that while many people accept a broad ethical value to prioritize treatments for the worst off, arriving at a single quantifiable measure for this concept is difficult and raises thorny questions about whether the goal should be to prioritize the absolute loss of health (“absolute QALY shortfall”) or the loss of health in relation to the amount of time patients have left to live (“proportional QALY shortfall”). Either approach creates “winners and losers” among treatments that often causes equity concerns and other concerns about unintended consequences. This value framework brings greater clarity and empiric results to these issues as part of the deliberation and voting on “contextual considerations” performed as part of every public meeting of our independent appraisal committees.

ICER uses a common set of cost-effectiveness thresholds for all assessments, including those for treatments of ultra-rare disorders, providing a uniform range of results from \$50,000 to \$200,000 per QALY and per evLYG for all assessments, for several reasons. First, there remain important equity concerns related to extending the threshold range higher for treatments just because they treat a small population.³⁷ In addition, the economic landscape for treatments of rare and ultra-rare conditions has shifted. Years ago, when drug prices were far lower on average, it could be reasonably argued that the profit required to sustain innovation in rare disease treatments required pricing that far exceeded standard cost-effectiveness thresholds. But in today’s market environment, it only takes \$100,000 per treatment course, multiplied by a mere 10,000 patients, to provide \$1 billion per year in revenue. We therefore judge that today it no longer seems necessary to make important exceptions to applying standard cost-effectiveness thresholds to analyzing the value of treatments of rare or ultra-rare conditions.

Our view of treatments for ultra-rare conditions includes the historical perspective that decision-makers have often accepted prices beyond standard cost-effectiveness ranges, particularly for treatments of very small ultra-rare populations. We will continue to include standard language to this effect when presenting health-benefit price benchmarks for these treatments. As our range for health-benefit price benchmarks remains \$100,000-\$150,000 per QALY and evLYG, we will provide a broader range of results symmetrically around this range, from \$50,000-\$200,000 per QALY/evLYG.

We believe this is a broad enough range to accommodate the needs of decision-makers in the US to think about their own desired interpretation of cost-effectiveness thresholds.

Although ICER uses a standardized threshold range across all assessments, our reports continue to include discussion of contextual factors and other important considerations for all therapies, including those for ultra-rare disease or short-term transformative treatments. We also acknowledge that, no matter the threshold or range selected, ICER and the broader HTA community have a responsibility to educate potential users of our work about the need to embed CEA in a broader decision-making structure that is sensitive to the benefits and disadvantages of treatments that do not feature in the outcomes of clinical trials, as well as the ethical dimensions that are always inherent in any priority-setting process.

Health-benefit Price Benchmarks

ICER uses the range of \$100,000-\$150,000 per QALY and per evLYG in presenting health-benefit price benchmarks. ICER continues to use the threshold range from \$100,000-\$150,000 per QALY as the standard for its health-benefit price benchmarks for all assessments, but health-benefit price benchmarks using \$100,000-\$150,000 per evLYG will also be provided.

ICER recognizes the variety of academic and conceptual work over the years that has explored methods for establishing cost-effectiveness thresholds.³⁸ There are two basic theoretical approaches to determining cost-effectiveness thresholds: 1) demand-side, or willingness to pay (WTP), and 2) supply-side, or opportunity cost.

Ryen and Svensson reviewed the literature on WTP for a QALY and found that results from studies based in the US differed by orders of magnitude, with the most recent (2010) estimate at approximately \$60,000 per QALY.³⁹ Demand-side approaches have often focused on measures of per capita GDP, surveys of individual WTP, or revealed choices (e.g., estimates from job choices). Benchmarks for cost-effectiveness thresholds have been frequently justified by estimates of societal WTP, which, based on earlier consensus efforts at the World Health Organization (WHO), have commonly been cited as approximately 1-3 times the per capita GDP of the country per additional QALY.^{40,41} However, Marseille et al. point out that thresholds based on per capita GDP have little theoretical basis, are too high to distinguish among most interventions, and are not likely to reflect affordability in many settings.⁴² WHO itself has recently commented on the “misuse” of its earlier recommendations, and has argued that thresholds in this range are likely to prove unaffordable over the long-term.⁴³

Attempts have also been made to use the value of statistical life (VSL) as a measure of societal WTP, especially in transportation and environmental assessments.⁴⁴ VSL estimates are based on evidence from market decisions such as wages for jobs with different risks of death, or on surveys that ask about similar risk-money tradeoffs. However, there are several important limitations of this

approach.⁴⁴ Using VSL estimates in this way conflates WTP to avoid risk and willingness to accept risk, which may be quite different. In addition, using VSL as an estimate of WTP requires the assumption that VSL can be converted to calculate the value of a life year, but how to “spread” the VSL over life years remains unresolved. Using data on job choice to determine WTP also requires several strong assumptions about the fairness and rationality of the labor market, such as that workers have free choice of employment across jobs with different levels of risk. Lastly, the literature finds a wide range of estimates for VSL across different studies, with Hirth et al. reporting upper-bound estimates that were greater than 20 times the lowest estimate.⁴⁵

Another suggestion as a basis for setting cost-effectiveness thresholds in the US has been to use prior funding decisions to benchmark WTP for future interventions. However, there is no certainty that previous funding choices were made with cost-effectiveness in mind. In addition, estimates of demand based on current funding may be distorted because health insurance is a tax-credited employment benefit, meaning that health insurance coverage decisions do not necessarily match population preferences.

In an important recent conceptual contribution, Phelps⁴⁶ built on earlier work he had done with Garber⁴⁷ to look at how the optimal (i.e., utility-maximizing) threshold would vary with income and relative risk aversion. In this recent work, Phelps estimated optimal WTP by specifying utility as a function of income and using estimates of relative risk aversion – a measure of the rate at which marginal utility changes as income changes – to calibrate the function. This analysis assumed a Weibull utility function, which was parameterized to have declining absolute risk aversion (DARA) because the quantity of risky assets rises with wealth, and increasing relative risk aversion (IRRA) because the share of risky assets declines with wealth, as observed by Arrow.⁴⁸

Results from these analyses confirm previous work suggesting that the optimal WTP threshold rises with income, as does the ratio of the threshold to income. That is, as income rises, trading off other goods and services for health care becomes less painful in terms of loss of utility and spending on health care should increase. Assuming an income of \$50,000 and plausible values for other parameters, Phelps found that the optimal threshold was approximately two times income, or approximately \$100,000-\$110,000 per QALY if using the mean personal income in the US (approximately \$54,000 in 2018).⁴⁹ Phelps notes that this work focuses on a representative, utility-maximizing individual, and expansion from this to decisions at the societal level may not be straightforward.

Phelps’ approach to estimating WTP represents an important contribution, but WTP may be considered a more relevant approach to thinking about thresholds in a consumer sovereignty-based (i.e., welfarist) system. Value may vary by individual income and over time, and it is not clear whether WTP should be measured at the individual or household level. In addition, all WTP methods need to account for the mix of those who can afford to pay something and those who cannot, as a “median voter rule” for this mixed population would give a different answer than

among those who can afford to pay some amount. Phelps has pointed out that a skewed income distribution means that the median voter model would almost always lead to lower thresholds than would be utility maximizing.⁴⁶ A central question in considering health economics is who captures the “value” of an intervention. Using a central measure of WTP, such as the median WTP, could lead to reduced access for those who have lower ability to pay. If an “average” WTP is selected, people with lower incomes may be forced to pay too much for health care to satisfy the WTP of the rich. Societal resources may be drawn into health spending from other domains of social spending that are much more important to people with lower incomes (such as public education). Some people with lower incomes are likely to be forced out of insurance markets all together.

In the US market-based system with multiple payers, there is a case for multiple thresholds based on WTP which may differ by payer type (e.g., government vs. commercial insurance).²⁵ However, there are broad requirements across the US health care system to fund all “medically necessary” care. We also believe that there exists a widely accepted ethical goal in the US to have a common standard of care available for all patients, albeit with acknowledged differences in access due to network constraints, out-of-pocket payment, and other benefit design features. That the US does not yet achieve the goal of a common standard of care available for all patients does not imply, in our view, that ICER should abstain from framing a range of cost-effectiveness that should apply broadly across many, if not all, health insurance systems in the US.

Turning from the WTP approach, the other major paradigm for determining cost-effectiveness thresholds is a supply-side approach based on the idea that thresholds should reflect the opportunity cost of additional health care spending. Opportunity cost approaches based on health care system outcomes and costs look at the trade-off between spending on a new intervention when that spending must come from curtailing current spending elsewhere in the health care system on existing interventions, or from reducing spending on other social goods outside the health care system, such as education or public safety. This approach has its strongest theoretical foundation in situations where the health care system budget can be considered fixed. In such cases, the threshold can be considered as reflecting the point at which a higher price for a new intervention will lead to more health being lost within the health care system than will be gained by the patients who will benefit from the new treatment.

The best recent evidence on opportunity cost suggests that the previous WHO-recommended ranges for cost-effectiveness of one to three times per capita GDP are too high.²¹ Claxton has argued for a lower cost-effectiveness threshold in the UK, US, and other countries, given the marginal productivity of the respective health care systems.²¹⁻²³ For the US, Claxton estimates an opportunity cost threshold of approximately \$30,000-\$40,000 per QALY.²³

More recently, there has been a seminal attempt to ground an opportunity cost analysis directly from US data. In this work, Vanness has estimated health opportunity costs for private plans in the US.²⁴ Taking account of the effect of premium changes on coverage and the morbidity and

mortality effects of loss of coverage, Vanness estimated the negative QALY impacts that result in the US health care system with rising costs and premiums. His research produces an estimate of \$84,000 per QALY as the threshold. Working within this paradigm, this means that any new intervention introduced at a price that leads to an incremental cost-effectiveness ratio greater than \$84,000 per QALY produces a net loss of health due to its impact on premium increases and thereby loss of insurance, especially among poorer members of the insurance pool. Vanness's work does not capture the potential impact of rising premiums on increasing deductibles and other out-of-pocket requirements that can lead to delayed or foregone care, nor does it capture the impact that rising premiums have on suppressing spending on other workplace benefits and wages. In some ways, therefore, it could be considered an upper-bound estimate of a threshold at which greater net losses occur despite the introduction of a treatment that will benefit those patients who can obtain it.

ICER uses the opportunity cost approach as the major theoretical foundation in its determination of the cost-effectiveness thresholds for health-benefit price benchmarks to inform decision-making, for several reasons. Despite the lack of an explicit overall budget for health care in the US, we believe the current environment of the US health care system indicates that we have reached a point where policymakers are no longer willing to accept cost increases in the US health care system that outpace growth in the overall economy. We hear this repeatedly from employers, unions, and other plan sponsors who are trying to maintain health benefits for their members. We hear this in broader concerns from consumer groups such as FamiliesUSA and AARP, who are aware of the opportunity costs faced by the public due to increasing health care costs. We hear it repeatedly from representatives of state government and state Medicaid programs, where rising health care costs have stripped out state spending on other needs such as education, police, and public infrastructure. And we also view the goals of several state laws as indicative. Maryland has a long-standing arrangement that limits hospital cost growth to the growth rate estimated for the state's overall economy.⁵⁰ Massachusetts already links policy actions to growth in health care costs that outstrip growth in the state per capita GDP; and recent initiatives may extend state oversight to prescription drugs as well.⁵¹

Overall, therefore, we believe that ICER functions in a system where health expenditure may continue to grow, but that it has reached the point at which policymakers sense that the opportunity cost for current spending is already substantial. This implies that an opportunity cost paradigm is justifiable as the predominant theoretical foundation for cost-effectiveness thresholds. We believe that the opportunity costs are real, both within the health care system and beyond, and that our goal should be to recommend prices that will ensure that new interventions are adopted at a price that leads to a net increase in health over the entire population. It is not a matter of saving money; it is a commitment to improving health.

Following this line of reasoning and reflecting on the most recent conceptual and empirical research, reducing the health-benefit price benchmark range to \$50,000-\$100,000 per QALY could

be contemplated. We note, however, that the top end of our price benchmark range is usually interpreted as a “ceiling” price beyond which a treatment will be viewed as not cost-effective. We are aware that the opportunity cost empirical data for the US need formal peer review and further delineation. It is reassuring that the most recent highly respected work using the WTP paradigm for determining thresholds arrived at a very similar approximate result: \$100,000 per QALY. And we believe there is some value in ICER retaining a consistent threshold range as a level playing field for all stakeholders. We therefore retain our current cost-effectiveness range to support health-benefit price benchmark recommendations. We recognize that single cost-effectiveness thresholds should not be used as a blunt decision rule, and that decision-makers may want to consider different thresholds given their own view of their opportunity costs and their interpretation of a treatment’s potential other benefits and contextual considerations.