# ICER 2019 Perspectives on Cost-Effectiveness Threshold Ranges

ICER conducted a seminar series in 2019 to share perspectives and inform thinking among academics and Health Technology Assessment (HTA) programs on approaches to determining operative cost-effectiveness thresholds for decision-making in health care. Health economists from around the world participated, with presentations from leading figures including Karl Claxton, PhD, MSc, from the University of York; Jens Grueger, PhD, from F. Hoffmann-La Roche & University of Washington; Sean Sullivan, PhD, MS, from the University of Washington; and Chris McCabe, PhD, MSc, from the University of Alberta. The perspectives below represent ICER's perspectives on this issue in 2019 and have not been updated to reflect subsequent academic work or policy development.

***Base-Case and Cost-Effectiveness Thresholds***

ICER provides incremental results at $50,000, $100,000, $150,000 and $200,000 per quality-adjusted life year (QALY) and per equal value of life years gained (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 or the loss of health in relation to the amount of time patients have left to live. 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.[1] 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 cost-effectiveness analysis (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.[2] 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.[3] 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.[4,5] 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.[6] 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.[7]

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.[8] 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.[8] 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.[9]

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, Phelps[10] built on earlier work he had done with Garber[11] 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.[12]

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).[13] 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.[10] 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).[14] 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.[15] 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.[15-17] For the US, Claxton estimates an opportunity cost threshold of approximately $30,000-$40,000 per QALY.[17]

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.[18 19] 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.[20] 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.[21]

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.

# References


1. 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 economics.* 2013;22(8):948-964.
2. 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.
3. Ryen L, Svensson M. The Willingness to Pay for a Quality Adjusted Life Year: A Review of the Empirical Literature. *Health economics.* 2015;24(10):1289-1301.
4. 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.
5. 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.
6. Marseille E, Larson B, Kazi DS, Kahn JG, Rosen S. Thresholds for the cost-effectiveness of interventions: alternative approaches. *Bulletin of the World Health Organization.* 2015;93(2):118-124.
7. Bertram MY, Lauer JA, De Joncheere K, et al. Cost-effectiveness thresholds: pros and cons. *Bulletin of the World Health Organization.* 2016;94(12):925-930.
8. 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.
9. 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.
10. Phelps CE. A New Method to Determine the Optimal Willingness to Pay in Cost-Effectiveness Analysis. *Value Health.* 2019;22(7):785-791.
11. Garber AM, Phelps CE. Economic foundations of cost-effectiveness analysis. *Journal of health economics.* 1997;16(1):1-31.
12. Arrow K. *Essays in the Theory of Risk-Bearing.* 2016/08/17 ed. Amsterdam: North-Holland Publishing Company; 1971.
13. US Bureau of Economic Analysis. Personal Income Per Capita [A792RC0A052NBEA]. In. Vol 2019: Federal Reserve Bank of St. Louis; 2018.
14. Pauly MV. The Questionable Economic Case for Value-Based Drug Pricing in Market Health Systems. *Value Health.* 2017;20(2):278-282.
15. 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.
16. 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.
17. 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 economics.* 2018;27(6):1017-1023.

18. 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.
19. Vanness DJ, Lomas J, Ahn H. A health opportunity cost threshold for cost-effectiveness analysis in the United States. *Annals of internal medicine.* 2021;174(1):25-32.
20. Maryland Health Services Cost Review Commission. Maryland's Total Cost of Care Model: Update for Providers. https://hscrc.state.md.us/Pages/tcocmodel.aspx. Accessed August 20, 2019.
21. 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.