



Published in final edited form as:

*Appl Health Econ Health Policy*. 2021 March ; 19(2): 155–162. doi:10.1007/s40258-020-00606-4.

## Is the Choice of Cost-effectiveness Threshold in Cost-utility Analysis Endogenous to the Resulting Value of Technology? A Systematic Review

William V. Padula, PhD<sup>1,2</sup>, Hui-Han Chen, MHS<sup>3</sup>, Charles E. Phelps, PhD<sup>4</sup>

<sup>1</sup>Department of Pharmaceutical & Health Economics, School of Pharmacy, University of Southern California, Los Angeles, CA

<sup>2</sup>Leonard D. Schaeffer Center for Health Policy & Economics, University of Southern California, Los Angeles, CA

<sup>3</sup>Division of Practice Advancement and Clinical Education, UNC Eshelman School of Pharmacy, University of North Carolina, Chapel Hill, NC

<sup>4</sup>Departments of Economics and Public Health Sciences, University of Rochester, Rochester, NY

### Abstract

**Background:** Cost-utility analysis (CUA) is widely used for health technology assessment; however, concerns exist that CUA analysts may suggest higher cost-effectiveness thresholds (CET) to compensate for technologies of relatively lower value.

**Objectives:** We explored whether selection of a CUA study's CET was endogenous to estimated incremental cost-effectiveness ratios (ICERs).

**Methods:** We systematically reviewed the U.S. cost-effectiveness literature between 2000 and 2017 where studies with explicit CET and ICERs were included. We classified the ratio of studies hypothesized to analyze cost-effective technologies at low CET (i.e. less than \$100,000/QALY) vs. higher CET (i.e. \$100,000 to \$150,000 /QALY) relative to their ICER, using a Chi-square test to examine whether technologies that were cost-effective at high CET would still be cost-effective at lower thresholds. We also performed fixed-effects linear regression exploring the associations between ICERs and reported CETs over time.

**Results:** Among 317 ICERs reviewed: (A) 185 had an ICER <\$50,000/QALY;(B) 53 had \$50,000 ICER, < \$100,000; (C) 20 had \$100,000 ICER <\$150,000; and (D) 59 had an ICER \$150,000. Chi-square testing showed a strong association ( $p < 0.001$ ) between estimated ICER values and chosen CET, illustrating a lack of independence between the two. The regression analysis indicated that CETs have a baseline value of \$52,000 and grow by \$0.37 for each dollar increase in the estimated ICER.

---

**Correspondence:** Dr. William Padula, USC Schaeffer Center, 635 Downey Way (VPD), Los Angeles, CA 90089; padula@usc.edu; +1-203-506-0161.

**Author Contributions:** All authors contributed equally to this manuscript in terms of conceptualization, data analysis and finalization of the manuscript in accordance with ICMJE guidelines.

**Financial Disclosures:** William Padula (Monument Analytics, consulting; Molnlycke Healthcare, scientific advisory board). Charles Phelps (Merck, consulting; Pfizer, consulting; Audentes Therapeutics, consulting).

**Conclusions:** Cost-effectiveness thresholds represent the hypothesis tests of typical CUAs. Our analysis highlights that most CUAs that cite high CETs also result in greater ICERs for the novel interventions that they investigate; thus, these interventions would otherwise not have been cost-effective at lower CETs. Selection of a CET may come after the ICER is calculated to infer value that suits a hypothesis.

### Keywords

Cost-utility; cost-benefit; cost-effectiveness; willingness-to-pay threshold; cost-effectiveness threshold; incremental cost-effectiveness ratio (ICER)

---

## 1. INTRODUCTION

Over the past 20 years since the first U.S. Panel on Cost-effectiveness in Health and Medicine, healthcare expenditures and the prices of medical services and biotechnologies have risen dramatically [1]. It has become important for researchers to provide objective information on the cost-effectiveness of different technologies so that low-value alternatives can be phased-out in favor of higher value biotechnologies and healthcare services [2]. Cost-effectiveness analysis (CEA) is a tool with the advantage to combine different aspects of value for a new healthcare services or biotechnology into a single composite measurement. By assessing the incremental costs and benefits (e.g., effectiveness, utility), decision-makers are able to quantify what additional resources they have to pay to gain the additional benefits.

Most western cost-utility analyses (CUA) use quality-adjusted life-years (QALYs) as a standardized measure of effectiveness as part of the key resulting statistic, an incremental cost-effectiveness ratio (ICER). The U.S. Panel recommends testing the hypothesis for a new technology's cost-effectiveness by comparing the ICER to a pre-determined cost-effectiveness threshold (CET); technology with ICERs that fall below this threshold would be considered cost-effective to the relevant perspective.[3] The CET is based on a societal "willingness-to-pay" for health technologies and healthcare services based on supply-side and demand-side economic approaches.[4] On the demand-side, CET represents value for money at a given price. On the supply-side, technologies with an ICER below a CET represent investments that do not draw resources away from better alternatives (i.e. "opportunity costs") [5]. However, there is little methodological governance with respect to the CET that should be chosen for CUA on a specific set of interventions [6]. The first and second U.S. Panels on Cost-effectiveness in Health and Medicine, for example, have not recommended a universal CET for any perspective [7, 8].

CETs reflect the maximum societal willingness-to-pay for an additional QALY while balancing the needs of multiple priorities across the public sector [9]. A commonly referenced CET, \$50,000 per QALY, first appeared in empirical studies in 1992 and its point of reference was widespread by the mid-1990s.[10] This CET might come from an arbitrary settlement for the range of \$20,000 to \$100,000 per QALY [8, 11]. Yet, it is still controversial whether this number reflects the true willingness-to-pay of the U.S. society. For example, taking into account a variety of factors might be necessary to adjust the

benchmark, where such factors include inflation rates, economic growth, and the different CETs across different diseases [11].

Previous studies argue that a higher CET should be adopted and different numbers have been assessed, including \$100,000/QALY and \$150,000/QALY based on the estimated one- to three-times Gross Domestic Product (GDP) per capita [11–13]. In particular, a 2014 article by Neumann and colleagues exploring the rise of CETs could have impacted downstream selection of CETs.[11] However, the range may be considered inconsistent and some studies proposed that the upper bound of the CET should be no more than one-times GDP per capita based on the principle that opportunity costs will increase, thereby withdrawing financial resources away from the needs of other patient cohorts or public health programming [5, 6]. According to the U.S. average income per capita of \$59,531 in 2017, \$50,000 to \$150,000 per QALY would be considered appropriate, although more than half of the US population actually make less than this amount individually or as a family, and as a product of national politics are not part of a single-payer reflecting pooled risk for this calculated range of CETs [14]. In a new method to calculate an optimal willingness-to-pay, Phelps estimated that an optimal CET would be about two-times income in the range of \$50,000, providing a best-estimate limit for the CET around \$100,000 per QALY if compared with per capita GDP in the US and many developed nations around the world.[15]

Since U.S. panel guidelines do not recommend a single CET benchmark for the CUA, the choice of CET depends on the researchers' discretion.[8] Consequently, while the scientific method prescribes that researchers hypothesize a CET prior to conducting an analysis, it remains possible that the choice of a CET is endogenous to the ICER that is anticipated or found for a given ICER. That said, researchers could create cost-effective results by 'Cherry-Picking a higher CET intentionally, which introduces potential bias on the interpretation of study outcomes. This study explores whether the choice of the CET was independent of the estimated ICER by reviewing recent published articles.

## 2. METHODS

### Study Design

We performed a systematic review of CUAs taking on U.S. societal perspectives following the PRISMA guidelines. We were assisted by a medical librarian at the Johns Hopkins Welch Medical Library (Baltimore, USA) to conduct a review of CUAs in Medline using the OVID search engine, as could be done to a high degree of fidelity with likely findings of a similar search through the Tufts CEA Registry or other formal economic study registries. The particular reason for restricting the review of CUAs using Medline was to ensure that they were peer-reviewed studies in publicly available literature with the OVID search engine to ensure consistency in the search results (Table 1). Formal, comprehensive search terms, including "cost-effectiveness," "cost-utility," "QALY," "quality-adjusted life-years," and "United States" were developed with the assistance of the medical librarian. Candidate studies included in the review had to be peer-reviewed and published in the range of January 2000 through February 2017, which was potentially an advantage over other registries which have a 1-2 year lag, if not more. The identified abstracts were reviewed and assessed by two independent reviewers (WVP and HHC), where only the original articles with published

full-text manuscripts were included and reports only encompassing systematic reviews and abstracts were excluded. The outcomes of interest, including ICERs and their referenced CET(s) of the articles, were extracted for further analysis (see Supplementary Material, Table e1).

As published ICERs and referenced CETs were analyzed, articles reporting either “dominant” or “dominated” strategies (i.e. where a negative ICER was calculated) were excluded because choosing a different WTP threshold would not influence the final interpretation. In the final Chi-square analysis, we also excluded papers that either reported ambiguous CETs or did not report referenced CETs.

### Statistical Approach

The extracted CETs and ICERs were categorized into two levels: (1) less than \$100,000/QALY; (2) between \$100,000/QALY and \$150,000/QALY. A Chi-squared test was performed to evaluate the independence between the reported ICERs and the selected CETs at the 95% confidence-level.

We also analyzed the associations between *CET* and the key explanatory variable, *ICER* over time using ordinary least squares (OLS) regression with fixed-effects, as well as linear regression with random-effects. These regression models were constructed by clustering studies by *year* of publication (Equation 1).[16] We also generated covariates to control for other potential study design factors that could have categorically impacted the results relative to the CET, including: time-*ICER* interactions; intervention type (e.g. drug, medical device or socio-behavioral intervention), comparator (i.e. placebo, other drug, or “other” comparator); and a dummy variable was created to explore the impact of the paper by Neumann and colleagues on shifts to higher ranges of CETs after 2014.[11] A fixed-effects form of the model (*i*) was tested in addition to a random-intercept ( $u_{0j}$ ) model clustering studies by year (*j*). Using stepwise regression, we removed covariates that did not offer strong associations, other than *ICER*.

$$E[y_{i,j}(CET)] = (\beta_0 + u_{0i}) + \beta_1 * ICER_{ij} + \beta_2 * year_{ij} + \beta_3 * ICER_{ij} * year_{ij} + \dots + \varepsilon_{ij}$$

(Equation 1)

## 3. RESULTS

We screened 2,125 articles and abstracts based on the inclusion and exclusion criteria, where 317 ICER results were extracted along with the referenced CETs (Figure 1). While more than half of the published ICERs (n=185) were less than \$50,000 per QALY, 58.4%, 75.1%, and 81.4% of the ICERs would be cost-effective by assuming the CETs to \$50,000/QALY, \$100,000/QALY, or \$150,000/QALY, respectively (Table 2).

While more than two-thirds of the ICERs were compared to a single CET benchmark, 13.6% (n=43) ICERs were compared to multiple CETs. Among 11.4% (n=36) of studies, the ICERs were not compared to traditional CET benchmarks; such studies were not included in the following Chi-squared analysis. In addition, 44 studies reported ICERs above \$150,000

per QALY which were also excluded since they would not have been cost-effective at any CET, regardless of the hypothesis being tested in the economic model.

We examined the association between the choices of CET and the ICERs. There were 194 studies with CETs and ICER results less than \$150,000 per QALY and referenced single CETs were included in the Chi-squared analysis. A 2-by-2 table, whose columns equal the referenced CETs (less than \$100,000 per QALY, and between \$100,000 and \$150,000 per QALY) and rows equal the ICER results (less than \$100,000 per QALY, and between \$100,000 and \$150,000 per QALY) was constructed (Table 3). From the Chi-squared test, and we found statistically significant results, implying that the ICER was more than likely to fit within a high CET ( $p < 0.001$ ). That is, a significant proportion of published papers that were cost-effective based on a CET between \$100,000-150,000 per QALY would not have been cost-effective if the authors had analyzed their results at a CET less than \$100,000 per QALY.

Of the ten studies reporting these higher ranges of CETs, one was published in 2002, and nine were published between 2015-2016. While the sample size is limited, these observations reflect a broad range of time over which studies have referenced CETs above \$100,000 per QALY.

Furthermore, fixed-effects linear regression identified statistical associations between ICER values and the reported CET. First, we explored the direct association between the ICER and CET in *Model 1* (Table 4). We observed that CETs had a baseline value of about \$52,203, and grew by \$0.37 for each dollar increase in the estimated ICER. Second, we controlled for the additional impact of the publication of the 2014 article by Neumann and colleagues, the timing of which was associated with an \$18,542 increase in the baseline CET, and grew by \$0.31 for each dollar increase in the estimated ICER for studies published after 2014. The results in *Model 2* indicate a potential overall shift in the use of greater CETs than previously.[11] Models controlling for *year*, time-interactions, or with random-effects did not offer improved fit or significant associations between time and CET as the fixed-effects model did. Other covariates such as comparators and intervention types did not provide statistically significant improvements to the models.

#### 4. DISCUSSION

Based on findings from our systematic review of CETs and statistical analysis, we cannot rule out that such thresholds are selected *a posteriori* to analysis. There is a significant association between CETs and positive results when interpreting the findings of CUAs with large ICERs above \$100,000 per QALY. This finding was disproportionate to the fact that most studies in the overall review had ICERs that fell below \$100,000 per QALY and accordingly reported CETs less than or equal to \$100,000 per QALY as well. Thus, higher ICERs are associated with a shift in the frequent use of higher CETs. This shift is particularly noticeable after 2014 when published commentary on cost-effectiveness methodology may have inadvertently guided economic modeling to replace lower CETs with values between \$100,000 to \$150,000 per QALY to complement health technology that came at greater opportunity costs.[11, 15]

These associations could be the product of several potential mechanisms. First, it is possible that low CETs are no longer realistic from the U.S. societal perspective given the medical price inflation in the past three decades since introducing the \$50,000/QALY threshold. Second, that ICERs are selected once the CUA is conducted, such that results are fitted to a CET that returns a positive result. For example, if inappropriate comparator is chosen, the ICER results may be misleading even if the the ICERs are less than the CET. Third, that negative results of ICERs above maximal thresholds (e.g. >\$100,000/QALY) are not being reported frequently. Fourth, that financial conflicts of interest could bias the threshold to fit the results of the study, as Bell and colleagues have previously identified.[17] While it is uncertain whether the CET was predetermined arbitrarily or retrospectively, the field's lack of direction from a governing body about a single CET has allowed authors to draw conclusions of their own work against the most convenient hypothesis test to fit their results.

While the use of CUA is limited in the US, CUA is widely used to determine the value of new technology in recent decades in order to facilitate decision-making processes in other countries, and may yet find a collective purpose in the U.S. through organizations such as the Institute for Clinical and Economic Review for purposes of value assessment. ICERs, the results of CUA, should then be compared to predetermined CETs. The Institute for Clinical and Economic Review has recently published a statement in its 2020 Value Assessment Framework indicating that “prices need to achieve between \$100,000 per additional QALY and \$150,000 per additional QALY...as the bookends of our ‘health-benefit price benchmark’.”[18] This position by the Institute could provide additional momentum for economic modelers to move away from the \$50,000 per QALY threshold.

Studies have referenced \$50,000 per QALY as the CET since the 1990s [10, 11]. The first publication to use a CET of \$50,000/QALY appeared in the literature in 1992, and the second paper did not appear until 1995 when this threshold garnered more frequent use [10, 19, 20]. However, because of medical price inflation, the continued use of this standard for CEA has been controversial [11]. Garber and Phelps explored CETs using “common practices” as a reference point and found a great deal of variability to validate \$50,000/QALY or recommend a different threshold [21]. Other societies, such as the U.K., have explored the use of a benchmark CET for most CEAs, and an increased CET for therapeutic areas with emergent needs or unusually costly technologies [22]. Some experts have argued that studies should adopt a more reasonable benchmark for CEA, leading investigators to explore the CET of the U.S. using more analytical approaches than simply deriving from national GDP [23]. Different sectors and researchers hence suggest a wide range of CET benchmarks. For example, using one- to three-times the annual per capita income as the standard has been proposed and endorsed by the World Health Organization, which has no theoretical basis [6]. Recent health economics studies suggest a lower CETs for CEA that reflect the opportunity costs [5, 6]. Whereas, others have recommended using an even higher CET ranging from \$200,000 to \$300,000 per QALY based on economic and healthcare expenditure growth [12, 24, 25]. Meanwhile, others have also indicated that a moderate standard, such as \$100,000 per QALY, should remain the target point for CUA [11].

Nonetheless, there is still no consensus about which threshold can truly represent the US societal CET [26]. Our review shows that the selection of a CET may be endogenous



to the resulting ICER of a CUA. In other words, to support the cost-effectiveness of the intervention, researchers tended to pick an appropriate CET to compensate for the lower value of the intervention. The interpretation of the study result is hence under the threat of distortion. To prevent this bias, we recommend that using a range of CETs will be more appropriate than only referencing a single CET, such that a CEA would note that while an intervention with an ICER of \$125,000/QALY is not cost-effective at traditional CETs below \$100,000/QALY, it would be of value to patients and payers willing to pay up to \$150,000/QALY.

Having multiple CETs would likely be preferable given the circumstances of the U.S. healthcare system's market heterogeneity of multiple payers as well as wide-ranging variability in patients' affordability and access to health services. Our results align with previous studies where no single threshold is appropriate for all decision-makers and multiple thresholds are preferable based on different availabilities of resources for policy-makers [11, 26]. Multiple CETs would promote competition among payers who wish to attract beneficiaries to their plans based on payers' interest in investing in beneficiary welfare through access to more high-cost services that still provide clinically meaningful benefit [27].

Our suggestion also aligns with the recommendations of previous studies and the Second Panel on Cost-Effectiveness in Health and Medicine [8, 11]. Since CETs should be based on the available budgets and the opportunity costs of decision makers, reporting multiple thresholds is desirable: decision makers can choose their own CET based on their budget [11]. While using multiple thresholds is desirable to fit the need of different stakeholders, using single benchmark is still beneficial to evaluate the comparative value for different healthcare technologies [11]. It would be preferable if the U.S. Panel also took an explicit stance on a recommended CET performed for specific perspectives (e.g. the U.S. societal perspective). Finally, further studies are needed to explore societal cost-effectiveness thresholds over different time periods.

### Limitations

Our study had several limitations. First, our study only explores associations between ICERs and reported CETs, rather than analyzing data that can be used to infer a direct causal relationship between the two. Second, the search criteria came from a single database, Medline. Third, it lacked adjustment for some potential confounders that could have been observed over time between published studies, such as availability of research funding from foundations, government or industry that could have curbed publication bias. Fourth, the analysis of *time* fixed-effects was limited since we could only depend on the publication year for reliable point of reference, and studies inconsistently reported a year of evaluation.

Lastly, the study was limited to studies published through 2017 which we could view in Medline using the OVID search engine. This is limiting in the current environment for two reasons: (A) the interruption identified in CETs after 2014 indicates that there may have been a shift in more previous years of data analysis; (B) there may be other CUAs not appearing on Medline that appear in other registries such as the Tufts CEA registry. However, these registries report lags in the availability of posted studies by 1-2 years or

more, making Medline a more current search engine. As such, we recommend that future research update information on this subject area, as well as explore the degree of congruence in literature searches between Medline and other registries of published economic studies.

## 5. CONCLUSION

In conclusion, with no formal guidance about the choice of CET, authors have considerable leeway in how they frame their conclusions for CEA and how society interprets the value of new health technology. We base this assertion on our investigation as to whether the hypothesized CET was independent of the resulting ICER for a CUA. Among 194 eligible ICER results, 94.8% of the referenced CETs reported were less than \$100,000 per QALY. In our primary model, CETs began \$52,000 per QALY and grew by \$0.37 per dollar increase in the ICER. Thus, while our study only explores associations between CETs and ICERs rather than a direct causal relationship, it can be concluded that the choices of the CETs may not be independent of the ICERs. Technology referencing CETs above \$100,000/QALY should be viewed with a critical eye towards the population that it could be used by, and whether the majority of payers could readily afford the technology without perpetuating opportunity costs that disenfranchise certain patient groups.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Funding:

William Padula is supported by a grant from the National Institutes of Health Office of Extramural Research (KL2 TR001854).

## Data Availability Statement

Our data are openly available in published literature through Medline. We have organized these data for public access in supplementary material provided (Table e1), which aggregate data used for analysis in this study.

## REFERENCES

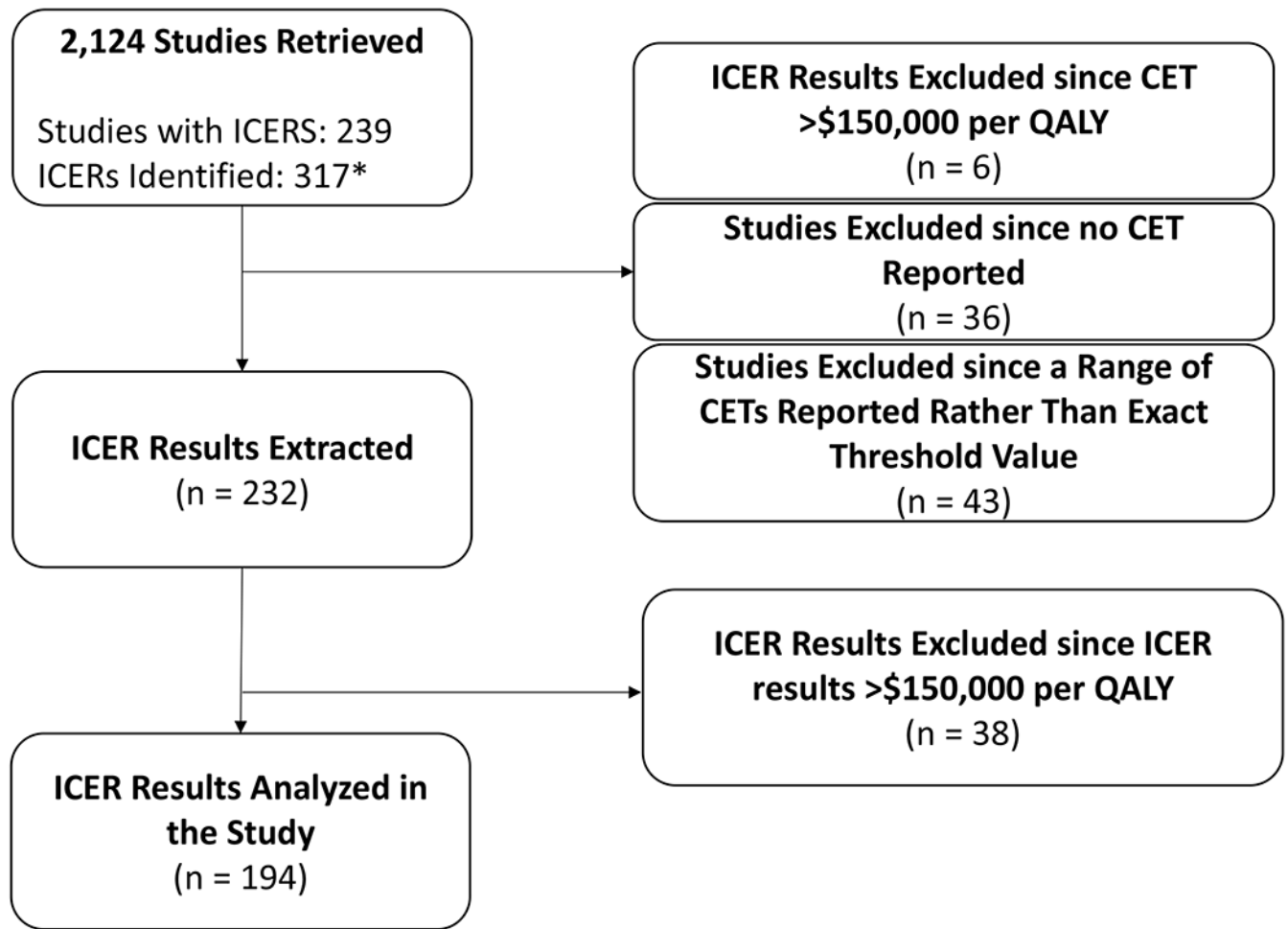
1. Kesselheim AS, Avorn J, Sarpatwari A. The High Cost of Prescription Drugs in the United States: Origins and Prospects for Reform. *JAMA*. 2016 Aug 23-30;316(8):858–71. [PubMed: 27552619]
2. Pandya A Adding Cost-effectiveness to Define Low-Value Care. *JAMA*. 2018 May 15;319(19):1977–8. [PubMed: 29710341]
3. Sanders GD, Neumann PJ, Basu A, Brock DW, Feeny D, Krahn M, et al. Recommendations for Conduct, Methodological Practices, and Reporting of Cost-effectiveness Analyses: Second Panel on Cost-Effectiveness in Health and Medicine. *JAMA*. 2016 Sep 13;316(10):1093–103. [PubMed: 27623463]
4. Neumann PJ, Kim DD, Trikalinos TA, Sculpher MJ, Salomon JA, Prosser LA, et al. Future Directions for Cost-effectiveness Analyses in Health and Medicine. *Med Decis Making*. 2018 Oct;38(7):767–77. [PubMed: 30248277]
5. Sculpher M, Claxton K, Pearson SD. Developing a Value Framework: The Need to Reflect the Opportunity Costs of Funding Decisions. *Value Health*. 2017 Feb;20(2):234–9. [PubMed: 28237201]



6. Woods B, Reville P, Sculpher M, Claxton K. Country-Level Cost-Effectiveness Thresholds: Initial Estimates and the Need for Further Research. *Value Health*. 2016 Dec;19(8):929–35. [PubMed: 27987642]
7. Gold M, Siegel JE, Russell LB, Weinstein MC. *Cost-Effectiveness in Health and Medicine*. New York: Oxford University Press; 1996.
8. Neumann PJ, Sanders GD, Russel LB, Siegel JE, Ganiats TG, editor. *Cost-effectiveness in Health and Medicine*. Second ed. New York: Oxford University Press; 2017.
9. Walker S, Griffin S, Asaria M, Tsuchiya A, Sculpher M. Striving for a Societal Perspective: A Framework for Economic Evaluations When Costs and Effects Fall on Multiple Sectors and Decision Makers. *Applied Health Econ Health Policy*. 2019 May 16.
10. Grosse SD. Assessing cost-effectiveness in healthcare: history of the \$50,000 per QALY threshold. *Expert Review Pharmacoeconomics & Outcomes Research*. 2008;8(2):165–78.
11. Neumann PJ, Cohen JT, Weinstein MC. Updating cost-effectiveness--the curious resilience of the \$50,000-per-QALY threshold. *N Engl J Med*. 2014 Aug 28;371(9):796–7. [PubMed: 25162885]
12. World Health Organization. *Macroeconomics and health: investing in health for economic development*. Geneva: World Health Organization; 2001.
13. Anderson JL, Heidenreich PA, Barnett PG, Creager MA, Fonarow GC, Gibbons RJ, et al. ACC/AHA statement on cost/value methodology in clinical practice guidelines and performance measures: a report of the American College of Cardiology/American Heart Association Task Force on Performance Measures and Task Force on Practice Guidelines. *J American College Cardiology*. 2014 Jun 3;63(21):2304–22.
14. The World Bank. GDP per capita (current US\$), 2015. 2016 [cited 2018 February 20]; Available from: <http://data.worldbank.org/indicator/NY.GDP.PCAP.CD>.
15. Phelps CE. A New Method to Determine the Optimal Willingness to Pay in Cost-Effectiveness Analysis. *Value Health*. 2019 Jul;22(7):785–91. [PubMed: 31277825]
16. Snijders TAB, Bosker RJ. *Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling*. 2nd ed. Los Angeles, CA: Sage Publications; 2012.
17. Bell CM, Urbach DR, Ray JG, Bayoumi A, Rosen AB, Greenberg D, et al. Bias in published cost effectiveness studies: systematic review. *BMJ*. 2006 Mar 25;332(7543):699–703. [PubMed: 16495332]
18. Institute for Clinical and Economic Review. *2020-2023 Value Assessment Framework*. Cambridge, MA: Institute for Clinical and Economic Review.; 2020.
19. Holtgrave DR, Qualls NL, Graham JD. Economic Evaluation of HIV Prevention Programs. *Annual Review Public Health*. 1996;17:467–88.
20. Freedberg KA, Tosteson AN, Cotton JD, Goldman L. Optimal management strategies for HIV-infected patients who present with cough or dyspnea: a cost-effectiveness analysis. *J General Internal Med*. 1992;7:261–72.
21. Garber AM, Phelps CE. Economic Foundations of Cost-effectiveness Analysis. *J Health Econ*. 1997;16:1–31. [PubMed: 10167341]
22. Towse A, Pritchard C. National Institute for Clinical Excellence (NICE): Is economic appraisal working? *Pharmacoeconomics*. 2002;20 Suppl 3:95–105. [PubMed: 12457430]
23. Bridges JF, Onukwugha E, Mullins CD. Healthcare rationing by proxy: cost-effectiveness analysis and the misuse of the \$50,000 threshold in the US. *Pharmacoeconomics*. 2010;28(3):175–84. [PubMed: 20067332]
24. Braithwaite RS. Willingness to pay for a quality-adjusted life year in search of a standard. *Med Care*. 2008;46(4):349–56. [PubMed: 18362813]
25. Hirth RA. Willingness to pay for a quality-adjusted life year in search of a standard. *Med Decis Making*. 2000;20(3):332–42. [PubMed: 10929856]
26. Weinstein MC. How much are Americans willing to pay for a quality-adjusted life year? *Med Care*. 2008 Apr;46(4):343–5. [PubMed: 18362811]
27. Pauly MV. The Questionable Economic Case for Value-Based Drug Pricing in Market Health Systems. *Value Health*. 2017 Feb;20(2):278–82. [PubMed: 28237209]

### Key Points for Decision Makers

- The selection of cost-effectiveness threshold (CET) may be endogenous to economic model outputs in terms of the incremental cost-effectiveness ratio (ICER) based on concern that cost-utility analyses use higher thresholds to compensate for low-value technology.
- Through systematic review, we illustrated an association linking the selection of higher CET (e.g. \$150,000/QALY) with ICERs that would not be cost-effective at traditional lower thresholds (e.g. \$50,000/QALY, or below \$100,000/QALY).
- Selection of a CET may come after the ICER is calculated to infer value that suits a hypothesis.



\* Some studies reported more than one ICER

**Figure 1.** Flow diagram of study selection based on inclusion criteria, reported incremental cost-effectiveness ratios (ICERs) and cost-effectiveness thresholds (CET).

**Table 1.**

Inclusion and Exclusion Criteria for Literature Review

| Inclusion Criteria   | Exclusion Criteria  |
|--|---|
| <ol style="list-style-type: none"> <li>1. English language</li> <li>2. U.S. perspective study</li> <li>3. The intervention arm of the study was drug</li> <li>4. Study included QALYs as one of the primary outcomes</li> <li>5. Study referenced a single cost-effectiveness threshold</li> <li><b>6. Peer-Reviewed and Published on Medline by 2017</b></li> </ol> | <ol style="list-style-type: none"> <li>1. Non-English language</li> <li>2. Non-U.S. perspective</li> <li>3. Study did not include clear-referenced cost-effectiveness threshold</li> <li>4. Study included cost-effectiveness threshold that is less than \$50,000 per QALY</li> <li>5. Dominant or dominated strategy</li> <li>6. Not Peer-Reviewed or Unavailable through Medline.</li> </ol> |

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

**Table 2.**

Characteristics of cost-utility analyses in the systematic review.

| Characteristics                                | Number (n) | Percent (%) |
|--|------------|-------------|
| <b>ICER (USD per QALY)</b>                     |            |             |
| < \$50,000                                     | 186        | 58.4        |
| \$50,000 - \$99,999                            | 53         | 16.7        |
| \$100,000 - \$149,999                          | 20         | 6.3         |
| \$150000                                       | 58         | 18.6        |
| <b>Cost-effectiveness threshold referenced</b> |            |             |
| Single   | 238        | 75.1        |
| Multiple <sup>a</sup>                          | 43         | 13.6        |
| No threshold referenced                        | 36         | 11.4        |

ICER, incremental cost-effectiveness ratio; QALY, quality-adjusted life-year.

<sup>a</sup>Category includes multiple and range of cost-effectiveness thresholds.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

**Table 3.**

Cost-effectiveness results categorized by ICER and Cost-effectiveness threshold.

| ICER                  | Cost-effectiveness Threshold |                       | P-value |
|-----------------------|------------------------------|-----------------------|---------|
|                       | < \$100,000                  | \$100,000 - \$150,000 |         |
| < \$100,000           | 175                          | 4                     | <0.001* |
| \$100,000 - \$150,000 | 10                           | 5                     |         |

Among those ICER results referencing single cost-effectiveness thresholds, 194 reported that both the ICER and cost-effectiveness threshold fell in the range of \$50,000 to \$150,000 per QALY.

\* Pearson Chi-squared (df=1) = 26.40.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript



**Table 4.**

Results of fixed-effects regression analyses exploring the associations between cost-effectiveness threshold (CET) and incremental cost-effectiveness ratios (ICER) of reported studies, controlling for continuous ICER values and time-interruptions.

| <i>CET</i>                | Model 1     |          |         | Model 2     |          |         |
|---------------------------|-------------|----------|---------|-------------|----------|---------|
|                           | Coefficient | SD       | P-Value | Coefficient | SD       | P-Value |
| <i>Intercept</i>          | 52,203.16   | 2,800.79 | <0.001  | 51,101.41   | 2,692.29 | <0.001  |
| <i>ICER</i>               | 0.37        | 0.05     | <0.001  | 0.31        | 0.05     | <0.001  |
| <i>Inter-ruption 2014</i> |             |          |         | 18,542.10   | 4,291.52 | <0.001  |
| R-Squared:                | 0.2058      |          |         | 0.2765      |          |         |

Definitions: CET: cost-effectiveness threshold; ICER, incremental cost-effectiveness ratio