International Journal of Technology Assessment in Health Care

cambridge.org/thc

Using QALYs versus DALYs to measure cost-effectiveness: How much does it matter?

Xue Feng, David D. Kim, Joshua T. Cohen, Peter J. Neumann and Daniel A. Ollendorf

Center for the Evaluation of Value and Risk in Health, Institute for Clinical Research and Health Policy Studies, Tufts Medical Center, Boston, MA, USA

Objectives. Quality-adjusted life-years (QALYs) and disability-adjusted life-years (DALYs) are commonly used in cost-effectiveness analysis (CEA) to measure health benefits. We sought to quantify and explain differences between QALY- and DALY-based cost-effectiveness ratios, and explore whether using one versus the other would materially affect conclusions about an intervention's cost-effectiveness.

Methods. We identified CEAs using both QALYs and DALYs from the Tufts Medical Center CEA Registry and Global Health CEA Registry, with a supplemental search to ensure comprehensive literature coverage. We calculated absolute and relative differences between the QALY- and DALY-based ratios, and compared ratios to common benchmarks (e.g., 1× gross domestic product per capita). We converted reported costs into US dollars.

Results. Among eleven published CEAs reporting both QALYs and DALYs, seven focused on pharmaceuticals and infectious disease, and five were conducted in high-income countries. Four studies concluded that the intervention was "dominant" (cost-saving). Among the QALY- and DALY-based ratios reported from the remaining seven studies, absolute differences ranged from approximately $2 to $15,000 per unit of benefit, and relative differences from 6–120 percent, but most differences were modest in comparison with the ratio value itself. The values assigned to utility and disability weights explained most observed differences. In comparison with cost-effectiveness thresholds, conclusions were consistent regardless of the ratio type in ten of eleven cases.

Conclusions. Our results suggest that although QALY- and DALY-based ratios for the same intervention can differ, differences tend to be modest and do not materially affect comparisons to common cost-effectiveness thresholds.

Method

Cite this article: Feng X, Kim DD, Cohen JT, Neumann PJ, Ollendorf DA (2020). Using QALYs versus DALYs to measure cost-effectiveness: How much does it matter? *International Journal of Technology Assessment in Health Care* **36**, 96–103. https://doi.org/10.1017/S0266462320000124

Received: 23 May 2019
Revised: 17 December 2019
Accepted: 19 February 2020

Key words:
QALY; DALY; Cost-effectiveness analysis; Medical-decision

© Cambridge University Press 2020. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted re-use, distribution, and reproduction in any medium, provided the original work is properly cited.

CAMBRIDGE UNIVERSITY PRESS

Cost-per-quality-adjusted life-year gained (QALY) and cost-per-disability-adjusted life-year averted (DALY) studies have become commonly used measures in the current practice of cost-effectiveness analysis (CEA) (1;2). In recent decades, childhood and adult mortality rates have continued to decline worldwide (3), whereas morbidity, or time lived with health loss, has become a more serious concern in high-income countries (HICs) and low- and middle-income countries (LMICs) alike. Consequently, QALYs and DALYs are being used increasingly frequently to assess the cost-effectiveness of interventions that affect quality as well as length of life. Because of this, a practical question for researchers and decision makers is which measure should be applied to a given intervention in a specific setting, whether defined by disease, geography, a country's per capita income, or another characteristic. The QALY-based measure has been recommended by many health technology assessment agencies in HICs, whereas the DALY-based measure is generally preferred in LMICs (4;5). Indeed, the number of published cost-per-DALY studies has substantially increased in LMICs over the past decade (2). One possible reason is that the freely and publicly available disability weights, which required for DALY calculations, can significantly reduce the cost of conducting CEA in resource-constrained settings. In contrast, cost-per-QALY studies, which use utility estimates for specific states and capabilities, may require more time and resources to collect or assemble (6).

QALYs and DALYs: Similarities and Differences

Conceptual Approach

Although the intent in using QALYs or DALYs in CEAs is similar, the theoretical and technical underpinnings of the two metrics differ (1). The concept of the QALY was developed in the 1960s; it represents the products of years lived and the associated utility values, ranging from 0 (dead) to 1 (perfect health). Utility estimates represent the perspective of an individual's values or preferences, based on the central tenet of "welfarist" economics—that individuals are the best judges of their own welfare, and improved societal welfare as the ultimate goal

is based on the sum of these individual utilities. However, QALYs also integrate so-called "extra-welfarist" elements to utility assessment, such as the contribution of particular states of health, functioning, and patient preferences to utility estimation (7;8). The primary application of QALYs has been the same ever since their initial use—to compare the benefits and risks of medical interventions (9).

In contrast, the DALY was developed in the 1990s by the Global Burden of Diseases, Injuries, and Risk Factors (GBD) initiative to assess burden of disease at a population level, to understand leading causes of health loss worldwide, and to compare population health across geographic settings (10). DALYs reflect the sum of years of life lost due to premature mortality and years lived with disability. The disability weights used for DALYs are inverse to that of utility weights, with "0" referring to no disability and "1" representing the dead state. DALYs also do not explicitly integrate extra-welfarist concepts; for example, disability weights are defined not based on surveys of individuals but based on expert opinion, as in the view of its developers a single set of weights anchored to specific diseases better facilitated cross-cultural comparisons than did some form of self-assessment (9). In addition, non-health effects are limited to age and sex alone. In recent years, GBD has refined its disability weights to attempt to isolate health loss from welfare loss and social context (11); these weights are intended to be universal and invariant to setting or population but are still undergoing further testing.

Empirical Comparisons

A number of other studies have discussed the theoretical differences between QALYs and DALYs (9;12–16), and have generally concluded that both measures have proven serviceable for resource allocation and priority-setting, but do differ in terms of estimation. For example, Sassi et al found that numeric differences between utility and disability weights may lead to further divergence between the QALYs and DALYs (12). Age weighting was also considered a major difference between the two measures (9), although the GBD no longer recommends such weighting. Airoldi et al. found that QALY gained is consistently larger than DALY averted because of the reference age used; differences tend to become larger for older ages (13). Given that an intervention may have differential impacts on population subgroups defined by age, the choice to adopt one measure over the other may further affect the process of healthcare decision making when considering potential interventions to fund. A recent study by Augustovski et al. used two models from empirical studies to evaluate the impact of using QALY- and DALY-based methods (16). The authors concluded that differences between the two approaches (e.g., the effects of discounting) could affect the magnitude of QALY and DALY estimates, and therefore influence policy decisions. However, the structural uncertainty introduced by use of the QALY versus DALY was similar to that associated with other key model assumptions. Despite these analyses, there remains a lack of empirical studies directly comparing the two measures to assess their relationship and explore whether the choice of one versus another affects decision making in practice. Hence, the objective of this study was to quantify differences between CEA using DALYs versus QALYs, and to assess the reasons for differences. We also evaluated whether using one versus the other measure would affect conclusions about the favorability of an intervention's cost-effectiveness.

Methods

Inclusion Criteria

Studies included English-language CEA articles that reported results using both cost-per-QALY and cost-per-DALY measures, published from 1996 through 2018.

Data Sources

We utilized two databases maintained by the Center for the Evaluation of Value and Risk in Health at Tufts Medical Center in Boston, Massachusetts: the CEA Registry (http://www.cearegistry.org), with information on 7,287 cost-per-QALY studies collected from 1976 to 2017, and the Global Health (GH) CEA Registry (http://www.ghcearegistry.org), which summarizes 620 cost-per-DALY studies from 1996 to 2017. The search strategies, data collection process, and review methods are similar for the two registries and have been described previously (1;17;18). We used the title and PubMed ID of the article to identify whether there were studies contained by both registries. If so, the identified study was deemed eligible based on our inclusion criteria, as both QALY- and DALY-based ratios were reported in the same article for the same intervention(s).

We also performed a supplemental search to identify articles published since 2018 using databases of PubMed, EMBASE, and Econlit to identify articles reporting results by both measures. We followed the same steps as mentioned above and used keywords of "QALYs," "quality-adjusted," "DALYs," and "disability-adjusted," to identify candidate papers.

Variables and Analysis of Data

We extracted information from the selected articles, including year of publication, intervention type, study region, disease area, study funder, study perspective, cost discount rate, DALY and QALY discount rate, age-weighting use, sources of disability weights, sources of utilities, cost-per-QALY gained and cost-per-DALY averted results in the base case, the use of a cost-effectiveness "threshold" for decision making as mentioned by the authors, and the conclusions of the study.

We quantified the differences between ratios by QALY and DALY measures based on their absolute and relative difference. Relative difference was defined as the absolute difference divided by the QALY-based ratio. Magnitudes of both types of differences were compared. We also counted the number of cases for which the cost-per-DALY was higher than the cost-per-QALY for each intervention studied. All costs estimated in non-U.S. currency were converted to United States (U.S.) dollars based on the present value year used in each article as we intended to evaluate differences within rather than between studies. In addition, we compared the QALY- and DALY-based ratios to commonly used cost-effectiveness thresholds, including those reported by the articles such as one time gross domestic product (GDP) per capita, as well as any country-specific thresholds mentioned in the articles. Because our sample size was expected to be small, primary analyses were descriptive in nature.

Furthermore, we estimated the net monetary benefit (NMB) based on QALY and DALY measures, respectively ($NMB = \Delta QALY\ [DALY] \times threshold - \Delta Cost$) so the results from both measures could be expressed in the same unit of U.S. dollars for further comparison. When calculating the NMB, we applied the threshold reported in each article, whether based on a

https://doi.org/10.1017/S0266462320000124 Published online by Cambridge University Press

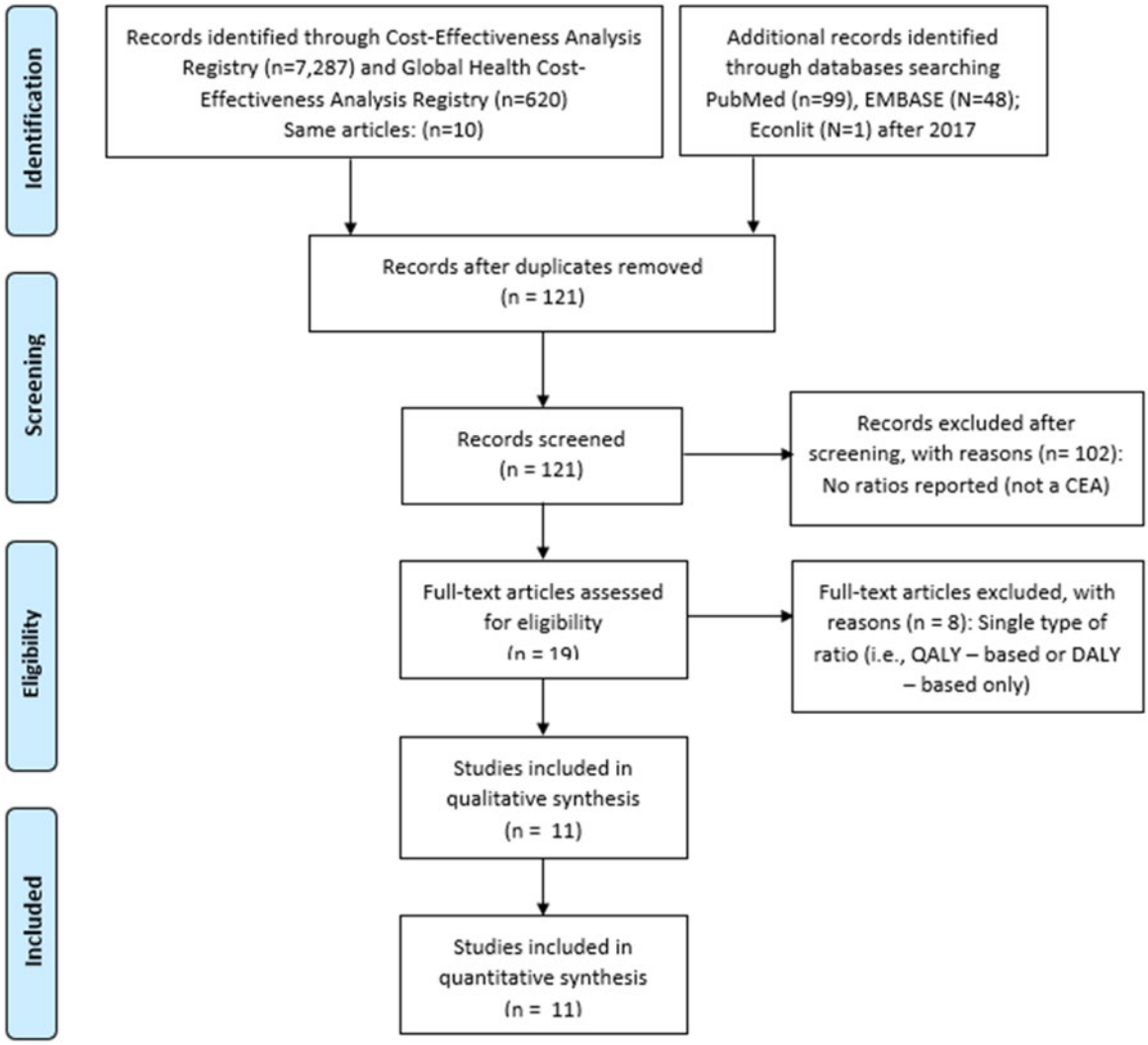


Fig. 1. Search strategy.

commonly used benchmark (e.g., 1× GDP per capita) or a country-specific estimate. Costs were presented in 2018 U.S. dollars for this analysis. The Pearson correlation coefficient was used to examine the relationship between the relative differences as assessed by NMB and the relative differences based on ratios.

Results

In total, we obtained eleven articles—ten articles from the two Tufts Medical Center registries and another 2018 article identified through the literature search (Figure 1) (19–29). Among the eleven articles in Table 1, seven (64 percent) focused on infectious diseases (i.e., HIV, TB, hepatitis B, hepatitis C, and rotavirus infections). Most of the articles (82 percent, 9/11) were published from 2015 to 2018. Five (45 percent) of the studies were from high income settings. Pharmaceutical interventions were assessed in seven studies (64 percent, 7/11); other types of interventions included immunization, care delivery, surgery, health education and behavior, legislation, and nutrition. Studies received funding from various sources such as government, foundations, academic institutions, healthcare organizations, the pharmaceutical industry, and other agencies. Studies were conducted using the perspectives of the healthcare sector (36 percent, 4/11) or healthcare payer (36 percent), or with a limited societal perspective (27 percent, 3/11). Most studies applied a discount rate of 3 percent for costs, QALYs, and DALYs. One study reported using age-weighting for the DALY measure. Cost-effectiveness threshold benchmarks of 1× or 3× GDP per capita were mentioned in all studies from LMICs, whereas country-specific thresholds (e.g., Australia: 50,000 AU$; The Netherlands: 20,000 euros) were used in the HICs. Disability weights from GBD sources were cited in eight articles (73 percent, 8/11), and utilities were obtained from a variety of sources, often not specific to the study setting. For example, utilities in a Zambia-based intervention cited a previous study in another African country (19); a Malawi study applied utilities from an Indian setting (29); and a Gambia study used utilities from multiple countries (26). Most of the included studies (64 percent, 7/11) applied a Markov model; other modeling techniques included decision-tree, stochastic simulation, and metapopulation and compartment modeling. Only two studies stated they used primary data from specific clinical trials (26;28) to inform effectiveness calculations.

Four articles reported that the intervention of interest was "cost-saving" relative to the comparator (i.e., no cost-effectiveness

https://doi.org/10.1017/S0266462320000124 Published online by Cambridge University Press

https://doi.org/10.1017/S0266462320000124 Published online by Cambridge University Press

Table 1. Study characteristics and methodology ($N = 11$)

	Article	Setting	Disease area	Intervention type	Study funder	Study perspective by reader	Threshold mentioned	Discount rate (%)	Source for disability weights
1	Ryan et al. (19)	Zambia	HIV infection	Pharmaceutical	Government	Healthcare payer	1× GDP per capita	3	GBD
2	Mangen et al. (20)	The Netherlands	Rotavirus infection	Immunization	Not determined	Limited societal	€20k	3	The Netherlands
3	Mihalopoulos et al. (21)	Australia	PTSD	Care delivery; health education and behavior; pharmaceutical; legislation	Government, non-profit organization	Healthcare sector	AU$50,000	3	Australia
4	Long et al. (22)	USA	Obesity	Nutrition; legislation	Government, foundation	Limited societal	NA	3	NA
5	Diel et al. (23)	UK	TB	Pharmaceutical	Not determined	Limited societal	€10k	3	GBD
6	Wolfson et al. (24)	Germany	TB	Pharmaceutical	Pharmaceutical industry	Healthcare payer	NA	3.5	GBD
7	Nguyen et al. (25)	Papua New Guinea, Australia	TB	Care delivery, Pharmaceutical, screening	Academic institutions	Healthcare payer	1× GDP per capita AU$35,000–120,000 AU$40,000–80,000	3	GBD
8	Nayagam et al. (26)	Gambia	Hepatitis B	Screening, pharmaceutical	Political and economic organization	Healthcare sector	1× GDP per capita $240 per DALY 3× GDP per capita	3	GBD
9	Aggarwal et al. (27)	India	Hepatitis C	Pharmaceutical	None	Healthcare payer	3× GDP per capita	3	GBD
10	Löfgren et al. (28)	Uganda	Groin hernia	Surgery	Government, foundation, academic institutions, healthcare organizations, pharmaceutical industry, other	Healthcare sector	1× GDP per capita	NA	GBD
11	Vetrini et al. (29)	Malawi	Diabetic retinopathy and macular edema	Screening, medical procedure	Foundation	Healthcare sector	1× GDP per capita, 3× GDP per capita	3	GBD

AU, Australian; GBD, global burden of diseases, injuries, and risk factors; GDP, gross domestic product; NA, not applicable; PTSD, post-traumatic stress disorder; TB, tuberculosis; UK, the United Kingdom; USA, the United States.

Table 2. Differences between cost-per-QALY and cost-per-DALY ($N = 11$)

	Article	ICER ($/unit of benefit [DALY or QALY])		Absolute difference	Relative difference (%)	QALY-based ratio > DALY-based ratio
		QALY-based	DALY-based			
1	Ryan et al. (19)	94	53	41	43.6	Yes
2	Mangen et al. (20)					
	Rotarix®	≈69,803	55,842	≈13,961	≈20.0	Yes
	RotaTeq®	≈76,356	61,540	≈14,815	≈19.4	Yes
3	Mihalopoulos et al. (21)					
	Adults	16,967	14,288	2,679	15.8	Yes
	Children	7,948	7,144	804	10.1	Yes
4	Long et al. (22)			Cost-saving		
5	Diel et al. (23)			Cost-saving		
6	Wolfson et al. (24)			Cost-saving		
7	Nguyen et al. (25)					
	Papua New Guinea—South Fly	3.5	5.3	1.8	52.0	No
	Australia—Torres Straits Islands	6,077	8,462	2,385	39.2	No
8	Nayagam et al. (26)	511	540	29	5.7	No
9	Aggarwal et al. (27)			Cost-saving		
10	Löfgren et al. (28)					
	Low-cost mesh	7.6	16.8	9.2	121.5	No
	Commercial mesh	33.3	58.2	24.9	74.7	No
11	Vetrini et al. (29)	400	766	366	91.5	No

DALYs, disability-adjusted life-years; QALYs, quality-adjusted life-years.

ratio was calculated) (Table 2). Among the seven remaining studies with the eleven intervention-specific QALY- and DALY-based ratios, cost-per-DALY results were higher than cost-per-QALY in six cases, whereas the reverse was seen in the other five instances. The magnitude of difference between the two measures also varied across studies. The relative differences between the two measures ranged from 6 to 122 percent, and absolute differences from approximately $2 to $15,000. However, the magnitude of difference was consistently modest, even in cases with seemingly large differences (Figure 2). For example, the study reporting an absolute difference of $15,000 between ratios for rotavirus vaccines in The Netherlands (20), had ratios that were both relatively high; as a result, the relative difference between ratios was only 19 percent. In contrast, the seemingly large relative difference of 122 percent was from a study of low-cost surgical mesh in a LMIC, with an absolute difference of only $9 between ratios (28). We were able to conduct our secondary analysis of NMB on seven interventions. In general, relative differences using these estimates were consistent with those directly employing the QALY- and DALY-based ratios (Pearson correlation coefficient = .67; Supplementary Table).

In many (73 percent) of these studies, global disability weights from the GBD studies were employed for DALY estimation, versus locally derived utility weights for QALYs. Few authors elaborated the possible reasons for the differences. For example, in the study of surgical mesh (28), the authors found that the estimate of the cost-per-QALY ratios were approximately half of the ratios estimated using the cost-per-DALY measures (relative differences: 122 and 75 percent for low-cost and commercial mesh, respectively); the authors posited that the GBD algorithm may have underestimated the magnitude of disability associated with groin hernia in the study country (Uganda). This may also explain the comparatively large relative difference (91.5 percent) seen in a study of screening and laser treatment for diabetic retinopathy and macular edema in Malawi (29). On the other hand, much smaller relative differences were observed in articles with disability weights and utilities from the same or similar contexts. For example, the relative differences between ratios were approximately 10 percent in an Australian analysis of a multi-component intervention for post-traumatic stress disorder, which featured utilities and disability weights that were both Australia-derived (20).

In Figure 2, we present the ratios of cost-per-QALY and cost-per-DALY compared with a set of threshold benchmarks for decision making for LMICs and HICs separately. Among eleven pairs of QALY- and DALY-based ratios, we identified only one instance of a change in favorability of results when compared to a cost-effectiveness threshold (29). For the remaining ten pairs with consistent conclusions, two pairs from the same study were not considered to be "cost-effective" interventions using the country-specific threshold or 1× GDP per capita (20), for both QALY- or DALY-based ratios. Both ratios for another intervention were slightly above the threshold of 1× GDP per capita in the study country (26).

https://doi.org/10.1017/S0266462320000124 Published online by Cambridge University Press

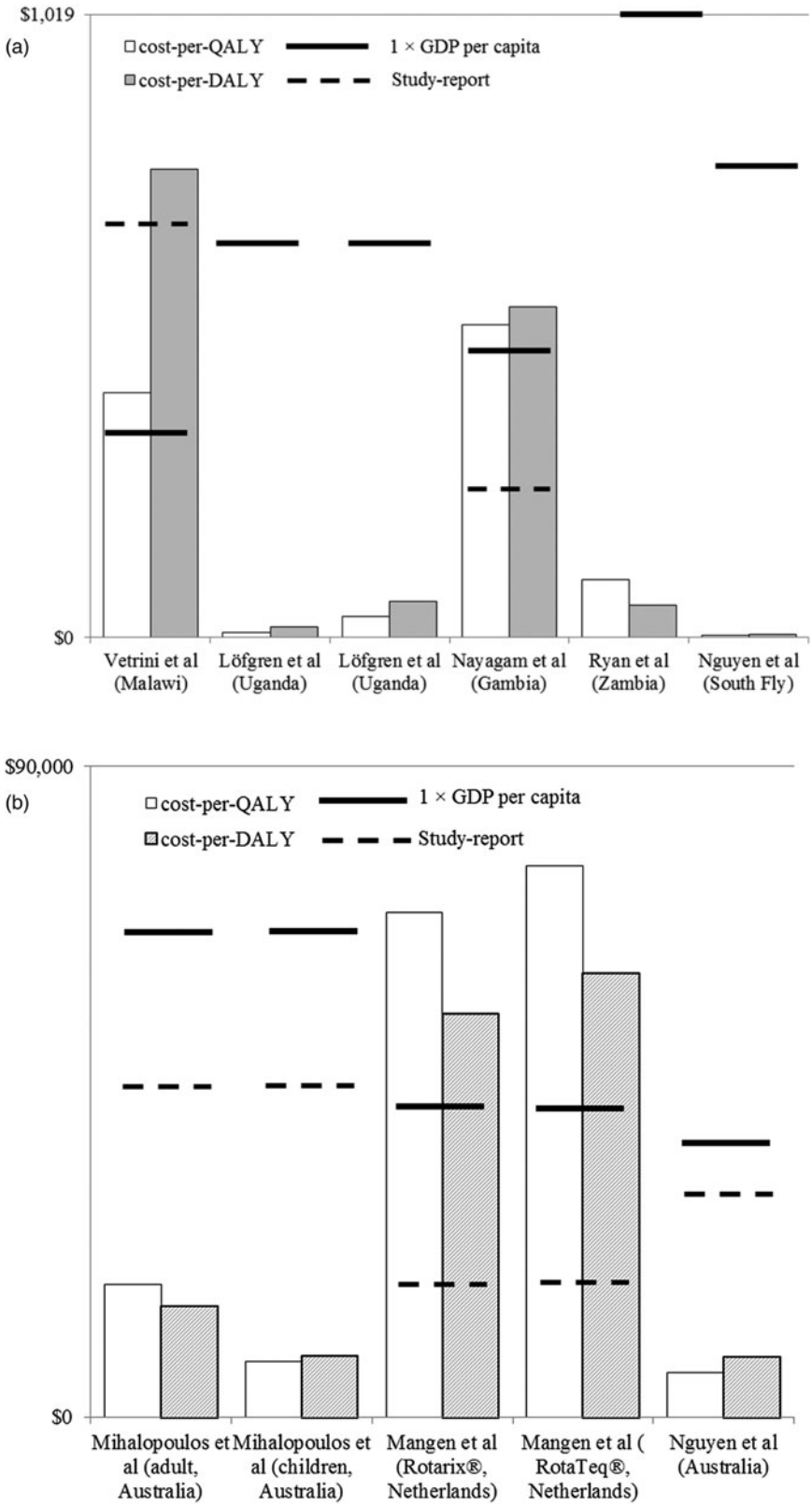


Fig. 2. Cost-per-QALY, cost-per-DALY, and thresholds: (a) LMICs and (b) HICs.
Note: Study-report: Australia: 50,000 AU$ in 2005 for Mihalopoulos et al.; The Netherlands: 20,000 euros for Mangen et al.; 40,000 AU$ in 2015 for Nguyen et al.; 679 US$ for Vetrini et al.; 240 US$ for Nayagam et al.
DALYs, disability-adjusted life-years; QALYs, quality-adjusted life-years; GDP, gross domestic product.

Discussion

Our study represents an attempt to quantify differences in estimates of cost-effectiveness based on QALY- and DALY measures when both were used in the same evaluation, and to explore possible reasons for these differences. Our findings suggest that differences were modest in relation to each ratio's magnitude in most cases. Perhaps more importantly, in the vast majority of cases, these differences would not affect CEA conclusions or decisions based on commonly used thresholds for cost-effectiveness.

https://doi.org/10.1017/S0266462320000124 Published online by Cambridge University Press

On the other hand, these modest differences may still have the potential to affect decisions to fund or not fund the health interventions; decision making can be influenced by many factors (e.g., the opportunity cost of other interventions) that vary within the specific contexts of each country. In addition, the motivation to use the two measures for the same intervention was not clearly stated in the included articles. For example, two studies mentioned that the two measures were the most commonly used metrics (21;25), and another posited that the use of the two measures may increase the robustness of the analyses (21). We cannot rule out the possibility of self-selection, however, potentially manifested here by focus on models, treatments, and conditions that would have ensured concordance of results between the two measures.

One of the major issues is that many of the included studies did not provide sufficient details on model specification to explain the factors associated with ratio-based differences. It is likely, however, that source of utilities and disability weights were a major driver. This may be a particular issue in LMICs, because respondents used as the basis of global estimates of disability weights were primarily from high-income settings in GBD studies (30–32), and because utility data often must be obtained from settings other than the location of interest for the study. In addition, comparatively small absolute differences in cost/QALY and cost/DALY ratios were often observed in studies targeted for LMICs, which reflects the relatively low costs of the interventions in these studies. For instance, the total cost of the intervention of low-cost mesh for groin hernia repair was only $49 (28). In such situations, differences in the method used to measure health gain may be less important given that the major driver of results is the low incremental cost itself. Still, this small absolute difference may impose substantial cost for the payers when considering budget planning for the covered population, particularly if the intervention will affect large numbers of individuals. Likewise, depending on population size, absolute differences in CEA estimates using DALYs versus QALYs may have an effect on price negotiations that could have quite considerable implications.

Our findings are consistent with those of previous studies, which concluded that the weight (disability weight vs. utility) used and age-weighting functions are major drivers of differences between QALY and DALY measures (9;13). However, age-weighting was used in only one of the included studies, suggesting that recent studies have adopted the 2010 guidance to remove these weighting functions. The cessation of use of age-weighting was in response to criticisms that the practice was potentially unethical and discriminatory (9;33), given that age-weighting assigns higher values to young- and middle-aged adults because of their higher potential for productivity. Therefore, differences in the utility and disability weights, as well as the sources of those estimates, are likely the major explanatory factors in our sample.

The conclusion about the interventions' "acceptable" cost-effectiveness was only affected by the type of ratio used in one case. In that case, a study-reported threshold of $679 (per QALY gained or DALY averted) was used; however, if a more common threshold such as one time GDP per capita ($333 in this case) had been used, the intervention would have been found to be cost-ineffective regardless of the type of ratio employed. With regard to policy making, country- and context-specific thresholds have been suggested to decide whether an intervention is considered a priority in healthcare planning (34). These thresholds may be more informative in the process of decision making when one also considers the budget for healthcare spending, and decision makers' willingness to divert funds from other healthcare interventions and/or consumption outside the healthcare sector. Whether or not decisions would differ for QALY- and DALY-based estimates using specific thresholds needs further exploration.

We acknowledge several limitations in our study. First, given likely differences in model structure, estimation, and programing language, among others, for the same intervention among different CEAs, it is likely not feasible to adjust for these differences or directly compare the cost-per-QALY and cost-per-DALY results generated from different studies. Our exclusion of such studies limited the number of articles to those that used both measures in the same evaluation, which in turn limited our sample size and precluded the use of hypothesis testing. Second, we cannot rule out the possibility that the observed consistency between QALY- and DALY-based measures may be due in part to publication bias, manifested by a predisposition to publish studies with consistent results, whether favorable or unfavorable. Moreover, the calculation of the differences between the two ratios and the use of a single threshold for decision making recommendations is based on an assumption that the QALY and the DALY reflect comparable constructs. As described previously, these measures reflect somewhat different domains of health and may not be readily exchangeable (11;30;35). If this holds true, then different thresholds are likely required to inform decision making. On the other hand, the differences in construction and interpretation of DALYs and QALYs are not likely to affect the interpretation of our findings from the perspective of the current application of CEA to decision making; however, as decision making thresholds for cost-effectiveness have remained relatively constant over time. Although we acknowledge this limitation, we are unaware of any empirical research to quantify differences in QALY- and DALY-based ratios, and so the full implications of our assumptions are not known. We note that findings were similar when QALY- and DALY-based results were presented using common units in our NMB calculations.

Despite these limitations, this is the first study using published CEAs to assess the potential relationship between QALYs and DALYs and to compare cost-effectiveness ratios with different thresholds. We find that, although nominal differences in results are observed, conclusions of the CEAs are not likely to change based on the use of QALYs versus DALYs to measure health gain, when the commonly used thresholds for CE are applied. Our findings should be of interest to policy makers and researchers in LMICs, particularly those who may be limited to DALY-based analyses because of constraints of resource and data collection costs, as well as those who do have the ability to estimate QALYs, but are concerned about the challenge of doing so in a climate dominated by DALY-based research.

Conclusions

Our results suggest that although QALY- and DALY-based ratios for the same intervention can differ, differences tend to be modest and are unlikely to materially affect resource allocation recommendations. On the other hand, the modest differences may still affect decision making process when considered from a broader perspective, including opportunity cost of other healthcare interventions, budgets for healthcare spending, and price negotiation. Although both QALYs and DALYs can produce cost-effectiveness estimates that assist in healthcare decision making, further studies are warranted to better improve the methodologies and applications of these measures to address local health needs and concerns.

https://doi.org/10.1017/S0266462320000124 Published online by Cambridge University Press

Supplementary Material. The supplementary material for this article can be found at https://doi.org/10.1017/S0266462320000124.

Financial Support. This study was funded by the Bill and Melinda Gates Foundation (grant no. OPP1171680).

Conflict of Interest. Dr. Feng reports no conflict of interest. Dr. Kim reports grants from the Bill and Melinda Gates Foundation, during the conduct of the study. Dr. Cohen reports grants from the Bill and Melinda Gates Foundation, during the conduct of the study; personal fees from Precision Health Economics, personal fees from Sage Therapeutics, personal fees from Merck Corporation, personal fees from Abbvie, personal fees from Indivior, personal fees from Sarepta Pharmaceuticals, outside the submitted work. Dr. Neumann reports advisory boards or consulting from Abbvie, Amgen, Avexis, Bayer, Congressional Budget Office, Vertex, Veritech, Janssen, Merck, Novartis, Novo Nordisk, Precision Health Economics, funding from The CEA Registry Sponsors by various pharmaceutical and medical device companies; grants from Amgen, Lundbeck, Bill and Melinda Gates Foundation, NPC, and Alzheimer's Association, NIH. Dr. Ollendorf reports grants from the Bill and Melinda Gates Foundation, during the conduct of the study; personal fees and other from Sarepta Therapeutics, LLC, personal fees from DBV Technologies, Inc., personal fees from EMD Serono, other from Gerson Lehman Group, other from The CEA Registry Sponsors, personal fees from Autolus, Inc., outside the submitted work.

References

1. **Neumann PJ, Anderson JE, Panzer AD, Pope EF, D'Cruz BN, Kim DD, Cohen JT**. Comparing the cost-per-QALYs gained and cost-per-DALYs averted literatures. *Gates Open Res*. 2018;**2**:5 (https://doi.org/10.12688/gatesopenres.12786.2)
2. **Center for Evaluation of Value and Risk in Health (CEVR)**. *Global health cost effectiveness analysis registry*. Boston, MA: CEVR, Tufts Medical Center, 2019. Available from: http://healtheconomics.tuftsmedicalcenter.org/ghcearegistry/. Accessed by March 15, 2019.
3. **Dicker D, Nguyen G, Abate D, Abate KH, Abay SM, Abbafati C, et al.** Global, regional, and national age-sex-specific mortality and life expectancy, 1950–2017: A systematic analysis for the Global Burden of Disease Study 2017. *Lancet*. 2016;**392**:1684–735.
4. **National Institute for Health and Care Excellence (NICE) international**. Bill and Melinda Gates Foundation Methods for Economic Evaluation Project. 2014. Available from: https://www.nice.org.uk/Media/Default/About/what-we-do/NICE-International/projects/Gates-Reference-case-what-it-is-how-to-use-it.pdf. Accessed by March 15, 2019.
5. **Wilkinson T, Sculpher MJ, Claxton K, Revill P, Briggs A, Cairns JA**. The international decision support initiative reference case for economic evaluation: An aid to thought. *Value Health*. 2016;**19**:921–8.
6. **Chapman RH, Berger M, Weinstein MC, Weeks JC, Goldie S, Neumann PJ**. When does quality-adjusting life-years matter in cost-effectiveness analysis? *Health Econ*. 2004;**13**:429–36.
7. **Brouwer WB, Culyer AJ, van Exel NJ**. Welfarism vs. extra-welfarism. *J Health Econ*. 2008;**27**:325–38.
8. **Coast J, Smith RD, Lorgelly P**. Welfarism, extra-welfarism and capability: The spread of ideas in health economics. *Soc Sci Med*. 2008;**67**:1190–8.
9. **Gold MR, Stevenson D, Fryback DG**. HALYs and QALYs and DALYs, oh my: Similarities and differences in summary measures of population health. *Annu Rev Public Health*. 2002;**23**:115–34.
10. **Murray CJ, Acharya AK**. Understanding DALYs. *J Health Econ*. 1997;**16**:703–30.
11. **Salomon JA, Vos T, Hogan DR, Gagnon M, Naghavi M, Mokdad A, et al.** Common values in assessing health outcomes from disease and injury: Disability weights measurement study for the Global Burden of Disease Study 2010. *Lancet*. 2012;**380**:2129–43.
12. **Robberstad B**. QALYs vs DALYs vs LYs gained: What are the differences, and what difference do they make for health care priority setting? *Norsk Epidemiologi*. 2005;**15**. https://doi.org/10.5324/nje.v15i2.217
13. **Sassi F**. Calculating QALYs, comparing QALY and DALY calculations. *Health Policy Plan*. 2006;**21**:402–8.
14. **Airoldi M, Morton A**. Adjusting life for quality or disability: Stylistic difference or substantial dispute? *Health Econ*. 2009;**18**:1237–47.
15. **Diel R, Lampenius N**. Cost-effectiveness analysis of interventions for tuberculosis control: DALYs versus QALYs. *Pharmacoeconomics*. 2014;**32**:617–26.
16. **Augustovski F, Colantonio LD, Galante J, Bardach A, Caporale JE, Zárate V, et al.** Measuring the benefits of healthcare: DALYs and QALYs–does the choice of measure matter? A case study of two preventive interventions. *Int J Health Policy Manag*. 2018;**7**:120.
17. **Neumann PJ, Thorat T, Zhong Y, Anderson J, Farquhar M, Salem M, et al.** A systematic review of cost-effectiveness studies reporting cost-per-DALY averted. *PLoS One*. 2016;**11**:e0168512.
18. **Neumann PJ, Thorat T, Shi J, Saret CJ, Cohen JT**. The changing face of the cost-utility literature, 1990–2012. *Value Health*. 2015;**18**:271–7.
19. **Ryan M, Griffin S, Chitah B, Walker AS, Mulenga V, Kalolo D, et al.** The cost-effectiveness of cotrimoxazole prophylaxis in HIV-infected children in Zambia. *AIDS*. 2008;**22**:749–57.
20. **Mangen MJ, van Duynhoven YT, Vennema H, van Pelt W, Havelaar AH, de Melker HE**. Is it cost-effective to introduce rotavirus vaccination in the Dutch national immunization program? *Vaccine*. 2010;**28**:2624–35.
21. **Mihalopoulos C, Magnus A, Lal A, Dell L, Forbes D, Phelps A**. Is implementation of the 2013 Australian treatment guidelines for posttraumatic stress disorder cost-effective compared to current practice? A cost-utility analysis using QALYs and DALYs. *Aust N Z J Psychiatry*. 2015;**49**:360–76.
22. **Long MW, Gortmaker SL, Ward ZJ, Resch SC, Moodie ML, Sacks G, et al.** Cost effectiveness of a sugar-sweetened beverage excise tax in the US. *Am J Prev Med*. 2015;**49**:112–23.
23. **Diel R, Hittel N, Schaberg T**. Cost effectiveness of treating multi-drug resistant tuberculosis by adding Deltyba™ to background regimens in Germany. *Respir Med*. 202015;**109**:632–41.
24. **Wolfson LJ, Walker A, Hettle R, Lu X, Kambili C, Murungi A, Knerer G**. Cost-effectiveness of adding bedaquiline to drug regimens for the treatment of multidrug-resistant tuberculosis in the UK. *PLoS One*. 2015;**10**:e0120763.
25. **Nguyen HT, Hickson RI, Kompas T, Mercer GN, Lokuge KM**. Strengthening tuberculosis control overseas: Who benefits? *Value Health*. 2015;**18**:180–8.
26. **Nayagam S, Conteh L, Sicuri E, Shimakawa Y, Suso P, Tamba S, et al.** Cost-effectiveness of community-based screening and treatment for chronic hepatitis B in The Gambia: An economic modelling analysis. *Lancet Glob Health*. 2016;**4**:e568–78.
27. **Aggarwal R, Chen Q, Goel A, Seguy N, Pendse R, Ayer T, Chhatwal J**. Cost-effectiveness of hepatitis C treatment using generic direct-acting antivirals available in India. *PLoS One*. 2017;**12**:e0176503.
28. **Löfgren J, Matovu A, Wladis A, Ibingira C, Nordin P, Galiwango E, Forsberg BC**. Cost-effectiveness of groin hernia repair from a randomized clinical trial comparing commercial versus low-cost mesh in a low-income country. *Br J Surg*. 2017;**104**:695–703.
29. **Vetrini D, Kiire CA, Burgess PI, Harding SP, Kayange PC, Kalua K, et al.** Incremental cost-effectiveness of screening and laser treatment for diabetic retinopathy and macular edema in Malawi. *PLoS One*. 2018;**13**:e0190742.
30. **Voigt K, King NB**. Disability weights in the global burden of disease 2010 study: Two steps forward, one step back? *Bull World Health Organ*. 2014;**92**:226–8.
31. **Nord E**. Uncertainties about disability weights for the Global Burden of Disease study. *Lancet Glob Health*. 2015;**3**:e661–2.
32. **Kim DD, Bacon RL, Neumann PJ Chapter 6: Assessing transferability of economic evaluations: A decision framework. In: Culyer A et al.**, editors. *Non-communicable disease prevention: Best buys, wasted buys, and contestable buys*. Open Book Publishers; 2019;91–117.
33. **Anand S, Hanson K**. Disability-adjusted life years: A critical review. *J Health Econ*. 1997;**16**:685–702.
34. **Leech AA, Kim DD, Cohen JT, Neumann PJ**. Use and misuse of cost-effectiveness analysis thresholds in low- and middle-income countries: Trends in cost-per-DALY studies. *Value Health*. 2018;**21**:759–61.
35. **Arnold M, Griffin S, Ochalek J, Revill P, Walker S**. A one stop shop for cost-effectiveness evidence? Recommendations for improving disease control priorities. *Cost Eff Resour Alloc*. 2019;**17**:7. doi: 10.1186/s12962-019-0175-6. eCollection 2019

https://doi.org/10.1017/S0266462320000124 Published online by Cambridge University Press