Bias in published cost effectiveness studies: systematic review

Chaim M. Bell
University of Toronto

Let us know how access to this document benefits you.
Follow this and additional works at: https://escholarship.umassmed.edu/qhs_pp

Part of the Biostatistics Commons, Epidemiology Commons, and the Health Services Research Commons

Repository Citation

This material is brought to you by eScholarship@UMassChan. It has been accepted for inclusion in Population and Quantitative Health Sciences Publications by an authorized administrator of eScholarship@UMassChan. For more information, please contact Lisa.Palmer@umassmed.edu.
Bias in published cost effectiveness studies: systematic review
Chaim M Bell, David R Urbach, Joel G Ray, Ahmed Bayoumi, Allison B Rosen, Dan Greenberg, Peter J Neumann

Abstract
Objective To investigate if published studies tend to report favourable cost effectiveness ratios (below $20 000, $50 000, and $100 000 per quality adjusted life year (QALY) gained) and evaluate study characteristics associated with this phenomenon.

Design Systematic review.

Studies reviewed 494 English language studies measuring health effects in QALYs published up to December 2001 identified using Medline, HealthSTAR, CancerLit, Current Content, and EconLit databases.

Main outcome measures Incremental cost effectiveness ratios measured in dollars set to the year of publication.

Results Approximately half the reported incremental cost effectiveness ratios (712 of 1433) were below $20 000/QALY. Studies funded by industry were more likely to report cost effectiveness ratios below $20 000/QALY (adjusted odds ratio 2.1, 95% confidence interval 1.3 to 3.3), $50 000/QALY (3.2, 1.8 to 5.7), and $100 000/QALY (3.3, 1.6 to 6.8). Studies of higher methodological quality (adjusted odds ratio 0.58, 0.37 to 0.91) and those conducted in Europe (0.59, 0.33 to 1.1) and the United States (0.44, 0.26 to 0.76) rather than elsewhere were less likely to report ratios below $20 000/QALY.

Conclusion Most published analyses report favourable incremental cost effectiveness ratios. Studies funded by industry were more likely to report ratios below the three thresholds. Studies of higher methodological quality and those conducted in Europe and the US rather than elsewhere were less likely to report ratios below $20 000/QALY.

Introduction
Cost effectiveness analysis can help inform policy makers on better ways to allocate limited resources. Some form of cost effectiveness is now required for health interventions to be covered by many insurers. The quality adjusted life year (QALY) is used to compare the effectiveness of a wide range of interventions. Cost effectiveness analysis produces a numerical ratio—the incremental cost effectiveness ratio—in dollars per QALY. This ratio is used to express the difference in cost effectiveness between new diagnostic tests or treatments and current ones.

Interpreting the results of cost effectiveness analysis can be problematic, making it difficult to decide whether to adopt a diagnostic test or treatment. The threshold for adoption is thought to be somewhere between $20 000 ($11 300, $16 500)/QALY and $100 000/QALY, with thresholds of $50-60 000/QALY frequently proposed.

Regardless of the true value of the willingness of society to pay, studies of healthcare interventions would be expected to report a wide range of incremental cost effectiveness ratios. When published ratios cluster around a proposed threshold, bias may exist, and health policies based on their values may be flawed.

To describe the distribution of reported incremental cost effectiveness ratios and characteristics of studies associated with favourable ratios, we systematically reviewed cost effectiveness studies in health care that used QALYs as an outcome measure. We hypothesised that authors tend to report favourable incremental cost effectiveness ratios, such as those below $50 000 per QALY.

Methods
We conducted a systematic literature search of Medline, HealthSTAR, CancerLit, Current Contents Connect (all editions), and EconLit databases for all original cost effectiveness analyses published in English between 1976 and 2001 that expressed health outcomes in QALYs. Cost effectiveness analyses are reported as dollars per QALY. We used a standard data collection form, and two reviewers independently evaluated each study and abstracted the data. Disagreements were resolved by consensus. Details of the Tufts-NEMC CEA Registry (formerly the Harvard School of Public Health CEA Registry) are available online (http://tufts-nemc.org/cearegistry).

For each article, we documented the name of the journal, the year of publication, the disease category, and the country where the study was carried out. We used the Science Citation Index database to assign an impact factor for the year before publication to each journal. The sources of funding were identified as “industry” (partial or complete funding by a pharmaceutical or medical device company indicated in the manuscript) or “non-industry.” Studies for which a funding source was not listed were identified as “not specified.” We also assigned a quality score to each article, ranging from 1 (low) to 7 (high), based on the overall quality of the study methods, assumptions, and reporting practices.

Because cost effectiveness analyses often compare several programmes and include scenarios specific to patient subgroups or settings, each study may have contributed more than one cost effectiveness ratio. All cost effectiveness ratios were converted to US dollars at the exchange rate prevalent in the year of publication. Because we wanted to test whether the ratios targeted certain thresholds of the willingness of society to pay, such as $50 000/QALY, we did not adjust the ratios to constant dollars.

Statistical analysis
We analysed the distribution of all incremental cost effectiveness ratios and of the smallest and largest ratios from each study. We excluded nine ratios for which both the incremental cost and the incremental QALYs were negative. Although such interventions may be economically efficient, decision makers might not want to adopt interventions associated with reduced health.
Generalised estimating equations were used to evaluate study characteristics associated with incremental cost effectiveness ratios below the threshold values of $20 000, $50 000, and $100 000, as recommended previously. We used these equations because they take into account the correlation of cost effectiveness ratios derived from within the same study. We estimated odds ratios for associations between study characteristics and the presence of a favourable cost effectiveness ratio. Adjusted odds ratios were estimated by fitting a non-parsimonious model that included a priori predictor variables.

We used SAS statistical software version 8.2 for all analyses. Two sided P values less than 0.05 were considered significant.

Results

We screened more than 3300 study abstracts and identified 533 original cost-utility analyses. Thirty nine studies were excluded because they did not report numerical incremental cost effectiveness ratios. In total, 1433 cost effectiveness ratios were reported in these 494 studies, with a median of 2.0 (interquartile range 1-3) and a range of 1-20 ratios per study. Overall, 130 incremental cost effectiveness ratios (9%) were reported as cost saving (they saved money and improved health simultaneously), 124 (9%) were dominated by their comparators (had worse health outcomes and increased costs), and 1179 (82%) increased costs but improved health outcomes.

Most studies were published in the 1990s (table 1). The citation impact factor in the year before publication was available for 449 studies (91%). Cardiovascular and infectious disease interventions were the most commonly studied. Most studies were from the United States. About 18% were sponsored by industry, almost half were sponsored by non-industry sources, and sponsorship could not be determined in 34% of studies.

Figure 1 shows the frequency distribution of all 1433 incremental cost effectiveness ratios. The median (interquartile range) ratio per QALY was $20 133 ($4520-74 400). Approximately half of the ratios (712; 50%) were below $20 000/QALY, two thirds (974; 68%) were below $50 000/QALY, and more than three quarters (1129; 79%) were below $100 000/QALY. When analysed according to study sponsor, median (range) ratios per QALY were $13 083 ($3600-35 000) for those sponsored by industry and $27 400 ($4600-96 600) for those with non-industry sponsors. The median (range) cost effectiveness ratio per QALY for studies with unknown sponsorship was $18 900 ($4 960–64 300). Restricting the analysis to the lowest and highest ratios reported by each study yielded median ratios of $8784/QALY and $31 104/QALY (fig 2).

Several study characteristics were associated with reporting incremental cost effectiveness ratios below one or all three thresholds (table 2). The more quoted journals with a citation impact factor above 4 were less likely to publish ratios below $20 000/QALY (crude odds ratio 0.60, 95% confidence interval 0.42 to 0.86) or $50 000/QALY (crude 0.56, 0.38 to 0.82) than less quoted journals with a lower impact factor. However, this finding was not significant within the multivariable model (table 2).

Studies funded by industry were more likely to report cost effectiveness ratios less than $20 000/QALY (adjusted odds ratio 2.1, 1.3 to 3.3), $50 000/QALY (3.2, 1.8 to 5.7), or $100 000/QALY (3.3, 1.6 to 6.8) than studies funded by non-industry sources (table 2). Studies carried out in the US and Europe were significantly less likely to find favourable incremental cost effectiveness ratios than studies carried out elsewhere. Studies with quality scores for methodology above 5.5 were significantly less likely to report ratios below $20 000/QALY (0.48, 0.33 to 0.70) and $50 000/QALY (0.57, 0.39 to 0.83). Within the multivariable model, the association with quality remained significant only for cost effectiveness ratios below $20 000/QALY (adjusted odds ratio 0.58, 0.37 to 0.91; table 2).

Discussion

About half of all cost effectiveness studies published over a 25 year period reported highly favourable incremental cost

![Fig 1 Frequency distribution of 1433 incremental cost effectiveness ratios for health interventions](image-url)
Table 2 Characteristics of studies associated with favourable incremental cost effectiveness ratios according to three threshold values. Values are odds ratios (95% confidence intervals).

<table>
<thead>
<tr>
<th>Study characteristic</th>
<th>Crude OR (95% CI)</th>
<th>Adjusted OR (95% CI)*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Publication year</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1992-95</td>
<td>1.6 (0.98 to 2.7)</td>
<td>1.6 (0.96 to 2.7)</td>
</tr>
<tr>
<td>1996</td>
<td>1.3 (0.94 to 1.9)</td>
<td>1.3 (0.87 to 1.8)</td>
</tr>
<tr>
<td>1997-2001</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>Journal impact factor</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;2</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>2-4</td>
<td>0.62 (0.42 to 0.91)</td>
<td>0.62 (0.41 to 0.94)</td>
</tr>
<tr>
<td>&gt;4</td>
<td>0.60 (0.42 to 0.86)</td>
<td>0.60 (0.38 to 0.82)</td>
</tr>
<tr>
<td>Disease category</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cardiovascular</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>Endocrine</td>
<td>1.3 (0.68 to 2.4)</td>
<td>1.3 (0.58 to 3.0)</td>
</tr>
<tr>
<td>Infectious</td>
<td>1.1 (0.66 to 1.7)</td>
<td>1.1 (0.58 to 2.4)</td>
</tr>
<tr>
<td>Musculoskeletal</td>
<td>1.4 (0.60 to 3.3)</td>
<td>1.4 (0.50 to 3.7)</td>
</tr>
<tr>
<td>Nephrological/psychiatric</td>
<td>0.76 (0.40 to 1.5)</td>
<td>0.76 (0.31 to 1.4)</td>
</tr>
<tr>
<td>Other</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>Study funding source</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-industry</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>Industry</td>
<td>2.1 (1.4 to 3.4)</td>
<td>2.1 (1.3 to 3.5)</td>
</tr>
<tr>
<td>United States</td>
<td>0.35 (0.21 to 0.57)</td>
<td>0.35 (0.18 to 0.66)</td>
</tr>
<tr>
<td>Other</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>Methodological quality</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1-4</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>4-5-5</td>
<td>0.92 (0.64 to 1.3)</td>
<td>0.92 (0.62 to 1.5)</td>
</tr>
<tr>
<td>5-7-10</td>
<td>0.48 (0.33 to 0.70)</td>
<td>0.48 (0.32 to 1.3)</td>
</tr>
</tbody>
</table>

GALY: quality adjusted life year.
*Adjusted for all other study characteristics.
†Impact factor in the year before publication.
‡Funding by a pharmaceutical or medical device manufacturer.
§Canada 41 (59%), Australia 18 (26%), Japan 2 (3%), New Zealand 2 (3%), South Africa 2 (3%), other 6 (7%).
¶Mean score from two reviewers.
What is already known on this topic

Cost effectiveness analysis is widely used to inform policy makers about the efficient allocation of resources.

Various thresholds for cost effectiveness ratios have been proposed to identify good value, but the distribution of published ratios with respect to these thresholds has not been investigated.

What this study adds

Two thirds of published cost effectiveness ratios were below $50 000 per quality adjusted life year (QALY) and only 21% were above $100 000/QALY.

Published cost effectiveness analyses are of limited use in identifying health interventions that do not meet popular standards of “cost effectiveness”.

Conclusions

More rigour and openness is needed in the discipline of health economics before decision makers and the public can be confident that cost effectiveness analyses are conducted and published in an unbiased manner. These considerations are a prerequisite for these analyses to compare health management strategies. A heightened awareness of the limitations of cost effectiveness analyses and potentially influential factors may help users to interpret the conclusions of these analyses.

The paper was presented in abstract form at the Fifth International Congress on Peer Review and Biomedical Publication in Chicago, IL, 18–18 September 2005.

Contributors: All authors contributed to the conception and design of the project, revised the article critically for important intellectual content, and provided final approval of the final version. CMB, DRU, JGR, and AB analysed and interpreted the data and helped to draft the article. CMB had access to all of the data and takes responsibility for the integrity of the data and the accuracy of the data analysis. CMB and PJN are guarantors.

Funding: Agency for Health Care Research and Quality (RO1 HS10919). CMB and JGR are recipients of a phase 2 clinician scientist award and a new investigator award, both from the Canadian Institutes of Health Research. DRU holds a career scientist award from the Ontario Ministry of Health. Competing interests: None declared.

6 Lucey A, Ferry D, Densley A, Taggart PN. How does a new technology have to be to warrant adoption and utilization? Tentative guidelines for using clinical and economic evaluations. CMAJ 1999;160:475-81.


31 Sharp BW. What can and should be done to reduce publication bias? The perspective of an editor. JAMA 1999;280:1389-1.

32 Chalmers TC, Frank CS, Reiman M. Minimizing the three stages of publication bias. JAMA 1999;281:1392-5.


(Accepted 22 December 2005)

doi 10.1136/bmj.38737.607558.80

St Michael’s Hospital, Toronto, Ontario, Canada M5B 1W8
Chalmers TC, Frank CS, Reiman M. Minimizing the three stages of publication bias. JAMA 1999;281:1392-5.