

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. 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 \$50 000-60 000/QALY frequently proposed.

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, 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 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. Sources of funding were identified as “industry” (partial or complete funding by a pharmaceutical or medical device company indicated in the manuscript) “non-industry” and “not specified.” We also assigned a quality score to each article, ranging from 1 (low) to 7 (high).¹

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.

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. These equations 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

St Michael's Hospital, Toronto, Ontario, Canada M5B 1W8

Chaim M Bell
assistant professor of medicine and health, policy, management, and evaluation

Joel G Ray
assistant professor of medicine and health, policy, management, and evaluation

Ahmed Bayoumi
assistant professor of medicine and health, policy, management, and evaluation

University Health Network, Toronto
David R Urbach
assistant professor of medicine and health, policy, management, and evaluation

University of Michigan, Ann Arbor, MI, USA

Allison B Rosen
assistant professor of medicine

Health Systems Management, Ben-Gurion University of the Negev, Beersheba, Israel

Dan Greenberg
senior lecturer

Center for the Evaluation of Value and Risk in Health, Institute for Clinical Research and Health Policy Studies, Tufts University School of Medicine, Boston, USA

Peter J Neumann
director

Correspondence to: C M Bell, St Michael's Hospital, Toronto, Ontario, Canada M5B 1W8
bellc@smh.toronto.on.ca

BMJ 2006;332:699-701



This is the abridged version of an article that was posted on bmj.com on 22 February 2006: <http://bmj.com/cgi/doi/10.1136/bmj.38737.607558.80>

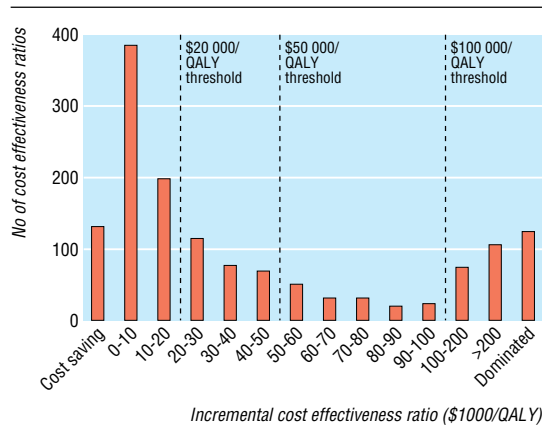


Fig 1 Frequency distribution of 1433 incremental cost effectiveness ratios for health interventions

by fitting a non-parsimonious model that included a priori predictor variables.

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.

See bmj.com for characteristics of the studies. Most were published in the 1990s. 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 sponsorship, median (range) ratios per QALY were \$13 083 (\$3600-33 000) 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. 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.

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. 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).

Discussion

About half of all cost effectiveness studies published over a 25 year period reported highly favourable incremental cost effectiveness ratios of less than \$20 000/QALY. More than half of the highest ratios reported by each study were below \$50 000/QALY. In multivariable analysis, location of the study, methodological quality, and sponsorship were associated with a favourable cost effectiveness ratio. Studies sponsored by industry were more than twice as likely as studies sponsored by non-industry sources to report ratios below \$20 000/QALY and over three times more likely to report ratios below \$50 000/QALY or \$100 000/QALY. Studies sponsored by industry were also more likely to be of lower methodological quality and to be published in journals with lower impact factors.

We restricted our analysis to studies that measured health outcomes with QALYs. It would be interesting to examine whether the use of alternative measures of health outcome results in more or less favourable cost effectiveness ratios. A limitation of our analysis is that some cost effectiveness analyses used in decision making may not have been published.² Analyses of cost effectiveness assessments submitted to the National

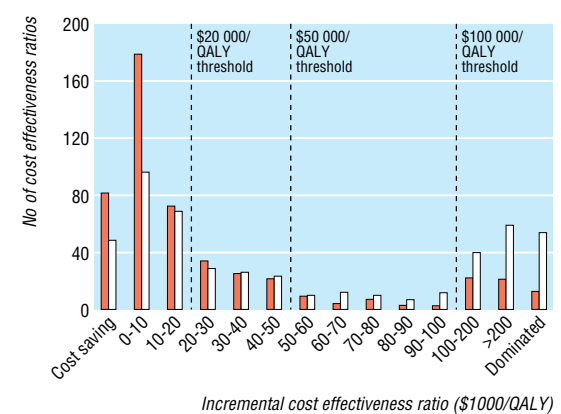


Fig 2 Frequency distribution of lowest (brown) and highest (white) incremental cost effectiveness ratios in each study

Institute for Health and Clinical Excellence (NICE) in the United Kingdom found that cost effectiveness ratios submitted by manufacturers were significantly lower than analyses of identical technologies performed by assessors from an academic centre.³

Publication bias

We found relatively few published incremental cost effectiveness ratios between \$50 000/QALY and \$100 000/QALY. There are three possible explanations. Firstly, they may reflect the true distribution of cost effectiveness ratios for healthcare interventions. Secondly, analysts may not be interested in studying interventions with mid-range cost effectiveness ratios or some journals may not want to publish such studies. Thirdly, some cost effectiveness analyses may be modelled to yield favourable ratios or studies with unfavourable ratios may be suppressed. Our results support concerns about the presence of significant and persistent bias in both the conduct and reporting of cost effectiveness analyses.⁴⁻⁵ It could be argued that all cost effectiveness analyses should be registered before they start, but this may be unrealistic given the way they are currently conducted.⁶

Recent attempts to standardise the conduct and reporting of economic analyses and modelling studies may help prevent the manipulation of studies.⁷⁻¹⁰ Electronic publishing could enhance transparency in modelling by making technical appendices available. Furthermore, distribution of the underlying decision analysis models to the public should be considered.

Journal editors and reviewers can help reduce publication bias. Potential conflicts of interest of study sponsors and authors need to be scrutinised. Journal editors may show bias by publishing studies with positive results but not studies with negative results, although this may not be common. However, differences between economic analyses may also reflect a more fundamental difference in the studies.¹¹

Conclusions

More rigour and openness is needed before decision makers and the public can be confident that cost effectiveness analyses are conducted and published in an unbiased manner.

The paper was presented in abstract form at the Fifth International Congress on Peer Review and Biomedical Publication in Chicago, IL, 16-18 September 2005.

Contributors: See bmj.com.

Funded by a grant from the 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.

- 1 Neumann PJ, Stone PW, Chapman RH, Sandberg EA, Bell CM. The quality of reporting in published cost-utility analyses, 1976-1997. *Ann Intern Med* 2000;132:964-72.
- 2 Hill SR, Mitchell AS, Henry DA. Problems with the interpretation of pharmacoeconomic analyses: a review of submissions to the Australian pharmaceutical benefits scheme. *JAMA* 2000;283:2116-21.
- 3 Miners AH, Garau M, Fidan D, Fischer AJ. Comparing estimates of cost effectiveness submitted to the National Institute for Clinical Excellence (NICE) by different organisations: retrospective study. *BMJ* 2005;330:65.
- 4 Freemantle N, Mason J. Publication bias in clinical trials and economic analyses. *Pharmacoeconomics* 1997;12:10-6.
- 5 Hillman AL, Eisenberg JM, Pauly MV, Bloom BS, Glick H, Kinoshita B, et al. Avoiding bias in the conduct and reporting of cost-effectiveness research sponsored by pharmaceutical companies. *N Engl J Med* 1991;324:1362-5.

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"

- 6 Rennie D. Peer review in Prague. *JAMA* 1998;280:214-5.
- 7 Canadian Coordinating Office for Health Technology Assessment. *Common drug review submission guidelines*. Ottawa: Canadian Coordinating Office for Health Technology Assessment (CCOHTA), 2005. www.ccohta.ca/CDR/cdr_pdf/cdr_submiss_guide.pdf (accessed 9 Feb 2006).
- 8 Sculpher M, Fenwick E, Claxton K. Assessing quality in decision analytic cost-effectiveness models. A suggested framework and example of application. *Pharmacoeconomics* 2000;17:461-77.
- 9 Philips Z, Ginnelly L, Sculpher M, Claxton K, Golder S, Riemsma R, et al. Review of guidelines for good practice in decision-analytic modelling in health technology assessment. *Health Technol Assess* 2004;8:1-172.
- 10 Weinstein MC, O'Brien B, Hornberger J, Jackson J, Johannesson M, McCabe C, et al. Principles of good practice for decision analytic modeling in health-care evaluation: report of the ISPOR task force on good research practices—modeling studies. *Value Health* 2003;6:9-17.
- 11 Baker CB, Johnsrud MT, Crismon ML, Rosenheck RA, Woods SW. Quantitative analysis of sponsorship bias in economic studies of antidepressants. *Br J Psychiatry* 2003;183:498-506.

(Accepted 22 December 2005)

doi 10.1136/bmj.38737.607558.80

Corrections and clarifications

Hanging in the balance

This news article by Rebecca Coombes gave the wrong age group for the bowel cancer screening programme in Scotland, due to be rolled out in 2007 (*BMJ* 2006;332:384, 18 Feb). The programme will target people aged 50-74 years (not 50-69, as we stated).

Obituary: Sir John Peel

Two errors occurred in this obituary by Caroline Richmond (*BMJ* 2006;332:366, 11 Feb). Firstly, Sir John Peel was born in Surbiton, not Bradford. Secondly, we were wrong to state that his third wife, Sally Barton, was a widow when she married him.

Optimising prenatal diagnosis of Down's syndrome

A last minute change made by editorial staff in this editorial by James P Neilson and Zarko Alfrevic resulted in an error in describing the detection of Down's syndrome (*BMJ* 2006;332:433-4, 25 Feb). The statement in the fifth paragraph that refers to full karyotyping "picking up more truly positive cases of Down's syndrome" is wrong. In fact, qf-PCR (quantitative fluorescent polymerase chain reaction) will detect all true cases of Down's syndrome.

Please do not resuscitate: Automatic refusal is as harmful as offering resuscitation to all

The authors of this letter, Carmelo Aquilina and colleagues (*BMJ* 2006;332:608-9, 11 Mar), have asked us to clarify that the penultimate author is Joyce (not Catherine) Tarrant.