



**ISPOR GOOD PRACTICES FOR QUALITY IMPROVEMENT  
OF COST-EFFECTIVENESS RESEARCH TASK FORCE REPORT**

*DRAFT*

William McGhan PharmD PhD<sup>1</sup>, Maiwen Al PhD<sup>2</sup>, Jalpa Doshi PhD<sup>3</sup>, Isao Kamae MD, DrPH<sup>4</sup>,  
Steven E. Marx PharmD, MS<sup>5</sup>, Donna Rindress PhD<sup>6</sup>

1. University of the Sciences, Philadelphia, PA, USA
2. Institute for Medical Technology Assessment (iMTA), Rotterdam, Netherlands
3. University of Pennsylvania, Philadelphia, PA, USA
4. Kobe University, Kobe, Japan
5. Abbott Laboratories, Chicago, IL, USA
6. BioMedCom Consultants Inc., Montreal, Quebec, Canada

**The sections of the QICER report are as follows:**

- Section 1** Introduction and Overview of Issues
- Section 2** Guidelines Around the Globe - What countries have guidelines and how well are they working?
- Section 3** Statistics and Science - Statistical issues in cost-effectiveness research & how can we improve?
- Section 4** Publication Quality - Guidelines: do they exist, are they used, and how can they improve?
- Section 5** Decision-makers, Practitioners and EBM - Barriers to use and application of cost effectiveness research.
- Section 6** Recommendations
- References**



**Section 1: INTRODUCTION AND OVERVIEW OF ISSUES**

***Introduction***

Quality assessment and continuous quality improvement has long been recognized as a vital process in all societal systems and organizations. In healthcare, critical reviews of treatments and quality reports on outcomes can help correct deficiencies and further advance efficiency and quality. Continuous quality improvement is integral to our global efforts to improve the economics and quality of life in all healthcare sectors and all patient populations. Healthcare researchers and practitioners are certainly aware of regional and local healthcare expenditures and ongoing publication of major health economic studies. But there is an important role for ISPOR in taking a macro review and examination of quality and trends in pharmacoeconomics, healthcare economics research and their resulting impact on global policies and practice.

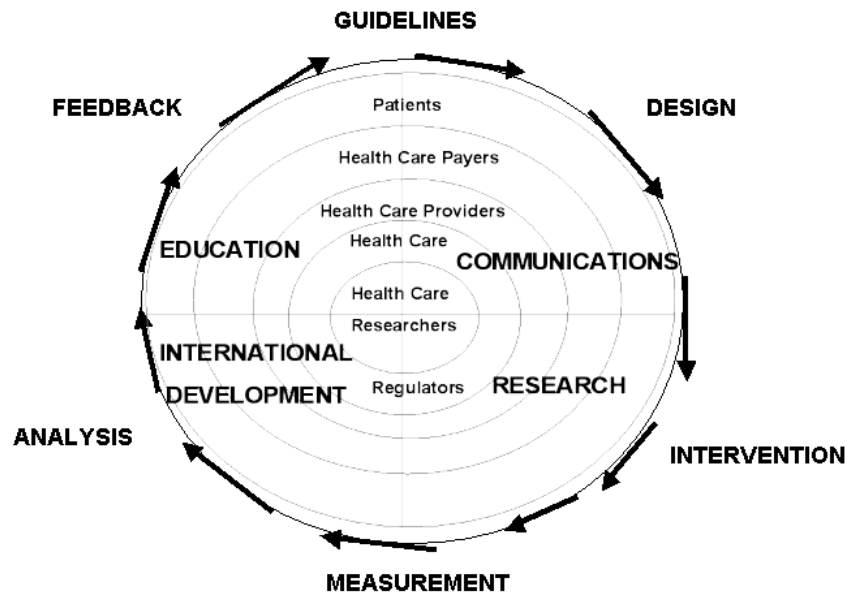
Methods and statistics used in pharmacoeconomics and health economic studies need to be periodically critiqued to improve the quality and usefulness of publications, reports, and practice guidelines. It is also essential to monitor the quality of healthcare economic study guidelines being used by journals, organizations and governments around the world for assessing new therapies and allocation of resources.

***Mission***

The mission of the ISPOR task force on Quality Improvement of Cost-effectiveness Research (QICER) is to generate periodic quality reports and make recommendations to facilitate the improvement of pharmacoeconomics and health outcomes research and their use in stimulating better healthcare and policy. These goals will be accomplished through periodic systematic reviews and surveys. The results and findings will be made available on the ISPOR website for membership comment and, after revision, published as white papers and reports that include recommendations for future ISPOR projects, educational programs, and member services.

In its first report, the focus of the task force has been limited to cost-effectiveness research. Broader issues in HEOR such as patient reported outcomes, health related quality of life, softwares, etc., have been considered outside the scope of this report, but are envisioned as fodder for future work.

**ISPOR VISION AND CONTINUOUS QUALITY IMPROVEMENT**



**Section 2: GUIDELINES AROUND THE GLOBE**

We felt it important to look at the role of guidelines in promoting the quality and improvement of health economic and health outcomes research (HEOR). It is usually assumed that the presence of guidelines leads to quality improvement, hypothesizing that established guidelines increase credibility and usefulness by defining generally accepted standards and the requirements of specific users. However, in this field, there is not a lot of evidence to support or disprove this assumption. A number of studies have evaluated the quality of research, but few have examined the relationship between the presence of guidelines and the quality of studies. We felt that two topics should be examined: the availability of HEOR guidelines and the impact of guidelines on the quality of HEOR.

A few authors have reviewed available guidelines, comparing and contrasting them<sup>1-3</sup> and there are other resources available as well ([www.ISPOR.org](http://www.ISPOR.org); [www.biomedcom.org/en/resources-BMC-databases.html](http://www.biomedcom.org/en/resources-BMC-databases.html)). All guidances are similar across the board, but there are some significant differences, generally due to their intended purpose, the audience to be addressed, regional, cultural or political variation, or author or sponsor preferences.

How can the impact of guidelines on HEOR quality be measured? Most journals do not have specific guidelines or requirements for HEOR (Section 4), so even though published articles are easily accessible, their quality, and the improvement of their quality are not easily linked to specific guidelines. Those who have measured quality of HEOR publications<sup>1,4-8</sup> have designed their own quality measures from among those guidelines currently available and generally accepted standards.

Several formulary evaluation bodies (such as NICE in the UK<sup>9</sup>, CADTH in Canada<sup>10</sup>, and PBS in Australia<sup>11</sup>) have developed specific guidances and requirements for HEOR studies submitted, but these studies are often not publicly available for evaluation. In a few cases, some of these bodies have performed, or allowed, evaluation of the studies

submitted to them and these have been presented publicly<sup>12</sup> or published<sup>13,14</sup> but these are often small sample size, qualitative and importantly, not easily comparable across jurisdictions.

### ***The evolution of guidelines***

Guideline development began in Australia in 1992, followed closely by Canada and a few academic groups in the US<sup>15-17</sup>. Over the last decade, many countries have produced their own guidances and others are in development. There are currently about 32 HE guidances from 28 countries (multiples from the US and Canada). Of these, 15 were produced by government bodies, 6 were developed by academic groups, 1 was produced by a healthcare insurer, and 10 were collaborations with various permutations between government agencies, academic groups and insurers. Seventeen of these were prepared to be part of formulary submission guidances or requirements. Table 1 summarizes those HEOR guidelines of which we are currently aware.

### ***Measuring and improving quality of guidelines***

No instrument has been found that permits quantitative measurement of guideline quality, or comparison of guidelines. What publicly available information there is has been rather qualitative; for example, in the early days of guideline development, Regence BlueShield found that guidelines were practical from a logistics perspective, improving the relevance and timeliness of information available for decisionmakers<sup>18</sup>.

Although guidelines have not been rigorously evaluated, a number of studies have been published looking at the quality of studies submitted to the guideline-producing bodies. In one of these, Hill et al evaluated 326 Australian submissions and found significant issues in interpretation of the guidelines and in conduct of studies.<sup>14</sup> In another Colmenero et al analyzed 53 economic submissions to US managed care organizations and found low levels of compliance with accepted standards of HEOR conduct.<sup>19</sup> No published studies however appear to have measured the relationship between guidelines and the

studies submitted. And among published papers evaluating the quality of submitted studies, measures of quality are not easily comparable.

If we are to assess the impact of guidelines on quality and improvement of quality of HEOR studies, a tool is required that is quantifiable, anchored to guidelines, and to generally accepted practices, and comparable across guidelines, studies and time. An instrument may already be available to fit this purpose, or a new one might be developed that incorporates the most relevant aspects of existing tools. Some already in use that might be adapted include: (a) Neumann et al measured the quality of economic analyses in several studies over the last decade;<sup>5,6</sup> (b) Chiou et al developed a grading system to measure the value and quality of HE analyses using the QHES instrument;<sup>1</sup> (c) Goetghebeur et al quantifiably assessed the quality of ten submissions to the Canadian Common Drug Review linking CADTH requirements to the studies submitted with quality of submitted dossiers.<sup>20,21</sup>

Table 1. Countries (regions or groups) that have developed health economic and outcomes research guidances	Source
Australia	11
The Baltics (Latvia, Estonia & Lithuania)	22
Austria	23
Belgium	24
Brazil	25
Canada (CADTH, Ontario)	26
China	27
Finland	28
France	29
Germany	30
Hungary	31
Ireland	32
Israel	33
Italy	34
The Netherlands	35
New Zealand	36
Norway	37
Poland	38,39
Portugal	40
Russia	41
Scotland	42
Slovak Republic	43
South Korea	44
Spain	45
Sweden	46
Switzerland	47
UK (England and Wales)	48
USA (task force, Gold Panel, AMCP, WellPoint)	49-52

***Future work***

Based on this preliminary review, there are three promising directions in which we could proceed with regard to guidelines as instruments in measuring and improving quality in HEOR:

1. Perform a formal evaluation of currently available instruments that might be used to quantify and compare the value of HEOR guidelines. If none are found that meet this need, develop, or promote development of one.
2. Once an instrument is selected or developed to assess guideline quality, promote harmonization of guidelines, allowing for differences in regional needs and politics.
3. Concurrently, evaluate available instruments or promote development of one to quantify the impact of HEOR guidelines on the quality of HEOR studies. The outcomes of such an instrument should allow comparison across all guidelines as well as comparison over time, to allow rigorous longitudinal evaluation of quality improvement.



**SECTION 3. STATISTICS AND SCIENCE**

***Introduction***

This section discusses the statistical issues in cost-effectiveness research, preferred approaches to address these issues, and recommendations for improving the statistical methods applied in the cost-effectiveness literatures. For ease of presentation, we discuss the statistical issues and their solutions separately for clinical trial based economic evaluations and decision modeling based studies.

***What are the statistical problems in cost-effectiveness research?***

***Clinical trial based economic evaluations***

A critical source of the evidence on costs and cost-effectiveness of new medical treatments comes from analyses of economic information collected prospectively in randomized clinical trials alongside clinical end points. The number of clinical trial-based economic evaluations has increased considerably over the last decade. In the same time frame, the field has matured substantially, including advancement and growing consensus about appropriate statistical methods for analysis of costs and cost-effectiveness alongside clinical trials.<sup>53</sup> Systematic reviews suggest that published studies on clinical trial based economic evaluations have begun to use some of these new statistical techniques.<sup>54,55</sup> Nevertheless, there are still a substantial number of studies using statistical methods of poor quality. In addition, there remain areas needing further research.

- a) Joint comparison of costs and effects and estimation of sampling uncertainty

A joint comparison of costs and effects using the incremental cost-effectiveness ratio (ICER) or the incremental net monetary (health) benefit (INB) is a useful decision tool to help determine whether the new therapy offers good value

relative to the alternative. The use of this tool is particularly important when there is a trade-off between costs and effects; that is, one therapy is both significantly more effective and more costly compared with the other therapy. If there is no trade-off between costs and effects, that is, when one therapy is significantly more effective and less costly when compared with the other therapy, this decision tool may not be necessary because that therapy is unambiguously dominant over its alternative. A third possibility occurs when the two treatments have the same effect. In this case, some authors have interpreted textbooks and guidelines on health economic evaluations to suggest that a cost minimization approach is sufficient (i.e., the lowest cost treatment is the treatment of choice) and there is no need to perform a joint comparison of costs and effects.<sup>56-59</sup> Nevertheless, as our understanding of sampling uncertainty for the comparison of costs and effects has grown, the cases where this interpretation is appropriate have shrunk. Because cost-effectiveness ratios and net monetary benefit estimated from trial data are the result of samples drawn from the population, one should report the uncertainty in this outcome that derives from such sampling.<sup>60</sup> Identification of methods such as confidence intervals for cost-effectiveness ratios<sup>61-64</sup> acceptability curves<sup>65</sup> and confidence intervals for net monetary benefit<sup>66</sup> for the measurement of this uncertainty have been important methodologic developments in the economic evaluation of medical therapies.<sup>67</sup> When one uses these methods, a finding of significantly lower cost and an indistinguishable clinical outcome need not guarantee that confidence that the significantly less expensive therapy is good value. As a result of uncertainty, the cost-minimization approach has been shown to be rarely appropriate as a method of analysis and the need for a joint comparison still remains under most circumstances.<sup>68</sup> Alternatively, because it is possible to have more confidence in the combined outcome of differences in costs and effects than in either outcome alone, observing no significant difference in costs and effects need not rule out that one can be confident that one of the two therapies is good value. In these cases, one should compare costs and effects, and one should report on their sampling uncertainty.

b) Analysis of cost data

For economic analysis, costs and cost differences between treatment groups should be expressed by the use of the arithmetic mean, and not medians, because this summary measure permits a budgetary assessment of treatment ( $N \times$  arithmetic mean = total cost) and is the statistic of interest for healthcare policy decisions.<sup>53</sup> Because of the often highly

skewed distribution of cost data, the normality assumption underlying the parametric *t*-test is often called into question and standard nonparametric tests (e.g., Mann–Whitney *U*-test or Wilcoxon rank sum test), or parametric tests on normalizing transformations (e.g., log transformation) are often used as a substitute. Yet these popular alternatives are not appropriate for drawing statistical inferences on differences in arithmetic mean costs.<sup>69-71</sup> For example, when one uses a *t*-test to evaluate the log of costs, the resulting *P*-value has direct applicability to the difference in the log of costs and to the difference in the geometric mean of costs. It may or may not be directly applicable to the arithmetic mean costs. Similarly, when one uses a Mann–Whitney *U*-test, one is testing differences in the median of costs. Thus, statistical inferences about these other statistics may not be representative of inferences about the differences in arithmetic mean, which is the statistic of interest. If one does not want to adopt a parametric *t*-test to directly test for differences in arithmetic mean costs, one can compare the arithmetic means by using a nonparametric bootstrap. This procedure has the added advantage of avoiding a parametric assumption about the distribution of costs. As a result, the nonparametric bootstrap has increasingly been recommended either as a check on the robustness of standard parametric *t*-tests, or as the primary statistical test for making inferences about arithmetic means for moderately sized samples of highly skewed cost data.<sup>71-73</sup>

Many clinical trial-based economic evaluations are limited to univariate analyses of costs. Even if treatment is assigned in a randomized setting, there are advantages to using multivariable techniques to analyze costs. Multivariable analysis of costs may be superior to univariate analysis because it improves the power for tests of differences between groups (by explaining variation due to other causes). It also facilitates subgroup analyses for cost-effectiveness, for example, more and less severe; different countries/centers, etc. Finally, it accounts for potentially large and influential variations in economic conditions and practice patterns by provider, center, or country that may not be balanced by randomization. Adoption of multivariable analysis does not, however, avoid the issues that arise in the univariate analysis of cost. For example, regressions on the logarithmic transformation of costs were previously considered an ideal remedy to the violation of the assumption of normally distributed error term that underlies ordinary least squares (OLS) regression. Nevertheless, as the shortcomings of multiple regression models of log transformed costs became more widely publicized<sup>70</sup>, the use of the generalized linear models have become the more acceptable alternative.<sup>74-76</sup>

c) Handling of censored cost data

Incomplete or censored cost data occur in most randomized trials that follow participants for clinically meaningful lengths of time; yet they are often not addressed in the analysis.. Whether cost data were incomplete, the amount of incomplete data and the statistical method adopted to address the problems posed by incomplete data should routinely be reported in trial-based analyses<sup>53</sup> The ISPOR RCT-CEA taskforce recommended that, “ignoring small amounts of missing data is acceptable if a reasonable case can be made that doing so is unlikely to bias treatment group comparisons.<sup>53</sup> However, no clear guidance exists for how much censoring is too much. Many studies in the literature have adopted naïve approaches wherein censored observations are either excluded from analysis (i.e., complete-case analysis) or included as though they were complete observations,(i.e., full sample analysis). In the first naïve approach, only the uncensored cases are used in the estimation of mean cost and this method is biased toward the costs of the patients with shorter survival times because patients with larger survival times are more likely to be censored.<sup>77,78</sup> Also completely discarding patients with censored data can lead to the loss of information and statistical power, which can be problematic if the percentage of censored cases is high. The second naïve approach which uses all cases without differentiating between censored and uncensored observations is always biased downward because the costs incurred after censoring times are not accounted for.<sup>78</sup> Although there exists a mix of approaches to impute the cost data, recent statistical interest in addressing censored cost data has led to the proposal of several methods of estimation that explicitly account for incomplete cost data due to loss-to-follow-up.<sup>77,79-87</sup> It is well-established that these methods are prone to less bias and return a better estimate of sampling variance than other naïve estimation methods.<sup>77,78,80,82,88,89</sup>

d) Sample size and power

Prior to the development of methods for assessing sampling uncertainty for the joint comparison of cost and effect, health economists commonly attempted to estimate sample size based on the larger of the sample sizes needed for estimating pre-specified cost and effect differences – i.e., what sample size was required to identify a \$1000 difference in costs, and what was required to identify a 10% reduction in mortality. With the development of methods for assessing uncertainty,

sample size calculations should now be based on the sample size needed to rule out that the net monetary benefits of the intervention are less than 0.<sup>90-93</sup> Often economic evaluations are piggy backed on clinical trials with a prespecified sample size. In such instances, researchers should estimate and report the power available to rule out cost-effectiveness ratios that exceed the maximum willingness to pay.

e) Evaluating transferability (generalizability) of trial results

Multinational clinical trials are the norm for the evaluation of new medical therapies. However, the presence of between-country heterogeneity in trials has led to a growing concern that the pooled or average economic results from multinational trials may not be reflective of the results that would be observed in individual countries that participated in the trial.<sup>94</sup> Common sources for concern about the representativeness of data from multinational trials include transnational differences in morbidity/mortality patterns; practice patterns (i.e., medical service use); and absolute and relative prices for medical service use (i.e., price weights). The use of trial-wide clinical results, trial-wide medical service use, and price weights from a single country has been one of the commonly proposed, potentially inadequate solutions to the problem of transferability (e.g., to tailor the results to the UK, simply use UK price weights, and conduct the analysis as if all participants were treated in the UK). A second potentially inadequate solution has been to use trial-wide clinical results, and country-specific medical service use and price weights. Both approaches have the failing that they ignore the fact that clinical and economic outcomes may influence one another. That is, differences in cost may affect practice patterns, which in turn may affect outcome; differences in practice pattern may affect outcome, which in turn may affect cost. The ISPOR Good Research Practices task force on Economic Data Transferability has recently recommended good research practices for dealing with aspects of transferability including three proposed statistical methods that use patient-level data to address transferability: detection of heterogeneity<sup>95,96</sup>, fixed effects models<sup>97, 98</sup>, and multilevel, or hierarchical models.<sup>99-106</sup>

*Decision Models*

The estimation of the full economic effects of health technologies generally requires the extrapolation of clinical trial evidence beyond the period of follow-up through the use of decision modeling techniques to synthesize data from various sources. The aim of the modeling study is to aid decisionmakers in making decisions under uncertainty. Obviously, the results of modeling studies will only be helpful to decisionmakers if the study is performed according to current standards. While the quality of cost-utility analyses has improved over time, still current studies do not address all issues appropriately.<sup>5</sup> This makes it clear that guidance is important for those performing modeling studies. In several countries, authorities have formulated guidelines (e.g., NICE 2004, CVZ 2006) and ISPOR has also published guidelines through the task force on Good Research Practice in Modeling Studies.<sup>107</sup> We discuss here some of the most important issues where quality might be improved as well as some new methodological topics that have emerged in the past few years.

a) Methods for evidence synthesis

The ISPOR task force on Good Research Practice in Modeling Studies suggested in their report that systematic review should be conducted on key model inputs. There are various ways of synthesizing the evidence found in various studies (e.g., fixed or random effects meta-analysis, either frequentist or Bayesian), but there is not one optimal method of synthesizing data currently available,<sup>108</sup> and the typical meta-analyses cannot straightforwardly be applied to synthesize data for cost-effectiveness models.

One reason is that meta-analysis has been developed to combine quantitative results of several similar studies into a pooled estimate of the treatment effect (e.g., odds ratio, relative risk, difference in change from baseline). It uses the magnitude of the treatment effect and its uncertainty from each individual study to produce a weighted mean of the treatment effect.<sup>109,110</sup> However, in modeling studies, the parameter to be estimated is not only a treatment effect like the odds ratio of having an event. Typically, models contain parameters like transition probabilities between disease states, event probabilities, rate ratios of treatment effects, quality of life or utility values, and costs. These parameters have

different distributions which need to be combined. Moreover, the placebo comparator needs to be modeled too, meaning that we are dealing with more heterogeneity than usually remains after the variance in treatment effect has been corrected for the variance in placebo-effect.

Second, meta-analyses have traditionally been performed on studies that compare the same intervention with the same comparator. However, comprehensive decision analytic models aim to identify the most cost-effective treatment among the entire spectrum of all relevant treatment options. This issue may be dealt with through so-called mixed treatment comparisons, which combine multiple different pairwise comparisons across a range of different interventions.<sup>111-114</sup> In mixed treatment comparisons, the relative effect of a treatment compared to a range of alternatives is estimated by including indirect comparisons of two interventions through a common comparator. Evidence from direct and indirect comparisons are analyzed simultaneously, which allows estimating treatment effects in the absence of head-to-head comparisons. Such mixed treatment comparisons are inevitable in modeling cost-effectiveness.

The fact that the choice between fixed and random effects model and between a Bayesian and a frequentist approach can have a large impact on the outcome of the model underlines the need for complete transparency in the reporting of a modeling study.

b) Probabilistic sensitivity analysis

Sensitivity analysis should always be an integral part of a modeling study: input parameters of the model are varied to see if and how the outcome changes.

Until recently, the common type of sensitivity analysis was a deterministic one-way or multi-way analysis, in which one or more (usually not more than three) parameters are varied between certain limits. However, such analysis does not pay attention to the fact that the limits are often less likely to occur than the baseline estimate. Consequently, in the last few years more and more studies include a probabilistic sensitivity analysis, in which all input variables are varied

simultaneously, according to probability distributions.<sup>115</sup> Such an analysis presents information on all possible outcomes, as well as on the likelihood of these outcomes.

In their report, the ISPOR task force on Good Research Practice in Modeling Studies stresses the need to include a sensitivity analysis as part of the modelling study. While those task force guidelines mention that deterministic and probabilistic analyses are equally appropriate, other guidelines such as the UK and Dutch guidelines strongly prefer a probabilistic sensitivity analysis as a way to correctly represent parameter uncertainty. However, there are still good reasons to also include deterministic sensitivity analysis in a modeling study, e.g., to account for other types of uncertainty such as uncertainty relating to the structure and assumption of decision models.<sup>116</sup> Series of sensitivity analyses may be done to look at the consequence of changing different assumptions and scenarios in a model.

c) Value-of-information

In the last few years much attention has been given to so-called Value-of-Information (VOI) analysis.<sup>117,118</sup> This type of analysis addresses the question what the value is of collecting additional information to eliminate or reduce uncertainty, since making the wrong decision comes with a cost that is equal to the benefits forgone due to the wrong decision. These expected costs of uncertainty can be determined by 1) the probability that a decision based on the current ICER is wrong and 2) the size of the opportunity loss if the wrong decision is made.

The first step in a VOI analysis is the estimation of the expected value of perfect information (EVPI), which is the amount the decisionmaker should be willing to pay to eliminate all uncertainty in the decision. The next step is to calculate the expected value of partial perfect information (EVPPI), which is simply the EVPI on one parameter or subset of parameters.<sup>119</sup> Based on the latter analyses, priorities may be set for further research. The final part of the VOI is to calculate the expected value of sampling information (EVSI), which is the amount the decisionmaker should be willing to pay to reduce uncertainty through a sample of a certain size and to set this against the costs of obtaining that sample.<sup>120</sup>

The ISPOR task force on Good Research Practice in Modeling Studies actually suggested that “The decision to obtain additional data to inform a model should be based on a balance between the expected value of the additional information and the cost of the information.”

d) Model validation

The final part of any model development should concern validation. Several types of validation may be distinguished: internal validation, between-model validation, predictive (or prospective) validation, and external validation.<sup>107,121</sup>

Internal validation concerns the comparison of model outputs with data used in the model development. While this type of validation is straightforward if the model is based on one source of data, it becomes more complicated if the model is based on a synthesis from various sources of data. It is possible that no model input can be chosen so that the model validates well against each separate data source. However, a good effort should at least be made to describe the deviations from the data sources and possible explanations.

Between-model validation involves the comparison of the current model and published or publicly available models. If the results differ, an attempt should be made to clarify whether these discrepancies are due to difference in model structure or model input.

Predictive validation aims at comparing model results to newly available data from the same data source that was used as model input. On the other hand, external validation concerns the comparison of model results to data from studies not used in the model development. Not always are these types of validation possible, if a model contains all data currently available, there is no data source for external validation. As the ISPOR task force on Good Research Practice in Modeling Studies remarks: “...it is not necessary that every data estimate or structural assumption be tested in prospective studies, in advance of model use”. However, that task force also stresses that models should never be regarded as immutable. They should be updated and possibly abandoned as new evidence becomes available to inform their structure or input values. If

models are inconsistent with the new evidence but have not been amended to calibrate against this evidence, the model should be abandoned until such recalibration has been finished.

***How can we make the science better?***

ISPOR has published best practices documents for the design, conduct, and reporting of economic analyses alongside clinical trials as well as decision modeling studies.<sup>53,107</sup> Whether explicit guidelines alone will foster improvements in the quality of future studies remains a question, given previous research that suggests that such guidelines have had minimal or slow impact in improving the quality of subsequent studies.<sup>3,59,122,123</sup>

Part of this problem may be that most of the advances in the statistical techniques for analyzing cost data have been published in highly technical economic or biostatistics journals. Although some applied researchers may not be reading such literature, many may have difficulty understanding the rationale for and implementation of these technical methods. There is a clear need for publications in more applied journals that focus on explaining these technical advances in an easily understandable format to conduct knowledge transfer to researchers who need to be applying these newer methods.

Another problem when trying to raise the quality of studies is that of space. Transparency of a model is of utmost importance if we want modeling studies to be taken seriously, not only within the health economics community but also for instance among clinicians. To achieve full transparency, the actual model should be available in electronic format, since it will often not be possible to describe every detail of the model sufficiently clearly in a paper due to space restrictions.

Additional efforts to improve the quality of future studies may involve peer reviewers for both funding agencies and journals being critical of studies that fail to apply best practices in cost-effectiveness research. For example, peer reviewers might be provided with a clear checklist incorporating all requirements. Thus, all studies are reviewed in a similar way, hopefully leading to increased quality. Regulatory and reimbursement authorities should also explicitly adopt

best practices guidelines and hold economic data submissions to these high standards while making reimbursement decisions.



**SECTION 4: JOURNALS AND PUBLICATION QUALITY**

Journal publication can play a critical role in quality improvement of HEOR research. This can be by establishing requirements and guidelines for the conduct and reporting of the various types of studies that comprise HEOR, through the peer-review process, by dissemination of studies, by peer feedback, and as an ongoing learning process for researchers. Although published work may include abstracts, posters and podium presentations, newsletters and other non peer-reviewed publications (such as educational texts, patient information and marketing materials), for the purposes of this first report the focus has been on peer-reviewed journal publications because these are most accessible and easiest to track.

There are a great many journals globally. To determine how many of these routinely accept and publish articles relevant to HEOR, and how many of these provide or require guidelines for the conduct and reporting of HEOR studies, a survey of the World Association of Medical Editors (WAME; [www.wame.org](http://www.wame.org)) was carried out. This organization represents over 965 biomedical journals, from over 91 countries, and all geographic regions of the world. As such WAME was an ideal source of information relevant to quality improvement in HEOR publication. In the survey which all WAME members were invited to participate through the WAME listserv, we asked about journal type, location, scope, circulation, whether they accepted HEOR articles and which types of studies, and how they found reviewers for HEOR. We also asked whether guidelines were recommended or mandatory for authors or reviewers, and if so, which ones.

Of the 965 journals represented in WAME, 55 (6%) responded to the survey. These came from 29 countries and all continents, with 45% representation for North America and Europe. Almost all (98%) were peer reviewed, and the majority (72%) international in readership. Most respondents (83%) were high level editorial staff. Journal readerships encompassed clinical and academic healthcare researchers (76% of respondents), healthcare decisionmakers, health service researchers, and medical generalists and specialists (50–67%), healthcare policymakers (40%), and other types of readers (37%) (students, patients, or the general public, the paramedical professions, other areas of academia).

The vast majority (92%) of journals accepted all or some types of HEOR work. Of the 10 categories of HEOR research published by respondents (figure 1), epidemiological burden of illness studies, database analyses and systematic reviews or meta-analyses were most commonly reported (77–79%), registry studies, clinical trials with economic or resource utilization data, economic burden of illness studies and epidemiological modeling studies were reported by about half the respondents (46–62%), and economic modeling, naturalistic clinical trials were least frequently reported (37–40%).

Although most journals recommended the International Committee of Medical Journal Editors (ICMJE) requirements<sup>1</sup> or specific guidances<sup>2</sup> directly in their Author Instructions, none of the journals provided their own HEOR guidance, and only 4 of the 54 responding journals recommended the BMJ health economic study guidelines<sup>59</sup>. The Cochrane website did however have links to some PRO and HEOR guidances.

About 58% of journals did not provide their reviewers with any guidelines for evaluating HEOR studies. For the 42% who said they did, in all cases, these were the same as author guidelines (e.g., instructions to authors, ICJME) and only rarely specific to HEOR.

Journals were asked if they would consider using a standard set of HEOR guidances from a recognized professional body to enhance the quality of published HEOR research in their journal; 91% said they would if these were made available.

When asked about the ease with which they found reviewers for HEOR papers, 27% of journals had great difficulty, 60% said it was difficult for some types of papers, and 6% had no difficulty. Almost 90% of respondents felt it would be useful if they had a pool of expertise available to perform reviews of HEOR for their journal. The areas of expertise specifically mentioned covered the spectrum of HEOR research: policy analysis, economic outcomes, resource utilization, clinical epidemiology, public health, preventive medicine, mental health (and other specialties), statistics, and methodologies.

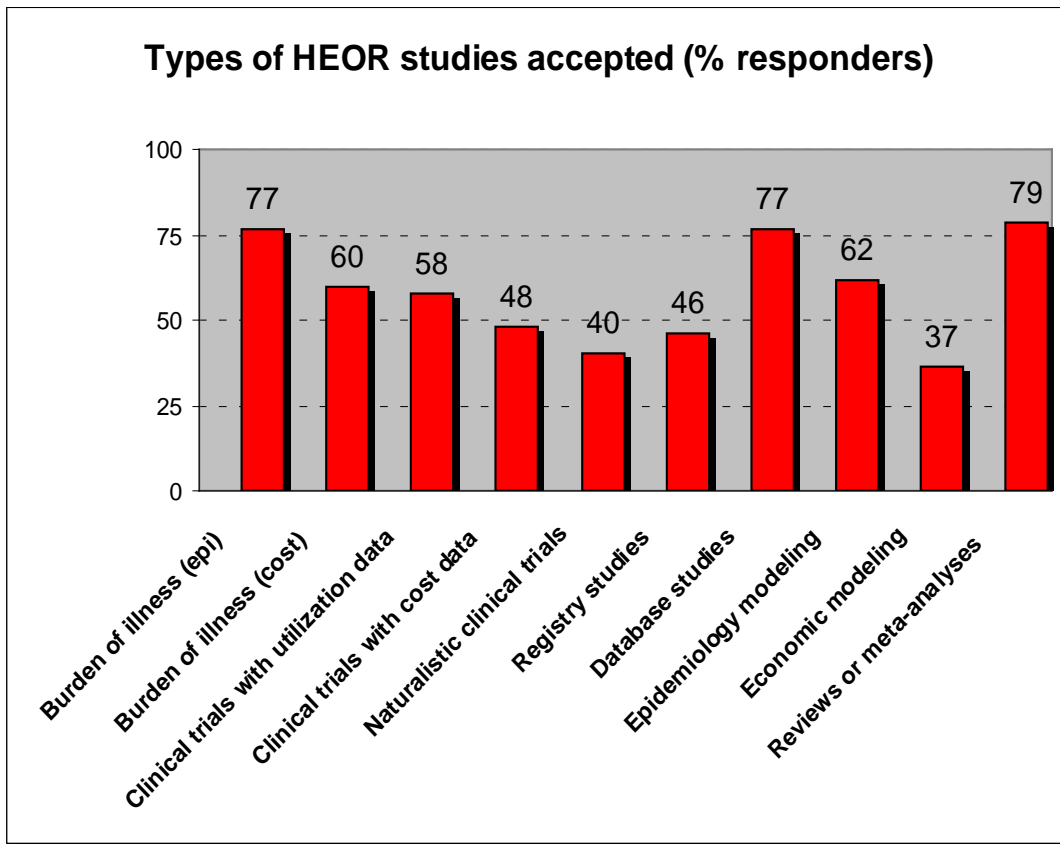
---

<sup>1</sup> which recommend no economic guidelines but some HEOR-relevant guidances (MOOSE [meta-analysis of observational studies in epidemiology], CONSORT [RCTs], STARD [studies of diagnostic accuracy], QUORUM [systematic reviews and meta-analyses] and STROBE [observational epidemiology studies])

<sup>2</sup> QUORUM, STROBE, CONSORT, STARD and two references for basic statistics

Although this survey had a small sample size, a fairly representative range of journals responded and several clear messages were heard. (1) Many biomedical journals accept and publish HEOR research. (2) Almost all do so without giving clear guidance to either authors or reviewers about quality standards for this type of research so discussion of quality control at either the manuscript production stage or the peer-review stage is currently moot. HEOR quality is currently left entirely up to the skill, knowledge and experience of each author and reviewer. (3) Most respondents expressed interest in clear guidances to which they could refer. And (4) many of the respondents reported difficulty finding expert reviewers for HEOR, and almost all were interested in having a larger pool of reviewers.

Figure 1. Types of HEOR articles published by responding journals



*Future plans*

It follows from the survey results that a first step in improving the quality of HEOR research in publications would be to (1) develop standard guidances to which journals are able to refer their authors and their reviewers; (2) lobby to establish these guidances within the ICJME Requirements to which most journals refer in their Author Instructions; and (3) provide some form of support in terms of additional expertise to those journals without appropriate reviewers. If such steps were undertaken, it would be worthwhile to subsequently resubmit the survey to WAME members to gauge change over time.

Besides working to support peer-reviewed journals, there are a number of groups which from time to time perform evaluations of HEOR research quality in the published literature.<sup>4,6,8,124</sup> It would perhaps be more useful for these evaluations to use quantitative quality measures that might be compared across evaluations, the evolution of which could be followed over time. This might be an ISPOR sponsored initiative, or one undertaken by one or more of the groups currently involved in such evaluations. Rather than reinventing the wheel, it would be worthwhile to examine possible quantitative measures already established, such as the QHES (Quality of Health Economic Studies) instrument<sup>1,7</sup>. A search for other instruments already in development is recommended.

Finally, once standardized, quantitative quality measures have been established, we recommend ongoing assessment of the quality of published HEOR and its reporting through longitudinal sampling of the literature and other publications, perhaps with annual or biannual reports.



**SECTION 5 DECISION-MAKERS, PRACTITIONERS AND EVIDENCE-BASED MEDICINE**

The process of healthcare decisionmaking in the era of evidence-based medicine (EBM) can be described in its simplest scheme by three steps:

- 1) generating evidence by researchers and compiling this evidence in databases
- 2) extracting evidence from these databases by decisionmakers or EBM practitioners
- 3) applying this evidence in healthcare settings at local levels

The three-step scheme is a clue as to what are the barriers to use and application of cost-effectiveness results by decisionmakers and clinical practitioners. The second step might cause a certain degree of information loss or inappropriate interpretation, called an evidence gap. As it is rare for decisionmakers to pick up full range of evidence from databases, there could exist a gap between stored evidence and that extracted by decisionmakers in a variety of forms such as partial evidence, abstracts, conclusions, executive summaries, commentaries, translations, etc.

It is generally recognized that practitioners in EBM and decisionmakers in healthcare tend to regard cost-effectiveness research as being of limited use even though it must be considered for rational resource allocation in healthcare. Some of obstacles for decisionmakers or EBM practitioners to properly recognize economic evaluations come from their insufficient knowledge or skills in cost-effectiveness analysis. Such insufficiencies are related to the ‘evidence gap’ mentioned above, and may lead to skepticism about cost-effectiveness research, often bringing decisionmakers to doubt the quality of the studies and the data.

One classical finding comes from a survey by Drummond et al<sup>125</sup> who surveyed almost 800 UK decisionmakers. The authors concluded that the use of health-economic evidence at local level was not extensive. The major reasons for such a limited use were either the inflexibility of budgets, limiting movement of resources between primary and secondary care, and the inability to free resources to adopt new treatments. All these issues are related to the third step in the above

decisionmaking scheme. Drummond et al also that decisionmakers were concerned about the studies themselves, specifically with respect to large numbers of assumptions and credibility of industry-funded studies. These are validity or credibility issues associated with the first step in the decisionmaking scheme.

In another survey, Duthie et al<sup>126</sup>, interviewed 17 pairs of UK NHS decisionmakers. Most information conveying health-economic outcomes, such as incremental cost-effectiveness ratios, quality-adjusted life years and willingness to pay, were either not understood or considered irrelevant by decisionmakers surveyed in their study. This issue is related to an evidence gap at the second step of decisionmaking. Those studies shed light on key concepts associated with the three-step scheme for healthcare decisionmaking: validity, generalizability, local level decisionmaking.

An interesting suggestion concerning overcoming the barriers to practical use of cost-effective evidence was investigated by Crump et al<sup>127</sup>, who interviewed 12 medical decisionmakers in the UK Leicestershire Health Authority, part of the European Network on Methodology and Application of Economic Evaluation Techniques (EUROMET) project<sup>128</sup>. Factors identified that encouraged decisionmakers to make more use of economic evaluations were: appraisal of studies by a trusted source; increasing flexibility in healthcare budgets; and more detailed explaining of the practical relevance of study results.

Barriers shown by previous studies were confirmed and more fully explored by Hoffmann et al<sup>129</sup>. The study drew on decisionmakers from two UK health authorities, Leicestershire and North Yorkshire, employing the National Health Service Economic Evaluation Database (NHS EED)<sup>130</sup> as a research vehicle. It confirmed that decisionmakers generally recognize the usefulness and necessity of published cost-effective evidence in informing their decisionmaking processes. However, they often regarded the value of studies limited due to poor generalizability, narrowness of research questions, and lack of methodological rigor, all of which are common seen in published articles. As identified by Crump et al, using trusted sources to appraise studies encouraged decisionmakers to use HE studies, as well as having a quality-scoring system for published studies and not just the critical summaries from NHS EED.

Hutton and Brown<sup>131</sup> discussed the findings of Hoffmann et al. They pointed out the general assumption has been that if decisionmakers do not find economic evaluations useful, then the way the evaluations are conducted or presented must be changed. The constraints of available resources and equity concerns could be changed, as well as the methods for measuring outcomes, but the basic approach would remain the same. Thus, if an economic decision framework is not satisfactory for decisionmakers, it must be assumed there is a different and superior model for decisionmaking. It is interesting that the discrepancy of modeling between researchers and decisionmakers might be interpreted as a further gap between the first and the third steps in the decisionmaking scheme.

Hutton and Brown claims questioned how decisions are made by health authorities for service planning and resource allocation without substantive economic input. In the absence of a clear understanding of the question, they thought economists might be more questioning of the validity of criticism of their work. To promote better use of cost-effective evidence among NHS decisionmakers, Hutton and Brown suggest:

- changing the process for NHS decisionmaking
- changing policies on funding local economic studies and expertise
- using decisionmakers with a better grasp of economics

They insisted that such changes might be more appropriate than modifying economic methods to inform unclear decision process. Regarding the classical decisionmaking behavior, Hutton and Brown observed that the Hoffmann study supported the concept that decisionmakers wanted to read commentaries rather than full abstracts, and they would prefer to be provided with quality scoring of studies.

Counter commentary by Hoffmann et al<sup>132</sup> to Hutton and Brown, agreed that there could be two approaches to addressing the problem: changes to the performance of economic evaluation and changes to the process of decisionmaking in the NHS. However, they continue to concentrate on one side of the problem, exploring ways to make economic studies more accessible without losing the key elements of critical appraisal. Nevertheless, they also state that: 1) it would be productive to examine some of the potential flaws in healthcare decisionmaking, 2) they share concern about requests for

quality scores which might lead to even less critical assessment of findings, and 3) cost-effectiveness analysis makes the shortcomings of the clinical data much more apparent. Regarding the quality of clinical effectiveness data in economic studies, we need to find out whether the problem is: the lack of good-quality data for economic evaluations *or* the lack of available good clinical data for economic studies performed in time of need. There is often a problem of a long and inexplicable lag between the publication of the first clinical data and the subsequent publication of the first cost-effectiveness study<sup>133</sup>; it was suggested that there should be better strategic planning connecting clinical and economic research plans.

Hoffmann and colleagues also expressed concern about generalizability of study findings at the local level, suggesting a stronger commitment to funding local studies and providing greater expertise at the local level to allow better adaptation of study findings. Do published cost-effectiveness studies communicate generalizable findings and could they be improved so as to make the generalizable messages self-evident. Such issues fit into the second step of the decisionmaking scheme.

In addition, Hoffmann and colleagues do not agree that there should be reorganization of the NHS decisionmaking process, as suggested by Hutton and Brown, to rectify some of the problems related to the use of economic evidence in decisionmaking. They felt such reorganization is likely to merely shuffle the pack rather than bring about real change. These arguments suggest many challenges remain for improving the application of cost-effective evidence at the local level.

Many challenges exist for EBM professionals, including better understanding of cost-effectiveness methodologies, economic evaluation alongside randomized-controlled trials, generalizability and transferability of evidence, and biases such as efficacy and efficiency gaps due to different study settings (RCTs vs real world; patients vs populations). Other challenges also exist, such as selection of appropriate comparators, ethical questions, priorities for patient outcomes (clinical benefit vs. cost-effectiveness).

Decisionmakers are also faced with such challenges as how to better understanding complex socioeconomic evaluations, how to improve decision processes, whether to take a subjective or objective perspective. Some methodological issues are key: measuring willingness-to-pay for different diseases, applying incremental cost-effectiveness ratios to budget impact analyses. Finally more difficult challenges remain to answer questions such as capturing multi-dimensional values, recognizing value-based healthcare, and ensuring social equity for patients.



**SECTION 6. QUALITY IMPROVEMENT OF COST-EFFECTIVENESS RESEARCH TASK FORCE**

**RECOMMENDATIONS**

The Quality Improvement Cost-effectiveness Research Task Force recommends that ISPOR implement the following:

*Guidelines*

- Promote harmonization of HE guidelines, allowing for differences in regional needs and politics
- Evaluate available instruments or promote development of one to quantify the impact of HE guidelines on the quality of HE studies
- Report periodically on countries using guidelines
- Evaluate periodically on quality of studies submitted to decisionmaking bodies (as public transparency increases)

*Methodologies*

- Promote publication of methodological guidelines in more applied journals in more easily understandable format to transfer knowledge to researchers who need to apply more rigorous methods
- Promote full availability of models in electronic format to combat space restrictions in hardcopy publications
- Promote consistency of methodological review for all HE studies
- Promote adoption of explicit best practices guidelines among regulatory and reimbursement authorities
- Periodically update all ISPOR task force reports
- Periodically review of use of ISPOR task force guidelines
- Periodically report on statistical and methodological challenges in HE
- Evaluate periodically whether ISPOR's methodological guidelines lead to improved quality

*Publications*

- Develop standard guidances to which journals are able to refer their authors and their reviewers
- Lobby to establish these guidances within the ICJME Requirements to which most journals refer in their Author Instructions
- Provide some form of support in terms of additional expertise to those journals without appropriate reviewers
- Periodically report on journals publishing CEA
- Periodically report on the quality of CEA publications

*Decisionmaking*

- Recognize annually at ISPOR meetings those countries (or agencies) using CEA well
- Recognize annually at ISPOR meetings those practitioners (or private companies) using CEA well
- Promote frequent publication/presentation of case studies of applied use of CEA concepts or guidelines
- Recognize annually at ISPOR meetings those practitioners/researchers supporting patient use of CEA in decisionmaking



**REFERENCES**

- 1 Chiou CF, Hay JW, Wallace JF, Bloom BS, Neumann PJ, Sullivan SD, et al. Development and validation of a grading system for the quality of cost-effectiveness studies. *Med Care*. 2003;41(1):32-44.
- 2 Hjelmgren J, Berggren F, Andersson F. Health economic guidelines--similarities, differences and some implications. *Value Health*. 2001;4(3):225-50.
- 3 Jefferson T, Smith R, Yee Y, Drummond M, Pratt M, Gale R. Evaluating the BMJ guidelines for economic submissions: prospective audit of economic submissions to BMJ and The Lancet. *JAMA*. 1998;280(3):275-7.
- 4 Neumann PJ, Olchanski NV, Rosen AB, Grenberg D, Chapman R, Stone PW, Nadai J. Are published cost-utility analyses improving? ISPOR; 2003 May 18-21; Arlington, VA.
- 5 Neumann PJ, Greenberg D, Olchanski NV, Stone PW, Rosen AB. Growth and quality of the cost-utility literature, 1976-2001. *Value Health*. 2005;8(1):3-9.
- 6 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(12):964-72.
- 7 Ofman JJ, Sullivan SD, Neumann PJ, Chiou CF, Henning JM, Wade SW, et al. Examining the value and quality of health economic analyses: implications of utilizing the QHES. *J Manag Care Pharm*. 2003;9(1):53-61.
- 8 Rosen AB, Greenberg D, Stone PW, Olchanski NV, Neumann PJ. Quality of abstracts of papers reporting original cost-effectiveness analyses. *Med Decis Making*. 2005;25(4):424-8
- 9 National Institute for Clinical Excellence. Guide to the methods of technology appraisal (reference N0515). On NICE website [updated 2008 Jun 25; cited 2008 Jul 17]. Available from: <http://www.nice.org.uk/page.aspx?o=taprocess>  
<http://www.nice.org.uk/page.aspx?o=423153>.
- 10 Canadian Agency for Drugs and Technologies in Health. Guidelines for the economic evaluation of health technologies: Canada. On CADTH website [updated 2006 Apr; cited 2006 Apr 4]. Available from: <http://www.cadth.ca>.
- 11 Commonwealth Department of Health and Ageing. Guidelines for the pharmaceutical industry on preparation of submissions to the Pharmaceutical Benefits Advisory Committee, including major submissions involving economic analyses. On Department website [updated 2002 Sep; cited 2006 Feb 14]. Available from:  
<http://www.health.gov.au/internet/wcms/Publishing.nsf/Content/health-pbs-general-pubs-guidelines-index.htm>.

- 12 Jordison CJ, Lee KM, Shukla VK, Tiernay M. An assessment of clinical evidence in economic evaluations submitted to the common drug review. Abstract presented at the 2007 CADTH Invitational Symposium; 2007 April 23-24.; Ottawa.
- 13 Gerard K, Smoker I, Seymour J. Raising the quality of cost-utility analyses: lessons learnt and still to learn. *Health Policy*. 1999;46(3):217-38.
- 14 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(16):2116-21.
- 15 Russell LB, Gold MR, Siegel JE, Daniels N, Weinstein MC. The role of cost-effectiveness analysis in health and medicine. Panel on Cost-Effectiveness in Health and Medicine. *JAMA*. 1996;276(14):1172-7.
- 16 Siegel JE, Weinstein MC, Russell LB, Gold MR. Recommendations for reporting cost-effectiveness analyses. Panel on Cost-Effectiveness in Health and Medicine. *JAMA*. 1996;276(16):1339-41.
- 17 Weinstein MC, Siegel JE, Gold MR, Kamlet MS, Russell LB. Recommendations of the Panel on Cost-effectiveness in Health and Medicine. *JAMA*. 1996;276(15):1253-8.
- 18 Mather DB, Sullivan SD, Augenstein D, Fullerton DS, Atherly D. Incorporating clinical outcomes and economic consequences into drug formulary decisions: a practical approach. *Am J Manag Care*. 1999;5(3):277-85.
- 19 Colmenero F, Sullivan SD, Palmer JA, Brauer CA, Bungay K, Watkins J, et al. Quality of clinical and economic evidence in dossier formulary submissions. *Am J Manag Care*. 2007;13(7):401-7.
- 20 Goetghebeur MM, Wagner M, Khoury H, Levitt RJ, Erickson LJ, Rindress D. Evidence and value: impact on decisionmaking - the EVIDEM framework and potential applications. *BMC Health Services Research*. 2008;270
- 21 Goetghebeur MM, Wagner M, Khoury H, Levitt R, Erickson LJ, Rindress D. Multicriteria decision analysis in formulary decisionmaking: a pilot study applying EVIDEM to historical cases in the Canadian setting. Submitted to *Open Health Services and Policy Journal*. 2009;
- 22 Behmane D, Lambot K, Irs A, Steikunas N, Hill S, Freemantle N. Baltic guideline for economic evaluation of pharmaceuticals (pharmacoeconomic analysis). On Latvian Medicines Pricing and Reimbursement Agency website [updated 2002 Aug 8; cited 2006 Feb 13]. Available from: <http://www.zca.gov.lv/english/guidelines.html>.
- 23 Walter E, Zehetmayr S. [Guidelines for health-economic evaluations in Austria]. *Wien Med Wochenschr*. 2006;156(23-24):628-32.

- 24 Cleemput I, Van Wilder P, rijens F, Huybrechts M, Ramaekers D. Recommendations pour les évaluations pharmacoéconomiques en Belgique. Belgium: 2008 Oct 4.
- 25 Ministry of Health Brazil. Health technology assessment guidelines for the Ministry of Health. On Ministry of Health Brazil website [updated 2007; cited 2009 Mar 1]. Available from:  
<http://www.ispor.org/PEguidelines/source/SECRETARIADECI%C3%84NCIATECNOLOGIAEINSUMOSESTRAT%C3%89GICOS-preliminaryversion.pdf>.
- 26 Ontario Ministry of Health and Long-Term Care. Ontario guidelines for economic analysis of pharmaceutical products. On Ministry website [updated 1994 Aug; cited 2006 Feb 14]. Available from:  
[http://www.health.gov.on.ca/english/providers/pub/pub\\_menus/pub\\_drugs.html](http://www.health.gov.on.ca/english/providers/pub/pub_menus/pub_drugs.html).
- 27 PharmacoEconomics China. Chinese PE guidelines key features. On PharmacoEconomics China website [updated 2006; cited 2009 Mar 1]. Available from: [http://www.pe-cn.org/project/guidelines\\_en.asp](http://www.pe-cn.org/project/guidelines_en.asp).
- 28 Ministry of Social Affairs and Health, Finland. Guidelines for preparing a health economic evaluation. On Pharmaceuticals Pricing Board website [updated 2003 Dec 30; cited 2006 Feb 14]. Available from:  
<http://www.stm.fi/Resource.phx/eng/orgis/board/pharmaboard/legislation.htx>.
- 29 Collège des Économistes de la Santé (CES). French guidelines for the economic evaluation of healthcare technologies: methodological recommendations. On CES website [updated 2004 Sep; cited 2006 Feb 14]. Available from: [http://www.ces-asso.org/PagesGB/default\\_gb.htm](http://www.ces-asso.org/PagesGB/default_gb.htm).
- 30 Institute for Quality and Efficiency in Healthcare (IQWiG). Methods for assessment of the relation of benefits to costs in the German Statutory Healthcare System. 2008 Sep 10.
- 31 Szende A, Mogyorosy Z, Muszbek N, Nagy J, Pallos G, Dozsa C. Methodological guidelines for conducting economic evaluation of healthcare interventions in Hungary: A Hungarian proposal for methodology standards. *Eur J Health Econ.* 2002;3(3):196-206.
- 32 National Centre for Pharmacoeconomics in Ireland, Irish Pharmaceutical Healthcare Association. Irish healthcare technology assessment guidelines: Version 1. On NCPE website [updated 2000 Feb 7; cited 2006 Feb 14]. Available from: <http://www.ncpe.ie/category.php?cid=26>.
- 33 Pharmaceutical Administration, Ministry of Health, Israel. Guidelines for the submission of a request to include a pharmaceutical product in the national list of health services. Guideline no. 53, version no: 6. On Ministry of Health

website [updated 2007 Feb; cited 2008 Sep 22]. Available from:

[www.health.gov.il/download/forms/a28\\_hachlala\\_eng.doc](http://www.health.gov.il/download/forms/a28_hachlala_eng.doc).

- 34 Capri S, Ceci A, Terranova L, Merlo F, Mantovani L, and the members of the Italian Group for Pharmacoeconomic Studies. Guidelines for economic evaluations in Italy: recommendations from the Italian Group of Pharmacoeconomic Studies. *Drug Inform J*. 2001;35(1):189-201.
- 35 Healthcare Insurance Board. Dutch guidelines for pharmacoeconomic research. Amstelveen, The Netherlands: Health Insurance Council; 2008 Apr 1.
- 36 Pharmaceutical Management Agency Ltd, New Zealand. A prescription for pharmacoeconomic analysis: Version 2. On PHARMAC website [updated 2007 Jul; cited 2008 Jul 17]. Available from:  
[http://www.pharmac.govt.nz/pharmo\\_economic.asp](http://www.pharmac.govt.nz/pharmo_economic.asp).
- 37 Norwegian Medicines Agency. Norwegian guidelines for pharmacoeconomic analysis in connection with applications for reimbursement. On NoMA website [updated 2005 Oct 27; cited 2006 Feb 14]. Available from:  
<http://www.legemiddelverket.no>.
- 38 Orlewska E, Mierzejewski P. Polish guidelines for conducting pharmacoeconomic evaluations. On ISPOR website [updated 2000; cited 2009 Mar 3]. Available from:  
<http://www.ispor.org/PEguidelines/source/PolishGuidelinesForConductingPharmacoeconomicsEvaluations.pdf>.
- 39 Orlewska E, Mierzejewski P. Polish guidelines for cost calculation in the economic evaluation of healthcare programs (project). *Farmakoekonomika*. 2003;1:3-6.
- 40 Alves da Silva EA, Pinto CG, Sampaio C, Pereira JA, Drummond M, Trindade R. Guidelines for economic drug evaluation studies. On ISPOR and Emilia Alves da Silva website [updated 1998 Nov; cited 2006 Feb 14]. Available from: <http://www.ispor.org/PEguidelines/COUNTRYSPECIFIC.asp>.
- 41 Ministry of Health, Russian Federation. Branch Standard: The Standardization System in the Russian Federation Healthcare System. Clinico-Economic Studies, General Provisions. On ISPOR website [updated 2002 Jun 29; cited 2006 Feb 14]. Available from: <http://www.ispor.org/PEguidelines/COUNTRYSPECIFIC.asp>.
- 42 Scottish Medicines Consortium. Guidance to manufacturers for completion of New Product Assessment Form (NPAF) (Clinical and Economic Sections combined). On Consortium website [updated 2007 Jun; cited 2006 Feb 14]. Available from: <http://www.scottishmedicines.org.uk>.

- 43 Ministry of Health Slovak Republic. Pharmacoeconomic guidelines in Slovak Republic. On Ministry of Health Slovak Republic website [updated 2008; cited 2009 Mar 1].
- 44 Health Insurance Review Agency. Korean PE guidelines. On Health Insurance Review Agency website [updated 2006 Jan; cited 2009 Mar 1]. Available from:  
[http://www.ispor.org/PEguidelines/source/Korean\\_PE\\_Guidelines\\_Korean\\_Version.pdf](http://www.ispor.org/PEguidelines/source/Korean_PE_Guidelines_Korean_Version.pdf).
- 45 Rovira J, Antonanzas F. Economic analysis of health technologies and programmes. A Spanish proposal for methodological standardisation. *Pharmacoeconomics*. 1995;8(3):245-52.
- 46 Pharmaceutical Benefits Board, Sweden. General guidelines for economic evaluations from the Pharmaceutical Benefits Board (LFNAR 2003:2). On Agency website [updated 2003 Apr 24; cited 2006 Feb 14]. Available from:  
[http://www.lfn.se/LFNTemplates/Page\\_\\_\\_\\_\\_442.aspx](http://www.lfn.se/LFNTemplates/Page_____442.aspx).
- 47 Swiss Federal Office of Social Security. Manual for the standardisation of clinical and economic evaluation of medical technology. 1995 May 9. Report No.: 2nd draft.
- 48 National Institute for Clinical Excellence. Updated guide to the methods of technology appraisal. On NICE website [updated 2008 Jun 25; cited 2008 Jul 17]. Available from: <http://www.nice.org.uk/page.aspx?o=taprocess>  
<http://www.nice.org.uk/page.aspx?o=423153>.
- 49 Academy of Managed Care Pharmacy. The AMCP format for formulary submissions. Version 2.1. On Foundation for Managed Care Pharmacy website [updated 2005; cited 2006 Feb 13]. Available from:  
<http://www.amcp.org/amcp.ark?c=pr&sc=link>.
- 50 National Pharmacy and Therapeutics Committee, WellPoint Pharmacy Management. Health technology assessment guidelines: drug submission guidelines for new products, new indications and new formulations. Version 5.1. On WellPoint Pharmacy Management website [updated 2005 Sep; cited 2006 Feb 14]. Available from:  
<https://www.wellpointrx.com/wps/portal/wpo/client/clinicalservices/nationalPharmacyTherapeutics>.
- 51 Siegel JE, Torrance GW, Russell LB, Luce BR, Weinstein MC, Gold MR. Guidelines for pharmacoeconomic studies: recommendations from the Panel on Cost-effectiveness in Health and Medicine. *Pharmacoeconomics*. 1997;11(2):159-68.
- 52 Task Force on Principles for Economic Analysis of Healthcare Technology. Economic analysis of healthcare technology. A report on principles. *Ann Intern Med*. 1995;123(1):61-70.

- 53 Ramsey S, Willke R, Briggs A, Brown R, Buxton M, Chawla A, et al. Good research practices for cost-effectiveness analysis alongside clinical trials: the ISPOR RCT-CEA Task Force report. *Value Health*. 2005;8(5):521-33.
- 54 Doshi JA, Glick HA, Polsky D. Analyses of cost data in economic evaluations conducted alongside randomized controlled trials. *Value Health*. 2006;9(5):334-40.
- 55 Barber JA, Thompson SG. Analysis and interpretation of cost data in randomised controlled trials: review of published studies. *BMJ*. 1998;317(7167):1195-200.
- 56 Drummond M, O'Brien B, Stoddart GL, Torrance GW. *Methods for the economic evaluation of healthcare programmes*. 2nd ed. New York: Oxford University Press; 1997.
- 57 Robinson R. Costs and cost-minimisation analysis. *BMJ*. 1993;307(6906):726-8.
- 58 Gold MR, Siegel JE, Russell LB, Weinstein MC. *Cost-effectiveness in health and medicine*. New York: Oxford University Press; 1996.
- 59 Drummond MF, Jefferson TO. Guidelines for authors and peer reviewers of economic submissions to the *BMJ*. The *BMJ* Economic Evaluation Working Party. *BMJ*. 1996;313(7052):275-83.
- 60 Glick H, Polsky D, Schulman K. Trial-based economic evaluation: an overview of design and analysis. In: Drummond M, McGuire A, editors. *Economic evaluation in healthcare: merging theory with practice*. Oxford: Oxford University Press; 2001.
- 61 O'Brien BJ, Drummond MF, Labelle RJ, Willan A. In search of power and significance: issues in the design and analysis of stochastic cost-effectiveness studies in healthcare. *Med Care*. 1994;32(2):150-63.
- 62 Chaudhary MA, Stearns SC. Estimating confidence intervals for cost-effectiveness ratios: an example from a randomized trial. *Stat Med*. 1996;15(13):1447-58.
- 63 Polsky D, Glick HA, Willke R, Schulman K. Confidence intervals for cost-effectiveness ratios: a comparison of four methods. *Health Econ*. 1997;6(3):243-52.
- 64 Willan AR, O'Brien BJ. Confidence intervals for cost-effectiveness ratios: an application of Fieller's theorem. *Health Econ*. 1996;5(4):297-305.
- 65 Van Hout BA, Al MJ, Gordon GS, Rutten FF. Costs, effects and C/E-ratios alongside a clinical trial. *Health Econ*. 1994;3(5):309-19.
- 66 Stinnett AA, Mullahy J. Net health benefits: a new framework for the analysis of uncertainty in cost-effectiveness analysis. *Med Decis Making*. 1998;18(2 Suppl):S68-S80.

- 67 Glick H, Briggs AH, Polsky D. Quantifying stochastic uncertainty and presenting results of cost-effectiveness analyses. *Pharmacoeconom Outcomes Res.* 2001;1:25-36.
- 68 Briggs AH, O'Brien BJ. The death of cost-minimization analysis? *Health Econ.* 2001;10(2):179-84.
- 69 Thompson SG, Barber JA. How should cost data in pragmatic randomised trials be analysed? *BMJ.* 2000;320(7243):1197-200.
- 70 Manning WG. The logged dependent variable, heteroscedasticity, and the retransformation problem. *J Health Econ.* 1998;17(3):283-95.
- 71 Barber JA, Thompson SG. Analysis of cost data in randomized trials: an application of the non-parametric bootstrap. *Stat Med.* 2000;19(23):3219-36.
- 72 Efron B, Tibshirani R. *An introduction to the bootstrap.* New York: Chapman & Hall; 1993.
- 73 Desgagne A, Castilloux AM, Angers JF, LeLorier J. The use of the bootstrap statistical method for the pharmacoeconomic cost analysis of skewed data. *Pharmacoeconomics.* 1998;13(5 Pt 1):487-97.
- 74 Manning WG, Mullahy J. Estimating log models: to transform or not to transform? *J Health Econ.* 2001;20(4):461-94.
- 75 McCulloch P, Nelder JA. *Generalized linear models.* 2nd edition ed. London: Chapman & Hall; 1989.
- 76 Blough DK, Madden CW, Hornbrook MC. Modeling risk using generalized linear models. *J Health Econ.* 1999;18(2):153-71.
- 77 Lin DY, Feuer EJ, Etzioni R, Wax Y. Estimating medical costs from incomplete follow-up data. *Biometrics.* 1997;53(2):419-34.
- 78 Raikou M, McGuire A. Estimating medical care costs under conditions of censoring. *J Health Econ.* 2004;23(3):443-70.
- 79 Etzioni R, Urban N, Baker M. Estimating the costs attributable to a disease with application to ovarian cancer. *J Clin Epidemiol.* 1996;49(1):95-103.
- 80 Lin DY. Linear regression analysis of censored medical costs. *Biostatistics.* 2000;1(1):35-47.
- 81 Lin DY. Proportional means regression for censored medical costs. *Biometrics.* 2000;56(3):775-8.
- 82 Bang H, Tsiatis AA. Estimated medical costs with censored data. *Biometrika.* 2000;87:329-43.
- 83 Lin DY. Regression analysis of incomplete medical cost data. *Stat Med.* 2003;22(7):1181-200.

- 84 Carides GW, Heyse JF, Iglewicz B. A regression-based method for estimating mean treatment cost in the presence of right-censoring. *Biostatistics*. 2000;1(3):299-313.
- 85 Lavori PW, Dawson R, Shera D. A multiple imputation strategy for clinical trials with truncation of patient data. *Stat Med*. 1995;14(17):1913-25.
- 86 Oostenbrink JB, Al MJ, Rutten-van Molken MP. Methods to analyse cost data of patients who withdraw in a clinical trial setting. *Pharmacoeconomics*. 2003;21(15):1103-12.
- 87 Oostenbrink JB, Al MJ. The analysis of incomplete cost data due to dropout. *Health Econ*. 2005;14(8):763-76.
- 88 Fenn P, McGuire A, Phillips V, Backhouse M, Jones D. The analysis of censored treatment cost data in economic evaluation. *Med Care*. 1995;33(8):851-63.
- 89 O'hagan A, Stevens JW. On estimators of medical costs with censored data. *J Health Econ*. 2004;23(3):615-25.
- 90 Al MJ, Van Hout BA, Michel BC, Rutten FF. Sample size calculation in economic evaluations. *Health Econ*. 1998;7(4):327-35.
- 91 Briggs AH, Gray AM. Power and sample size calculations for stochastic cost-effectiveness analysis. *Med Decis Making*. 1998;18(2 Suppl):S81-S92.
- 92 Laska EM, Meisner M, Siegel C. Power and sample size in cost-effectiveness analysis. *Med Decis Making*. 1999;19(3):339-43.
- 93 Willan AR, O'Brien BJ. Sample size and power issues in estimating incremental cost-effectiveness ratios from clinical trials data. *Health Econ*. 1999;8(3):203-11.
- 94 Glick H, Doshi J, Sonnad S, Polsky D. *Economic evaluation in clinical trials*. Oxford: Oxford University Press; 2007.
- 95 Cook JR, Drummond M, Glick H, Heyse JF. Assessing the appropriateness of combining economic data from multinational clinical trials. *Stat Med*. 2003;22(12):1955-76.
- 96 Gail M, Simon R. Testing for qualitative interactions between treatment effects and patient subsets. *Biometrics*. 1985;41(2):361-72.
- 97 Willke RJ, Glick HA, Polsky D, Schulman K. Estimating country-specific cost-effectiveness from multinational clinical trials. *Health Econ*. 1998;7(6):481-93.
- 98 Koopmanschap MA, Touw KC, Rutten FF. Analysis of costs and cost-effectiveness in multinational trials. *Health Policy*. 2001;58(2):175-86.
- 99 Rice N, Jones A. Multilevel models and health economics. *Health Econ*. 1997;6(6):561-75.

- 100 Manca A, Rice N, Sculpher MJ, Briggs AH. Assessing generalisability by location in trial-based cost-effectiveness analysis: the use of multilevel models. *Health Econ.* 2005;14(5):471-85.
- 101 Grieve R, Nixon R, Thompson SG, Normand C. Using multilevel models for assessing the variability of multinational resource use and cost data. *Health Econ.* 2005;14(2):185-96.
- 102 Grieve R, Nixon R, Thompson SG, Cairns J. Multilevel models for estimating incremental net benefits in multinational studies. *Health Econ.* 2007;16(8):815-26.
- 103 Pinto EM, Willan AR, O'Brien BJ. Cost-effectiveness analysis for multinational clinical trials. *Stat Med.* 2005;24(13):1965-82.
- 104 Willan AR, Pinto EM, O'Brien BJ, Kaul P, Goeree R, Lynd L, et al. Country specific cost comparisons from multinational clinical trials using empirical Bayesian shrinkage estimation: the Canadian ASSENT-3 economic analysis. *Health Econ.* 2005;14(4):327-38.
- 105 Manca A, Lambert PC, Sculpher M, Rice N. Cost-effectiveness analysis using data from multinational trials: the use of bivariate hierarchical modeling. *Med Decis Making.* 2007;27(4):471-90.
- 106 Thompson SG, Nixon RM, Grieve R. Addressing the issues that arise in analysing multicentre cost data, with application to a multinational study. *J Health Econ.* 2006;25(6):1015-28.
- 107 Weinstein MC, O'Brien B, Hornberger J, Jackson J, Johannesson M, McCabe C, et al. Principles of good practice for decision analytic modeling in healthcare evaluation: report of the ISPOR Task Force on Good Research Practices-- Modeling Studies. *Value Health.* 2003;6(1):9-17.
- 108 Riley RD, Simmonds MC, Look MP. Evidence synthesis combining individual patient data and aggregate data: a systematic review identified current practice and possible methods. *J Clin Epidemiol.* 2007;60(5):431-9.
- 109 DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials.* 1986;7(3):177-88.
- 110 Sutton AJ, Abrams K, Jones DR. *Methods for meta-analysis in medical research.* London: Wiley; 2000.
- 111 Ades AE. A chain of evidence with mixed comparisons: models for multi-parameter synthesis and consistency of evidence. *Stat Med.* 2003;22(19):2995-3016.
- 112 Lu G, Ades AE. Combination of direct and indirect evidence in mixed treatment comparisons. *Stat Med.* 2004;23(20):3105-24.
- 113 Caldwell DM, Ades AE, Higgins JP. Simultaneous comparison of multiple treatments: combining direct and indirect evidence. *BMJ.* 2005;331(7521):897-900.

- 114 Jansen JP, Crawford B, Bergman G, Stam W. Bayesian Meta-Analysis of Multiple Treatment Comparisons: An Introduction to Mixed Treatment Comparisons. *Value Health*. 2008;
- 115 Doubilet P, Begg CB, Weinstein MC, Braun P, McNeil BJ. Probabilistic sensitivity analysis using Monte Carlo simulation. A practical approach. *Med Decis Making*. 1985;5(2):157-77.
- 116 Manning WG, Fryback DG, Weinstein M. Reflecting uncertainty in cost-effectiveness analysis. In: Gold M, Siegel J, Russell L, Weinstein M, editors. *Cost-effectiveness in health and medicine*. Oxford: Oxford University Press; 1996.
- 117 Claxton K. The irrelevance of inference: a decisionmaking approach to the stochastic evaluation of healthcare technologies. *J Health Econ*. 1999;18(3):341-64.
- 118 Claxton K, Sculpher M, Drummond M. A rational framework for decisionmaking by the National Institute For Clinical Excellence (NICE). *Lancet*. 2002;360(9334):711-5.
- 119 Brennan A, Kharroubi S, O'hagan A, Chilcott J. Calculating partial expected value of perfect information via Monte Carlo sampling algorithms. *Med Decis Making*. 2007;27(4):448-70.
- 120 Ades AE, Lu G, Claxton K. Expected value of sample information calculations in medical decision modeling. *Med Decis Making*. 2004;24(2):207-27.
- 121 Kim LG, Thompson SG. Uncertainty and validation of health economic decision models. *Health Econ*. 2009;
- 122 Gore SM, Jones G, Thompson SG. The Lancet's statistical review process: areas for improvement by authors. *Lancet*. 1992;340(8811):100-2.
- 123 Altman DG, Goodman SN. Transfer of technology from statistical journals to the biomedical literature. Past trends and future predictions. *JAMA*. 1994;272(2):129-32.
- 124 Bell CM, Urbach DR, Ray JG, Bayoumi A, Rosen AB, Greenberg D, et al. Bias in published cost-effectiveness studies: systematic review. *BMJ*. 2006;332(7543):699-703.
- 125 Drummond M, Cooke J, Walley T. Economic evaluation under managed competition: evidence from the UK *Soc Sci Med*. 1997;45(4):583-95.
- 126 Duthie T, Trueman P, Chancellor J, Diez L. Research into the use of health economics in decisionmaking in the United Kingdom--Phase II. Is health economics 'for good or evil'? *Health Policy*. 1999;46(2):143-57.
- 127 Crump B, Drummond MF, Alexander S, Devaney C. Economic evaluation in the United Kingdom National Health Service. In: Graf von der Schulenburg J-M, editor. *The influence of economic evaluation studies on healthcare decisionmaking: a European survey*. Amsterdam: I.O.S. Press; 2000.

- 128 Hoffmann C, Graf von der Schulenburg JM. The influence of economic evaluation studies on decisionmaking. A European survey. The EUROMET group. *Health Policy*. 2000;52(3):179-92.
- 129 Hoffmann C, Stoykova BA, Nixon J, Glanville JM, Misso K, Drummond MF. Do healthcare decisionmakers find economic evaluations useful? The findings of focus group research in UK health authorities. *Value Health*. 2002;5(2):71-8.
- 130 NHS Centre for Reviews and Dissemination. Improving access to cost-effectiveness information for healthcare decisionmaking. The NHS economic evaluation database. York: University of York; 2001.
- 131 Hutton J, Brown RE. Use of economic evaluation in decisionmaking: what needs to change? *Value Health*. 2002;5(2):65-6.
- 132 Hoffmann C, Stoykova B, Nixon J, Glanville J, Misso K, Drummond M. Re: Hutton J, Brown R. Use of economic evaluation in decisionmaking: what needs to change? (Editorial) *Value Health* 2002;5:65-6. *Value Health*. 2002;5(2):114.
- 133 Stoykova B, Drummond M, Barbieri M, Kleijnen J. The lag between effectiveness and cost-effectiveness evidence of new drugs. Implications for decisionmaking in healthcare. *Eur J Health Econ*. 2003;4(4):313-8.