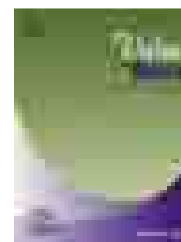




ELSEVIER

Available online at [www.sciencedirect.com](http://www.sciencedirect.com)

ScienceDirect

journal homepage: [www.elsevier.com/locate/jval](http://www.elsevier.com/locate/jval)

## Cost-Effectiveness Analysis Alongside Clinical Trials II—An ISPOR Good Research Practices Task Force Report

Scott D. Ramsey, MD, PhD<sup>1,2,\*</sup>, Richard J. Willke, PhD<sup>3</sup>, Henry Glick, PhD<sup>4</sup>, Shelby D. Reed, PhD, RPh<sup>5</sup>, Federico Augustovski, MD, MSc, PhD<sup>6</sup>, Bengt Jonsson, PhD<sup>7</sup>, Andrew Briggs, MSc, DPhil<sup>8</sup>, Sean D. Sullivan, PhD<sup>1,2</sup>

<sup>1</sup>Hutchinson Institute for Cancer Outcomes Research, Fred Hutchinson Cancer Research Center, Seattle, WA, USA; <sup>2</sup>Schools of Medicine and Pharmacy, University of Washington, Seattle, WA, USA; <sup>3</sup>Outcomes & Evidence Lead, CV/Metabolic, Pain, Urology, Gender Health, Global Health & Value, Pfizer, Inc., New York, NY, USA; <sup>4</sup>Division of General Internal Medicine, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA; <sup>5</sup>Duke Clinical Research Institute, Duke University, Durham, NC, USA; <sup>6</sup>Institute for Clinical Effectiveness and Health Policy (IECS), University of Buenos Aires, Buenos Aires, Argentina; <sup>7</sup>Department of Economics, Stockholm School of Economics, Stockholm, Sweden; <sup>8</sup>William R. Lindsay Chair of Health Economics, University of Glasgow, Glasgow, Scotland, UK

### ABSTRACT

Clinical trials evaluating medicines, medical devices, and procedures now commonly assess the economic value of these interventions. The growing number of prospective clinical/economic trials reflects both widespread interest in economic information for new technologies and the regulatory and reimbursement requirements of many countries that now consider evidence of economic value along with clinical efficacy. As decision makers increasingly demand evidence of economic value for health care interventions, conducting high-quality economic analyses alongside clinical studies is desirable because they broaden the scope of information available on a particular intervention, and can efficiently provide timely information with high internal and, when designed and analyzed properly, reasonable external validity. In 2005, ISPOR published the Good Research Practices for Cost-Effectiveness Analysis Alongside Clinical Trials: The ISPOR RCT-CEA Task Force report. ISPOR initiated an update of the report in 2014 to include the methodological developments over the last 9 years. This report provides updated recommendations reflecting advances in several areas related to trial design, selecting data elements, database

design and management, analysis, and reporting of results. Task force members note that trials should be designed to evaluate effectiveness (rather than efficacy) when possible, should include clinical outcome measures, and should obtain health resource use and health state utilities directly from study subjects. Collection of economic data should be fully integrated into the study. An incremental analysis should be conducted with an intention-to-treat approach, complemented by relevant subgroup analyses. Uncertainty should be characterized. Articles should adhere to established standards for reporting results of cost-effectiveness analyses. Economic studies alongside trials are complementary to other evaluations (e.g., modeling studies) as information for decision makers who consider evidence of economic value along with clinical efficacy when making resource allocation decisions.

**Keywords:** clinical trial, cost-effectiveness, economic, guidelines.

Copyright © 2015, International Society for Pharmacoeconomics and Outcomes Research (ISPOR). Published by Elsevier Inc.

### Introduction

Data from clinical trials that evaluate medicines, medical devices, and procedures are now commonly used to assess the value for money of these interventions. A growing number of clinical trials include specific data on resource use and outcome for assessment of cost-effectiveness. This growth reflects both widespread interest in economic information for new technologies and the regulatory and reimbursement requirements of many countries that now consider evidence of economic value along with clinical

effectiveness. In the last decade, researchers have improved the methods used for the design, conduct, and analysis of data for economic evaluation collected alongside clinical trials. Despite these advances, the literature reveals a great deal of variation in methodology and reporting of cost-effectiveness analyses (CEAs). Improving the quality of these studies will enhance their credibility and usefulness to decision makers worldwide.

Early health technology assessments (HTAs) or joint regulatory/HTA advice is used to communicate the potential needs for payer decisions on reimbursement and funding [1]. Payers are

\* Address correspondence to: Scott D. Ramsey, Hutchinson Institute for Cancer Outcomes Research, Fred Hutchinson Cancer Research Center, 1100 Fairview Avenue North, M3-B232, Seattle, WA 98109.

E-mail: [sramsey@fredhutch.org](mailto:sramsey@fredhutch.org)

1098-3015/\$36.00 see front matter Copyright © 2015, International Society for Pharmacoeconomics and Outcomes Research (ISPOR).

Published by Elsevier Inc.

<http://dx.doi.org/10.1016/j.jval.2015.02.001>

## Background to the Task Force

In 2005, ISPOR published the Good Research Practices for Cost-Effectiveness Analysis Alongside Clinical Trials: The ISPOR RCT-CEA Task Force report ([http://www.ispor.org/workpaper/research\\_practices/Good\\_Research\\_Practices-Cost\\_Effectiveness\\_Analysis\\_with\\_Clinical\\_Trials.pdf](http://www.ispor.org/workpaper/research_practices/Good_Research_Practices-Cost_Effectiveness_Analysis_with_Clinical_Trials.pdf)) to address issues related to trial design, selecting data elements, database design and management, analysis, and reporting of results. The findings included the following: trials should be designed to evaluate effectiveness (rather than efficacy), should include clinical outcome measures, and should obtain health resource use and health state utilities directly from study subjects. Collection of economic data should be fully integrated into the study. Analyses should be guided by an analysis plan and hypotheses. An incremental analysis should be conducted with an intention-to-treat approach. Uncertainty should be characterized. Articles should adhere to established standards for reporting results of cost-effectiveness analyses.

After 9 years, the cochairs determined an update was appropriate. They submitted a proposal to the Health Science Policy Council in January 2014 to review the methodological and applied studies published since the original report. The proposed new report would reflect the advances that should be considered standard of care for these types of studies. The ISPOR Board of Directors approved the proposal in February 2014.

The task force is composed of experts in designing and conducting clinical trials, modeling, economic evaluation,

statistical methods, and quality-of-life research. Task force members were selected to represent a diverse range of perspectives, including public and private research centers, academia, hospitals, and the pharmaceutical industry. Task force members are also internationally based, and include individuals from the United Kingdom, Sweden, Argentina, Canada, and the United States.

The task force met approximately every 6 weeks by teleconference to develop an outline and discuss issues to be included in the report. In addition, task force members met in person at ISPOR International meetings and European congresses. All task force members reviewed many drafts of the report and provided frequent feedback in both oral and written comments.

Preliminary findings and recommendations were discussed in a forum presentation at the 2014 ISPOR Annual European Congress in Amsterdam. In addition, written feedback was received from the first and final draft reports' circulation to the ISPOR Economic Evaluation Review Group. The task force discussed comments on a series of teleconferences and in person at the ISPOR Amsterdam Conference. All comments were considered, and most were substantive and constructive.

Comments were addressed as appropriate in subsequent versions of the report. All written comments are published at the ISPOR Web site on the task force's Webpage: [http://www.ispor.org/TaskForces/Cost\\_Effectiveness\\_Analysis\\_Clinical\\_Trials-GRPIITF.asp](http://www.ispor.org/TaskForces/Cost_Effectiveness_Analysis_Clinical_Trials-GRPIITF.asp). The task force report and Webpage may also be accessed from the ISPOR homepage ([www.ispor.org](http://www.ispor.org)) via the purple Research Tools menu, ISPOR Good Practices for Outcomes Research, heading: Economic Evaluation Methods.

important stakeholders in data from clinical trials; without third-party payments, patients will have very limited access to new treatments. In addition, the number of audiences for clinical and economic trials is growing, and now include governments, third-party insurers, delivery systems, health care professionals, and patients. Thus, there is an increased focus on whether economic results from clinical trials are externally valid for different jurisdictions. Furthermore, the growing prevalence and cost of chronic progressive diseases such as rheumatoid arthritis, multiple sclerosis, and cancer has increased the number of economic evaluations that are based on a combination of clinical trial data and external data representative of routine clinical practice [2]. Finally, the increasing number of randomized trials that evaluate comparative effectiveness in clinical practice, for example, using data from patient registries, provides new opportunities for economic evaluations alongside randomized controlled trials (RCTs) [3,4].

To foster improvements in the conduct and reporting of trial-based economic analyses, the International Society for Pharmacoeconomics and Outcomes Research (ISPOR) published the Good Research Practices for Cost-Effectiveness Analysis Alongside Clinical Trials: The ISPOR RCT-CEA Task Force report in 2005 [5]. To address changes in standards for designing, conducting, and reporting cost-effectiveness in clinical trials, ISPOR initiated an update of the report in 2014 to include the methodological developments over the last 9 years. This report's recommendations reflect advances in several areas related to trial design, selecting data elements, database design and management, analysis, and reporting of results (see Table 1).

The intended audience for this report is researchers in academia, industry, and government who design and implement these studies, decision makers who evaluate clinical and economic evidence for reimbursement and formulary coverage policies, and students.

The authors recognize that advances in methodology for joint clinical/economic analyses will continue and that clinical/

**Table 1 – Significant updates to 2014 Task Force Report since the 2005 ISPOR RCT-CEA Task Force Report.**

Trial design	<ul style="list-style-type: none"> <li>• Developments in improving external validity of trial-based CEA</li> <li>• Considerations regarding value-of-information (VOI) analysis for sample size calculations</li> <li>• Discussion of appropriate comparator characteristics</li> <li>• Additional considerations for timing of health economic data collection</li> </ul>
Data elements and database management	<ul style="list-style-type: none"> <li>• Use of VOI analysis to help determine value of sample information</li> <li>• Use of electronic medical records and other new technologies for data collection</li> <li>• Discussion of mapping disease-specific instruments to preference weights</li> <li>• Collection of provider-, site-, and jurisdiction-specific data</li> </ul>
Analysis	<ul style="list-style-type: none"> <li>• Recent approaches to analyzing preference scores</li> <li>• Examples for modeling outcomes beyond the time horizon of the trial</li> <li>• Use of provider-, site-, and jurisdiction-specific data in cross-country transferability adjustments</li> <li>• Additional approaches to select subgroup analysis</li> <li>• Expanded uncertainty analysis section</li> </ul>
Reporting	<ul style="list-style-type: none"> <li>• Addition of Consolidated Health Economic Evaluation Reporting Standards (CHEERS) elements</li> </ul>
References	<ul style="list-style-type: none"> <li>• Added 70 new references across most topics</li> </ul>

CEA, cost-effectiveness analysis; RCT, randomized controlled trial.

economic trials are heterogeneous in nature. Therefore, the report highlights areas of consensus, emerging methodologies with a diversity of professional opinions, and issues for which further research and development are needed.

Common issues that are fundamental to all CEAs, such as choice of discount rate for costs and outcomes or types of costs that will be included (direct medical and nonmedical, indirect costs for loss of productivity, etc.), will not be addressed in this article. These topics are well described elsewhere [6].

### Initial Trial Design Issues

The quality of economic information that is derived from trials depends on attributes of the trial's design. The commonly used phrase that economic analyses are conducted "alongside" clinical trials is indicative of an important practical design issue: *assessing relative value is rarely the primary purpose of an experimental study*. Nevertheless, when the decision is made to conduct an economic evaluation alongside a clinical trial, it is important that the economic investigator contributes to the design of the study to ensure that the trial will provide the data necessary for a high-quality economic evaluation.

#### Appropriate Trial Design

The development of HTA as a method to inform decisions about the use of medicines has put new demands on clinical trials. Such studies focus on relative effectiveness; that is, studies should directly compare effective interventions and should include patients who are typical of day-to-day health care settings. In some regions, specific guidelines that have implications for trial design have been developed for the early assessment of pharmaceuticals [7].

It is generally acknowledged that pragmatic effectiveness trials are the best vehicle for economic studies [8,9]. It is usually necessary, however, to undertake economic evaluations earlier in the development cycle when the focus is on efficacy (including phase III or even phase II drug trials), to provide timely information for pricing and reimbursement.

Joint RCT-CEA evaluations can provide important information for adoption decisions, including early evaluations (e.g., phases II and III drug trials) when there are little or no additional data to be synthesized, evaluations of trials with large enrollments and long follow-up, and in cases in which trials reveal dramatic health improvements for interventions such that very early posttrial adoption is considered. Even trials that do not satisfy these conditions may be valuable, particularly if they report their data in such a way that their results can be synthesized with those of other studies.

Some have questioned the use of clinical trials as a vehicle for economic evaluations because of the often artificial nature of trials and patient selection issues related to clinical practice among other concerns [9]. Although these issues are important challenges for trial-based economic evaluations, new developments have increased the value of such studies. During the last decade, some HTA authorities and payers have had greater roles and influence on the planning and design of early confirmatory clinical trials [10,11]. The European Commission has funded the Shaping European Early Dialogues for health technologies, for early dialogues between member HTA agencies and developers of pharmaceutical devices, and several pilot projects have been conducted [12]. Although in formative stages, there is a growing harmonization between the evidence requirements by regulators and HTA/reimbursement agencies and payers, suggesting that in the future, economic evaluations alongside RCT will also meet objectives of external validity [13,14].

When an economic investigator is asked to participate in designing an economic evaluation alongside a clinical trial, he or she should first consider the extent to which that trial is likely to be judged as an appropriate vehicle for an economic evaluation, after adaptation in relevant aspects. Not all trials are appropriate for economic analyses, for example, those that are exploratory in nature such as dose finding studies, or have design issues that limit their potential to affect clinical practice (and therefore of limited interest to decision makers). For example, if so much trial-specific health care is mandated that all potential differences in cost are overwhelmed, or if participants are followed differentially on the basis of outcome, it is unlikely that an economic assessment will be informative.

#### Identifying and Addressing Threats to External Validity/Generalizability

The "artificiality" of most clinical trials can pose serious threats to external validity. These threats stem from

1. protocol-driven resource use (which could bias costs in each treatment arm upwards if included and downwards if excluded, but it is generally difficult to know how this will bias the difference between treatments);
2. comparators that are uncommon or not recommended in clinical practice;
3. recruiting that is not representative of the larger patient population (e.g., large, urban, academic hospitals); inclusion of study sites from countries with varying access and availability of health care services (e.g., rehabilitation, home care, or emergency services);
4. restrictive inclusion and exclusion criteria (patient population, disease severity, comorbidities); and
5. artificially enhanced compliance.

Perhaps most important among these factors is having a relevant comparator, that is, one that represents best current or most common practice (ideally both). The external validity can best be increased by designing the trial to be more naturalistic, that is, designed to evaluate the effectiveness of interventions in real-life routine practice conditions [6,15].

Additional challenges arise with international trials. There can be tremendous differences in treatment pathways, patient and health care provider behavior, supply and financing of health care, and unit costs (prices) between countries [16–23]. Pooled results may not be representative of any one country, but the sample size is usually not large enough to analyze countries separately. Such problems with heterogeneity are related to not only economic evaluations but regulatory decisions as well [24]. The largest outcomes trials, however, are often international and can provide excellent opportunities for meaningful economic evaluation; many examples are available (e.g., [25–29]). Mitigation of the threats mentioned is typically handled by a careful combination of identifying country-specific resource types, cost drivers, and unit costs—often via consultation with country-specific clinical and economic experts during the design phase—and proper analytical approaches, as discussed subsequently.

#### Sample Size and Power

In an ideal world, the economic appraisal would be factored into sample size calculations using standard methods [30,31] based on asymptotic normality, value of information [32,33], or simulation [34]. It is common for the sample size of the trial, however, to be based on primary clinical outcomes alone. As a consequence, the economic comparisons can be underpowered. In cases in which researchers wish to set up formal hypotheses for economic

analyses, these should be stated a priori including the thresholds (e.g., \$50,000 or \$100,000/quality-adjusted life-year [QALY]) and the power to detect when incremental analysis meets or exceeds those thresholds [35]. Depending on the intended audience for the evaluation, power calculations may also be useful for cost-related end points.

Recently, economic investigators have become interested in Bayesian techniques for specifying the appropriate sample size for economic evaluation based on value-of-information methods [36–38]. These methods have become popular in the methodological literature, but have not yet been used routinely in practice. Although value of information is compelling at a theoretical level, and is worthy of mention in relation to some other areas covered below, there can be many practical challenges associated with its use [39].

### Study End Points and Comparators

The choice of the primary end point in a clinical study is unlikely to correspond with the ideal end point for economic evaluation. For example, the use of composite clinical end points is common in clinical trials (e.g., fatal events and nonfatal events combined) to provide greater statistical power. Cost per composite clinical end point, however, is often an unsatisfactory summary measure for an economic analysis, in part because the different outcomes are rarely of equal importance. In addition, economic evaluation of surrogate end points that are not well linked to final end points is not recommended (e.g., progression-free survival for many cancers). It is recommended that clinical end points used in economic evaluations be presented in disaggregated form, and that end points are included that can be used to determine value in a common measure that allows comparison across treatments and diseases.

Clinical end points that focus on the impact of a treatment on how a patient feels, functions, and survives are most useful for an economic evaluation [40]. We recommend weighting end points (e.g., by utilities) so that they yield a measure of QALYs in the case of CEA, or a monetary benefit measure in the case of cost-benefit analysis. For example, preference-weighted quality-of-life scores are typically collected within the trial at regular intervals, for example, every 3 months if there is an expectation of rapid change and less frequently when changes are expected to have reached a steady state. The resulting scores can be combined with data on survival to estimate QALYs.

If possible, intermediate or surrogate end points (e.g., percent cholesterol reduction) as the measure of benefit should be avoided. Intermediate outcome measures, however, are often used when the costs of conducting a long-term trial are prohibitive. When use of intermediate outcomes or surrogate end points is unavoidable, additional evidence is needed to link them with long-term costs and outcomes. If such a link is not reliable or is unavailable, the economic investigator should argue for follow-up sufficient to include clinically meaningful disease end points.

There is little that differentiates the choice of a comparator for trials versus other types of CEAs. As already noted, the ideal comparator likely is one that is widely used in practice. But even this relatively uncontroversial recommendation can be problematic for those trials for which there is disagreement about the acceptability or cost-effectiveness of potential comparators. In general, if there is flexibility within the trial as to the choice of a comparator, selecting one that most represents practice in the region(s) of interest is recommended.

### Appropriate Follow-Up

Economic analyses ideally include lifetime costs and outcomes of treatment. Yet, clinical trials rarely extend beyond a few years and are often conducted over much shorter periods. In practice, consideration of the follow-up period for the trial involves the

relationship between intermediate end points gathered in the short run and long-term disease outcomes—the stronger that relationship, the more a reliance on intermediate end points can be justified.

A key design consideration is to determine the appropriate points at which to gather medical resource use and data to measure health-related preference weights. Practical considerations include the use of scheduled contacts to collect information direct from patients, ensuring that data collection points coincide with expected changes in health-related quality of life (HRQOL) (where QALYs are to be estimated from attaching weights to clinical events) and that frequency is sufficient to minimize problems associated with patient recall [41,42]. It is recommended that baseline measures of HRQOL and medical resource use (e.g., at the point of enrollment in the trial) be collected. Any QALY analysis should adjust for any baseline imbalance in HRQOL [43], and baseline measures of medical resource use can be used to reduce the variance in incremental cost estimates [44]. Measurement of resources and preferences must be frequent enough to capture important changes in consumption and HRQOL, such that reasonable estimates of area under the curve can be constructed.

### Data Elements

Prospective collection of patient-level resource use and health preference-based data within a clinical trial requires careful planning. Decisions about which data elements to collect should be driven by their potential to affect the results of the study. One approach for identifying these elements is to conduct or review analyses of resource use patterns in routine practice to identify cost drivers and cost variation between patients. When resources and time allow, a decision model can be used to estimate the expected value of this (sample) information. Within the model, the economic investigator can apply a range of values for one or more model parameters expected to be generated from the trial, along with incremental costs for data collection, to compare the expected net benefit with new (sampled) information from the trial versus the expected net benefit with existing information [45,46]. In practice, however, economic investigators may not have been consulted early enough to allow adequate time or resources to carry out a value-of-information analysis.

### Patient-Level Data: Resource Use

Consistent with our previous report, we recommend prioritization of high-cost resources as well as those that are expected to differ between treatment arms, without distinction as to whether they are related to disease or intervention [44]. The scope of resources considered should include direct medical and nonmedical resources and indirect or productivity costs across patients and caregivers. Nonmedical resources and productivity effects can be particularly important if the intervention requires substantial commitments on the part of patients and caregivers (e.g., long travel times and intensive home therapy) relative to direct medical care. Their relevance will be driven by the study perspective(s) planned for the economic evaluation, as related to the study question. Although we acknowledge pragmatic pressures to streamline data collection in clinical trials, we caution against narrow collection of resources given that the treatment may have unanticipated effects and the trial may offer the last opportunity to collect these data within a randomized study design.

The frequency with which resource use data are collected should account for the levels of resource usage expected among patients enrolled in the trial, the ability to verify patient-reported data through electronic medical records or other sources, and the characteristics (e.g., cognitive abilities) of the trial participants. To



assist participants with accurate recall, economic investigators should consider using memory aids such as diaries to record medical visits and events, and should inform participants that they will be asked to report this information throughout the trial [47]. Based on our experience, it is often most feasible to schedule reporting for various types of resources over the same period of follow-up and coincide with trial visits. In any case, the time horizon over which resource use and cost are collected should be carefully defined and include the entire time period between collection points.

Resource use data are typically collected using a trial's case report form. The level of overlap between data elements crucial to both clinical and economic outcome assessments typically obviates consideration of separate data collection protocols. Given that most case report forms are customized for each trial, data collection processes are typically not validated and can lead to variability in the content and quality of the information [48]. This variability reduces the ability to pool resource use data across studies and to make direct comparisons between studies. To improve the quality and uniformity of data generated from trials, we recommend using validated instruments for resource data collection [49–51] or when incorporating productivity costs [52–54]; for resource data collection, use secondary sources when possible to confirm the utilization reported by patients.

#### **Patient-Level Data: Preference-Based Outcomes**

Information to derive individual-level preference weights should be collected from patients participating in the trial to generate QALYs for a cost-utility analysis. Preference-weighted health state classification systems are more widely used in clinical trials than are direct elicitation methods such as time trade-off or standard gamble because they are both easier to administer and are considered to yield a measure of preferences from the general public. Examples of these classification systems include the EuroQol five-dimensional questionnaire [55], including its newly developed five-level version, which may be more sensitive to changes in health status than is the three-level version [56,57], any of the three versions of the health utilities index [58–60], the Quality of Well-Being Scale [61,62], the six-dimensional health state short form (derived from short-form 36 health survey) [63,64], or the recently introduced Assessment of Quality of Life-8D [65]. As with resource use data, the frequency with which these measures are administered will depend on the medical condition under study and the expected timing of effects from the treatments being evaluated.

The literature on mapping disease-specific instruments to preference weights has grown extensively over the last several years, largely to address concerns that generic instruments do not represent relevant dimensions of health for specific conditions and are thus not responsive to changes that can be detected with disease-specific instruments [66,67]. Although we acknowledge the need for valid and reliable preference instruments that reflect important changes in health status across conditions, a wide range of methodological approaches to and concerns about this mapping remain. Many regression-based mapping algorithms exhibit poor model fit [68], generate biased estimates [69], and underestimate actual variance [70]. Few have undergone external validation [71,72]. Given these concerns, economic investigators may wish to consider coupling a disease-specific instrument with a generic instrument to allow for scenario analysis.

In cases in which preference weights based on generic or disease-specific instruments may not seem appropriate (e.g., interventions in which the process of care may affect utility), economic investigators must also consider whether to apply preference weights derived from traditional time trade-off and standard gamble exercises or newer valuation methods such as

discrete choice experiments [73,74]. For multinational trial analyses or adaptations, economic investigators should consider whether country-specific estimates may be desirable for the stakeholders who will review the results from the analysis [28,75].

#### **Patient-Level Data: Data Collection/Tracking**

Technological advances offer the possibility of tracking resource use, time estimates, and health status directly from trial participants via the Web, smartphones, or mobile health applications. These options offer the possibility of collecting relevant data more proximate to the time when clinical events occur, at more frequent intervals, or at random times throughout the day, potentially increasing accuracy and reducing recall bias [76]. Research is needed to validate and compare these technology-enabled data collection methods to traditional self-report and interview strategies [77,78].

#### **Provider-, Site-, and Jurisdiction-Level Data**

As the trend continues toward more multisite and more global representation in clinical trials, concerns about generalizability and transferability have grown [28,75,79]. To provide greater contextual information about participating providers and sites and to enable investigation into factors that may affect the net economic benefit of a given intervention across jurisdictions, we encourage economic investigators to collect additional provider-level, site-level, or country-level information on practice patterns and resource use.

#### **Valuation of Resources**

Trial-based economic evaluations require the ascertainment of unit costs (or price weights) to value resources. The specificity of unit costs will be dependent on the level of resource use data collected within the trial, the perspective of the study, the availability of the estimates, the time horizon, and the acquisition costs required to obtain the estimates [15,80]. For inpatient care, medical tests and procedures, bundled payments systems, or classifications such as Diagnosis Related Groups, Healthcare Resource Groupings, or *International Statistical Classification of Diseases and Related Health Problems* codes are often used to map resource use to appropriate unit costs. Although costs may be highly correlated between some coding systems [81], the level of specificity of coding systems varies for specific medical events, and economic investigators must be careful to consider the impact that different systems can have on a study's findings [82,83].

The approach to valuation/assignment of unit costs to medical and nonmedical resources requires trade-offs across accuracy, feasibility, generalizability, and cost [15,84]. In many cases, adjustment for differential timing, imputation for estimates that are not available, adaptation to different settings or coding systems (e.g., Diagnosis Related Groups to Healthcare Resource Grouping and *International Classification of Diseases, Ninth Revision* to *International Statistical Classification of Diseases, 10th Revision*), and conversion to different currencies may be necessary before assigning unit costs to quantities of resource. In other cases, particularly for new medical interventions, supplemental micro-costing exercises may be necessary to develop accurate cost estimates. Drummond et al. [15], Glick et al. [44], and Luce et al. [84] provide useful references on issues related to costing.

#### **Database Design and Management**

The full integration of clinical and economic data helps to ensure high-quality processes to obtain valid and complete data across study sites. Data capture for clinical trials is increasingly electronic and

streamlined. Most data capture systems require that sites manually enter data into an electronic case report form that is transmitted to a central trial database. These systems provide the opportunity for database programmers to apply real-time edit checks and queries to minimize errant data from sites and to alert sites to data elements that may be missing or overdue. We strongly recommend early and regular monitoring of economic data collection.

The next generation of electronic data capture includes the use of electronic health record data to (partially) populate clinical trial databases and the development of distributed data networks that do not require the data to be transferred to a central data repository for analysis [85,86]. Although these new models could drastically improve the efficiency of large-scale clinical trials, future studies will be needed to understand the meaning, quality, and completeness of data to support trial-based economic evaluations.

Informed consent for clinical trials should be inclusive of language to allow for collection of medical resource use and HRQOL data. In cases in which trial data can be augmented with hospital bills, health claims, or productivity reports for the economic evaluation, explicit language must be included in patient consent forms to allow the release of these data.

Whether data are collected by paper or electronically, consistency across data elements collected for clinical, economic, and safety end points is crucial. As such, early and ongoing collaboration among economic investigators carrying out disparate analyses is necessary to ensure that reliable results are generated across studies.

If trial investigators plan to release patient-level trial records at some point following the publication of trial results, consideration should be made as to whether economic data can also be released without compromising patient confidentiality. Such release of patient-level records would aid with subsequent modeling and meta-analyses.

## Analysis

### Guiding Principles

The analysis of economic measures should be guided by a data analysis plan. A prespecified plan is particularly important if formal tests of hypotheses are to be performed. Any tests of hypotheses that are not specified within the plan should be reported as exploratory. The plan should specify whether generalized linear model, least squares regression, or other multivariable analysis will be used to improve precision and to adjust for treatment group imbalances. The plan should also identify any selected subgroups and state the type of analysis, for example, intention-to-treat or modified intention-to-treat, that will be conducted. The plan should be finalized before trial data are unblinded; publication of the analysis plan before the completion of the trial is a best practice [87–89].

Although it is unlikely that all studies will use the same approaches for the analysis of resource use, cost, preference, and cost-effectiveness, there are several analysis features that should be common to all analyses of economic data derived from clinical trials:

1. The intention-to-treat population should be used for the primary analysis.
2. A common time horizon(s) should be used for accumulating costs and outcomes; a within-trial assessment of costs and outcomes should be conducted, even when also modeling or projecting beyond the time horizon of the trial.
3. An assessment of uncertainty is necessary for each measure (standard errors or confidence intervals for point estimates; *P* values for hypothesis tests).

4. A (common) real discount rate should be applied to future costs and, when used in a CEA, to future outcomes.
5. If data for some subjects are missing and/or censored, the analytic approach should address this issue consistently in the various analyses affected by missing data.

### Trial Costs

The purpose of clinical trial cost analysis is to estimate costs, cost differences associated with treatment, the variability of differences, and test whether the differences occurred by chance.

Once resources have been identified and valued, differences between groups must be summarized. Sample/arithmetical mean cost differences are generally considered the most appropriate and robust measure [90]. Nevertheless, cost data often do not conform to the assumptions for parametric statistical tests for comparing differences in arithmetic means [44,91–93]. They are usually right-skewed because it is impossible to incur costs less than zero and there are typically small numbers of high-resource-use patients. In addition, if subsets of data such as hospitalizations are being analyzed, there can be excessive numbers of participants with costs equaling zero. In most cases, the nonparametric bootstrap is an appropriate method to compare means and calculate confidence intervals [94–96] although parametric methods provide a unique solution and have been shown to provide coverage equal or superior to that provided by bootstrap methods [97,98].

Other common nonparametric tests (e.g. Kruskal-Wallis or Wilcoxon) compare medians (or other characteristics of the distribution of cost), but not means and thus are less appropriate [99–101]. Analyses of data that have been transformed to normalize the distribution can translate between-group differences in variance, kurtosis, and so forth into what appear to be differences in means. Retransformation to the original scale of costs must account for differences in variance and so forth and must include transformation of error terms [102–106].

The same distributional issues that affect univariate analysis of costs also affect use of costs as a dependent variable in multivariable analysis. The underlying distribution of costs should be carefully assessed to determine the most appropriate approach to draw inferences about or estimate between-group cost differences [107]. Generalized linear models commonly, but not necessarily, using a log link and gamma family should be considered for multivariable analysis of cost [103,108]. When most subjects in the study have nonzero costs, ordinary least squares can be used with other methods to provide a reference point or a check of robustness of results. If differences in resource use or subsets of costs are to be estimated, similar considerations regarding the appropriateness of statistical approaches based on distributional assumptions should be applied.

When study participants use large amounts of medical services that are unrelated to the disease or treatment under study, it may be difficult to detect the influence of the treatment on total health care costs. One approach to addressing this problem is to conduct secondary analyses that evaluate costs that are considered related to the disease or treatment under study. If such analyses are performed, it is important to prespecify services that were deemed “disease-related,” and to display costs for each component in the treatment and control arms.

### Trial Outcomes

When one of the trial’s clinical end points is also used as the outcome for the CEA (e.g., in-trial mortality), it is generally most transparent to adopt the methods used in the clinical analysis for the primary analysis plan, particularly if the clinical result is cited in product labeling or a publication. In some cases, the clinical analysis methods are not appropriate for economic analysis (e.g., the clinical

analysis focuses on relative treatment differences, whereas the economic analysis relies on absolute treatment differences, or the clinical analysis focuses on time to first event, whereas the economic analysis considers all events). Composite measures of outcome are also generally not appropriate for economic analyses. If other outcomes are used for the economic analysis, the linkage between clinical and economic measures should be clearly specified.

Using nonclinical effectiveness end points such as QALYs involves both construction and analysis. Health state utilities/preference scores, either collected directly from trial patients or imputed on the basis of observed health states, can be transformed into QALYs using standard area-under-the-curve methods [15,109]. As with the analysis of cost, multivariable analysis should be considered for drawing inferences about or estimating between-group differences in outcome. Generalized linear models, regression estimators based on features of the beta distribution [110], or limited dependent variable mixture models [111] should be considered for the analysis of preference scores. As stated previously, for the analysis of both costs and effects, explanatory variables should include baseline measures of costs or effects [43,44]. Analytic refinements may include adjusting for ceiling effects [112] and modeling of longitudinal effects [113,114].

Because failure to reject the hypothesis about the equality of two therapies is not the same as finding that outcomes of two therapies are identical, CEA should still be performed if the clinical study fails to demonstrate a statistically significant difference in clinical end points [115,116]. In such situations, a cost-minimization analysis can provide useful information to those who are interested in understanding the budget impact implications of the trial, but cost-minimization should not be the primary or sole form of analysis.

### Missing and Censored Data

Missing data are inevitable in economic analyses conducted alongside trials. Such data can include item-level missingness and missing because of censoring. In analyzing data sets with missing data, one must determine the nature of the missing data and then define an approach for dealing with the missing data. Missing data may bear no relation to observed or unobserved factors in the population (missing completely at random), may have a relationship with observed variables (missing at random), or may be related to unobserved factors (not missing at random) [44,117]. Eliminating cases with missing data is not recommended because it may introduce bias or severely reduce the power to test hypotheses. Nevertheless, ignoring small amounts of missing data (e.g., <5% of the observations) is acceptable if the amount and pattern of missing data are similar across treatment groups and a reasonable case can be made that doing so is unlikely to bias treatment group comparisons.

Imputation refers to replacing missing fields with estimates. If one chooses to impute missing data, most experts recommend multiple imputation approaches because they reflect the uncertainty that is inherent when replacing missing data [118–120]. Most commonly used statistical software packages include programs for imputation of missing data. A review of these programs can be found at <http://www.multiple-imputation.com> [121].

Censoring can be addressed with a number of approaches [122,123]. Most assume that censoring is either completely at random [124] or at random [123,125–128]. Nevertheless, nonrandom censoring is common, and external data sources for similar patients may be required to both identify and address it.

### Summary Measures

One or more summary measures should be used to characterize the relative value of treatments in the clinical trial. Three general

classes of summary measures are available that differ in how incremental costs and outcomes are combined into a single metric:

1. Ratio measures (e.g., incremental cost-effectiveness ratios) are obtained by dividing the incremental cost by the incremental health benefit.
2. Difference measures (e.g., net monetary benefits) rely on the ability to define a common metric (such as monetary units) by which both costs and outcomes can be measured [129–131].
3. Probability measures (e.g., acceptability curves) characterize the likelihood the new treatment will be deemed cost-effective based on incremental costs and outcomes [132,133].

The difference measures and probability measures are calculated for specific values of “willingness-to-pay” or cost-effectiveness thresholds. Because these values may not be known and/or vary among health care decision makers, one should evaluate the summary measure over a reasonable range of values.

### Uncertainty

Results of economic assessments in trials are subject to a number of sources of uncertainty, including sampling uncertainty, uncertainty in parameters such as unit costs and the discount rate, and—when missing data are present—imputation-related uncertainty.

#### Sampling Uncertainty

Because economic outcomes in trials are the result of a single sample drawn from the population, one should report the variability in these outcomes that arises from such sampling. Variability should be reported for within-group estimates of costs and outcomes, between-group differences in costs and outcomes, and the comparison of costs and outcomes. One approach for reporting this variability is to construct a confidence interval for the cost-effectiveness ratio or for net monetary benefit or to construct an acceptability curve. A second is to quantify the value of eliminating the uncertainty by estimation of the expected value of information.

Confidence intervals for cost-effectiveness ratios, confidence intervals for net monetary benefit, and acceptability curves are different, but related, measures that allow us to identify whether we can be confident that a therapy’s cost per unit of outcome, for example, a QALY, is less than one’s maximum willingness to pay. Both parametric and nonparametric methods can be used to construct these measures [44,134–136]. Decision makers can thus identify values of willingness to pay for which they 1) can be confident that the therapy is good value for the cost; 2) can be confident that the therapy is not good value; and 3) cannot be confident that the therapies’ values differ from one another. One advantage of the confidence interval for the cost-effectiveness ratio is that its limits define these values of willingness to pay; one advantage of the acceptability curve is that it defines these values for varying levels of confidence that range from 0% to 100%.

Value of information allows quantification of the value of eliminating the uncertainty that exists in the data [32,45,137]. Results of value-of-information analyses can help decision makers determine how much they should be willing to spend for research meant to increase our certainty, to prioritize research across disease areas, and to identify specific sources of uncertainty that are more and less important to target for further research [138,139].

#### Parameter Uncertainty

Uncertainty should be assessed for any parameter estimates that, when varied, have the potential to influence policy. Examples include unit costs and the discount rate. One approach to this assessment is sensitivity analysis: for example, if one uses a



discount rate of 3%, one may want to assess the impact of this assumption by repeating the analysis but using a 1% rate or a 5% rate.

A second approach is the assessment of the value of information related to parameter uncertainty (expected value of partial perfect information) [140]. Measures of sampling uncertainty and sensitivity analysis for parameter uncertainty are complements, not substitutes. Thus, when conducting sensitivity analysis, one should report both the revised point estimate and revised 95% confidence intervals that result from the sensitivity analysis.

#### *Imputation Uncertainty*

Finally, some methods used to address missing or censored data (e.g., use of an imputed mean) may artificially reduce estimates of sampling uncertainty. One should make efforts to address this shrinkage when reporting sampling uncertainty, for example, by bootstrapping the entire imputation and estimation process.

#### *Estimating Country-Specific Costs for Multinational Studies*

It is common to apply country-specific unit costs for pooled trial resource use to estimate country-specific costs. In practice, this approach yields few qualitative differences in summary measures of cost-effectiveness among countries with similar levels of economic development but may not adjust for important country-specific differences [141,142]. Rather, intercountry differences in population characteristics and treatment patterns are more likely to influence summary measures between countries rather than differences in unit costs. Recommended approaches to address this issue include [141,143–145]

1. hypothesis tests of homogeneity of results across countries (and adjusting the resource use in other countries to better match those seen in country X);
2. multivariable cost or outcome regressions to adjust for country effects (e.g., include country dummies or adjusted gross national product per capita as covariates); and
3. multilevel random effects model with shrinkage estimators.

All three methods allow for the inclusion of site-specific characteristics that are recommended in the “Provider-, Site-, and Jurisdiction-Level Data” section.

#### *Including Costs and Effects beyond the Time Horizon of the Trial*

The cost-effectiveness observed within the trial may be substantially different from what would have been observed with continued follow-up. Well established, published models (preferred) or those developed specifically for the trial are used to project costs and outcomes that could have been observed had observation been prolonged. When modeling beyond the follow-up period for the trial, it is important to project costs and outcomes over the expected duration of treatment and its effects.

Direct modeling of long-term costs and outcomes is feasible when the trial period is long enough, or if at least a subset of patients are observed for a longer time and provide a basis for estimating other patients' outcomes. A number of approaches are feasible [146]. Parametric survival models estimated on trial data are generally recommended for such projections, unless models based on other data or methods can be justified [147,148].

In cases in which such direct modeling is not feasible, it may be possible to “marry” trial data to long-term observational data in a model. In either case, good modeling practices should be followed. Examples of projection models used for trials include those from clopidogrel versus aspirin in patients at risk of ischaemic events [149], Global Utilization of Streptokinase and

Tissue Plasminogen Activator for Occluded Coronary Arteries [150], Eplerenone's neuroHormonal Efficacy and SURvival Study [27], Sudden Cardiac Death in Heart Failure Trial [151], and the National Emphysema Treatment Trial [152]. The reader is referred to the consensus position of the ISPOR Modeling Good Research Practices Task Force reports for discussion of modeling issues [153–159].

Cost-effectiveness ratios should be calculated at various time horizons (e.g., 2, 5, 10 years, or as appropriate for the episode of the disease), both to accommodate the needs of decision makers and to provide a “trajectory” of summary measures over time. The effects of long-term health care costs (i.e., after the episode of care) not directly related to treatment should be taken into account as well as possible [160]. As always, assumptions used must be described and justified, and the uncertainty associated with projections must be taken into account.

#### *Subgroup Analysis*

Proper subgroup analysis can be vital to decision makers [161] and should be prespecified. Nonetheless, the dangers of spurious subgroup effects are well known. For example, the probability of finding a statistically significant difference due solely to random variation increases with the number of differences examined unless the alpha level is adjusted appropriately. The focus should be on testing treatment interactions on the absolute scale, with a justification for choice of scale used. In cases in which prespecified clinical interactions are significant, subgroup analyses may be justified. Methods [162] include stratification by subgroup [163], n-of-1 trials [164], latent mixture models [165], and Bayesian methods [166].

### **Reporting the Methods and Results**

Economic analysis has various audiences. Correspondingly, detailed and comprehensive information on the methods and results should be available to interested readers in a format that facilitates interpretation. Journal word limits often necessitate parsimony in reporting; therefore, we recommend that in addition to the main report, detailed technical appendices be made available online.

A number of organizations including ISPOR have developed minimum reporting standards for clinical trials and economic analyses [167–171]. The principles and suggested format in these guidance documents should be adhered to in the reporting of economic studies. We have highlighted issues of particular importance to economic studies conducted alongside clinical trials.

Reporting of the methods and results should include elements identified in the Consolidated Health Economic Evaluation Reporting Standards statement [169,170].

#### *Trial-Related Issues*

1. A brief, general description of the clinical trial, including patients' demographic characteristics, trial setting (e.g., country, tertiary care hospital), inclusion and exclusion criteria and protocol-driven procedures that influence external validity, intervention and control arms, time horizon for the intervention and follow-up, and a link to the registry posting of the trial (e.g., clinicaltrials.gov) and
2. Key clinical findings.

#### *Data for the Economic Study*

1. Delineation between data collected as part of the trial versus data collected outside of the trial;



2. Description of all outcomes assessments and schedule of data collection;
3. Source of unit costs, published utility weights; and
4. Amount of missing and censored data.

### Methods of Analysis

1. Construction of costs and outcomes, including the discount rate used;
2. In cases in which the main clinical end point is used in the denominator of the incremental cost-effectiveness ratio and different methods were used to analyze this end point in the clinical and economic analyses, any differences in the point estimates should be explained;
3. Methods for addressing missing and censored data;
4. Statistical methods used to compare resource use, costs, and outcomes;
5. Methods and assumptions used to project costs and outcomes beyond the trial period; and
6. Deviations from the prespecified analysis plan and justification for these changes.

### Results

1. Resource use, costs, outcome measures, including point estimates and measures of uncertainty;
2. Results within the time horizon of the trial;
3. Results with projections beyond the trial (if conducted);
4. Graphical displays of results not easily reported in tabular form (e.g., cost-effectiveness acceptability); and
5. Curves (joint density of incremental costs and outcomes).

For further guidance on reporting standards, please see Consolidated Health Economic Evaluation Reporting Standards (CHEERS)—Explanation and elaboration: A report of the ISPOR Health Economic Evaluations Publication Guidelines Good Reporting Practices Task Force [169,171].

Finally, data from economic analyses performed in the context of trials may also be used in independent cost-effectiveness models based on decision analysis or meta-analyses [6]. To facilitate synthesis of economic information from multiple trials, authors should report means and standard errors for incremental costs and outcomes and their correlation.

### Conclusions

Since publication of the first ISPOR RCT-CEA Task Force report 10 years ago, the role and use of economics in health care coverage and reimbursement decisions have increased substantially around the world. The focus of this report was methods for CEA conducted alongside randomized clinical trials designed to test the efficacy or effectiveness of drugs, devices, surgical procedures, or screening interventions, including pragmatic trials. This report provides guidance on the methods, with the goal of making trial-based economic evaluation as useful as possible for decision makers.

Most of the recommendations from the 2005 report remain valid and appropriate. This report emphasizes innovations in several areas, including the use of value-of-information techniques, the increased role of diverse stakeholders in study design, the use of newer multivariate methods, and standards for modeling outcomes beyond the trial observation period.

This task force maintains that when designed, analyzed, and interpreted appropriately, economic evaluations alongside randomized clinical trials are important sources of information

for decision makers. To be useful as a stand-alone evaluation, however, a trial must be designed to represent the population, duration of treatment, clinical practice, types of outcomes, and other factors most relevant to the clinical situation to which the decision is being applied. Similar caveats generally apply to the clinical results of the trial. When these conditions are not satisfied, modeling-based corrections can often adapt the results to the conditions appropriate for the decision context.

Methods for designing, conducting, and reporting economic analyses alongside RCTs will continue to evolve and improve over time, reflecting changes in knowledge and the evolving needs of decision makers. As these methods are identified, tested, and validated, they will be included in future versions of this guidance document [11].

### Acknowledgments

We gratefully acknowledge the individual contribution of Andrew Willan, MSc, PhD, Senior Scientist at SickKids Research Institute, and Professor, Biostatistics, Dalla Lana School of Public Health, University of Toronto, Toronto, Ontario, Canada.

In addition, we thank the following reviewers for valuable written feedback on earlier drafts of this report: Anita Burrell, John Cook, Francis Fatoye, Parul Gupta, Marlene Gyldmark, Alex Hakuzimana, Jinhai (Stephen) Huo, Don Husereau, Ivan Kocic, Paul C. Langley, Robert Launois, Bengt Lilius, Jo Mauskopf, Kimberly A. McGuigan, Belinda Acuña Mohr, Marjorie Neidecker, Kaganda Paschal Nyendo, Margaret K. Pasquale, Elvira Bel Prieto, Brian Seal, Cristina Teruzzi, Randy Vogenberg, Wendy Wan, and Lonnie Wen. ISPOR member comments contribute to the high-quality consensus nature that characterizes ISPOR Good Practices for Outcomes Research reports.

Finally, the coauthors thank Elizabeth Molsen, Director, Scientific & Health Science Policy Initiatives, for her support on this project.

### REFERENCES

- [1] Berntgen M, Gourvil A, Pavlovic M, et al. Improving the contribution of regulatory assessment reports to health technology assessments—a collaboration between the European Medicines Agency and the European Network for Health Technology Assessment. *Value Health* 2014;15:634–41.
- [2] Jönsson B. Relative effectiveness and the European pharmaceutical market. *Eur J Health Econ* 2011;12:97–102.
- [3] Luce BR, Drummond MF, Dubois RW, et al. Principles for planning and conducting comparative effectiveness research. *J Comp Eff Res* 2012;1:431–40.
- [4] Lauer MS, D'Agostino RB Sr. The randomized registry trial—the next disruptive technology in clinical research? *N Engl J Med* 2013;369:1579–81.
- [5] Ramsey SD, Willke RJ, Briggs AH, et al. Good research practices for cost-effectiveness analysis alongside clinical trials: the ISPOR RCT-CEA Task Force report. *Value Health* 2005;8:521–33.
- [6] Gold MR, Siegel JE, Russell LB, Weinstein M. *Cost-Effectiveness in Health and Medicine*. New York: Oxford University Press, 1996.
- [7] EUnetha. HTA Core Model® for rapid relative effectiveness assessment of pharmaceuticals. Available from: <http://www.eunetha.eu/outputs/new-application-hta-core-model-hta-core-model-rapid-relative-effectiveness-assessment-pharma>. [Accessed December 19, 2014].
- [8] Schwartz D, Lellouch J. Explanatory and pragmatic attitudes in therapeutic trials. *J Chronic Dis* 1967;20:637–48.
- [9] O'Brien B. Economic evaluation of pharmaceuticals: Frankenstein's monster or vampire of trials? *Med Care* 1996;34:DS99–108.
- [10] Schulpher MJ, Claxton K, Drummond M, McCabe C. Whither trial-based economic evaluations for health care decision making? *Health Econ* 2006;15:677–87.
- [11] Backhouse ME, Wonder M, Homby E, et al. Early dialogue between the developers of new technologies and pricing and reimbursement agencies: a pilot study. *Value Health* 2011;14:608–15.

- [12] Haute Autorite de Sante. Shaping European Early Dialogues for health technologies (SEED). Available from: [http://www.has-sante.fr/portail/jcms/c\\_1700958/fr/seed-shaping-european-early-dialogues-for-health-technologies](http://www.has-sante.fr/portail/jcms/c_1700958/fr/seed-shaping-european-early-dialogues-for-health-technologies). [Accessed December 19, 2014].
- [13] European Medicines Agency-EUnetHTA. EMA-EUnetHTA three-year work plan 2013-2015. Available from: [http://www.ema.europa.eu/docs/en\\_GB/document\\_library/Other/2013/11/WC500154588.pdf](http://www.ema.europa.eu/docs/en_GB/document_library/Other/2013/11/WC500154588.pdf). [Accessed December 19, 2014].
- [14] Eichler H-G, Bloechl-Daum B, Abadie E, et al. Relative efficacy of drugs: an emerging issue between regulatory agencies and third party payers. *Nat Rev Drug Discovery* 2010;9:277–91.
- [15] Drummond MF, Sculpher MJ, Torrance GW, et al. *Methods for the Economic Evaluation of Health Care Programmes*. (3rd ed.). Oxford: Oxford University Press, 2005.
- [16] O'Brien BJ. A tale of two or more cities: geographic transferability of pharmacoeconomic data. *Am J Man Care* 1997;3:S33–9.
- [17] Schulman K, Burke J, Drummond M, et al. Resource costing for multinational neurologic clinical trials: methods and results. *Health Econ* 1998;7:629–38.
- [18] Koopmanschap MA, Touw KCR, Rutten FFH. Analysis of costs and cost-effectiveness in multinational trials. *Health Policy* 2001;58:175–86.
- [19] Gosden TB, Torgerson DJ. Converting international cost-effectiveness data to UK prices. *BMJ* 2002;325:275–6.
- [20] Sullivan SD, Liljas B, Buxton MJ, et al. Design and analytic considerations in determining the cost-effectiveness of early intervention in asthma from a multinational clinical trial. *Control Clin Trials* 2001;22:420–37.
- [21] Buxton MJ, Sullivan SD, Andersson FL, et al. Country-specific cost-effectiveness of early intervention with budesonide in mild asthma. *Eur Respir J* 2004;24:568–74.
- [22] Gerdtham UG, Jönsson B. Conversion factor instability in international comparisons of health care expenditure. *J Health Econ* 1991;10:227–34.
- [23] World Bank. *World Development Indicators 2001*. Washington, DC: The World Bank, 2001.
- [24] European Medicines Agency. *Guideline on the investigation of subgroups in 4 confirmatory clinical trials*, EMA/CHMP/539146/2013. 2014. Available from: [http://www.ema.europa.eu/docs/en\\_GB/document\\_library/Scientific\\_guideline/2014/02/WC500160523.pdf](http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2014/02/WC500160523.pdf). [Accessed December 19, 2014].
- [25] Briggs AH, Glick HA, Lozano-Ortega G, et al. Towards a revolution in COPD health (TORCH) investigators. Is treatment with ICS and LABA cost-effective for COPD? Multinational economic analysis of the TORCH study. *Eur Respir J* 2010;35:532–9.
- [26] Sullivan SD, Buxton M, Andersson LF, et al. Cost-effectiveness analysis of early intervention with budesonide in mild persistent asthma. *J Allergy Clin Immunol* 2003;112:1229–36.
- [27] Weintraub WS, Zhang Z, Mahoney EM, et al. Cost-effectiveness of eplerenone compared with placebo in patients with myocardial infarction complicated by left ventricular dysfunction and heart failure. *Circulation* 2005;111:1106–13.
- [28] Drummond M, Barbieri M, Cook J, et al. Transferability of economic evaluations across jurisdictions: ISPOR Good Research Practices Task Force report. *Value Health* 2009;12:409–18.
- [29] Pinto EM, Willan AR, O'Brien BJ. Cost-effectiveness analysis for multinational clinical trials. *Stat Med* 2005;24:1965–82.
- [30] Laska EM, Meisner M, Siegel C. Power and sample size in cost-effectiveness analysis. *Med Decis Making* 1999;19:339–43.
- [31] Briggs A, Tambour M. The design and analysis of stochastic cost-effectiveness studies for the evaluation of health care interventions. *Drug Inf J* 2001;35:1455–68.
- [32] Koerkamp BJ, Spronk S, Stijnen T, Hunink MGM. Value of information analyses and economic randomized controlled trials: the treatment of intermittent claudication. *Value Health* 2010;13:242–50.
- [33] Willan AR, Eckermann S. Optimal clinical trial design using value of information methods with imperfect information. *Health Econ* 2010;19:549–61.
- [34] Al MJ, van Hout B, Michel BC, et al. Sample size calculation in economic evaluations. *Health Econ* 1998;7:327–35.
- [35] Gardner MJ, Altman DG. Estimation rather than hypothesis testing: confidence intervals rather than P values. In: Gardner MJ, Altman DG, eds. *Statistics with Confidence*. London: BMJ, 1989.
- [36] Claxton K, Posnett J. An economic approach to clinical trial design and research priority setting. *Health Econ* 1996;5:513–24.
- [37] Willan AR, Pinto EM. The expected value of information and optimal clinical trial design. *Stat Med* 2005;24:1791–806.
- [38] Willan AR. Optimal sample size determinations from an industry perspective based on the expected value of information. *Clin Trials* 2008;5:587–94.
- [39] Cooke JA, Hislop J, Adewuyi TE, et al. Assessing methods to specify the target difference for a randomised controlled trial: DELTA (Difference Elicitation in Trials) review. *Health Technol Assess* 2014:18.
- [40] EUnetHTA. *GUIDELINE—endpoints used for relative effectiveness assessment of pharmaceuticals: CLINICAL ENDPOINTS*. 2013. Available from: <http://www.eunetha.eu/sites/5026.fedimbo.belgium.be/files/Clinical%20endpoints.pdf>. [Accessed December 19, 2014].
- [41] Bhandari A, Wagner T. Self-reported utilization of health care services: improving measurement and accuracy. *Med Care Res Rev* 2006;63:217–35.
- [42] Evans C, Crawford B. Patient self-reports in pharmacoeconomic studies. *Pharmacoeconomics* 1999;15:241–56.
- [43] Manca A, Hawkins N, Sculpher MJ. Estimating mean QALYS in trial-based cost-effectiveness analysis: the importance of controlling for baseline utility. *Health Econ* 2005;14:487–96.
- [44] Glick HA, Doshi JA, Sonnad SS, Polsky D. *Economic Evaluation in Clinical Trials*. (2nd ed.). New York: Oxford University Press, 2015.
- [45] Claxton K. The irrelevance of inference: a decision-making approach to the stochastic evaluation of health care technologies. *J Health Econ* 1999;18:341–64.
- [46] Ades AE, Lu G, Claxton K. Expected value of sample information calculations in medical decision modeling. *Med Decis Making* 2004;24:207–27.
- [47] Goossens ME, Rutten-van Mölken MP, Vlaeyen JW, et al. The cost diary: a method to measure direct and indirect costs in cost-effectiveness research. *J Clin Epidemiol* 2000;53:688–95.
- [48] Chernyak N, Ernsting C, Icks A. Pre-test of questions on health-related resource use and expenditure, using behaviour coding and cognitive interviewing techniques. *BMC Health Serv Res* 2012;12:303.
- [49] Tang K. Estimating productivity costs in health economic evaluations: a review of instruments and psychometric evidence. *Pharmacoeconomics* 2015;33:31–48.
- [50] DIRUM. *Database of Instruments for Resource Use Measurement*. Available from: <http://www.dirum.org>. [Accessed July 31, 2014].
- [51] Ridyard CH, Hughes DA, DIRUM Team. Development of a database of instruments for resource-use measurement: purpose, feasibility, and design. *Value Health* 2012;15:650–5.
- [52] Reilly MC, Zrobek AS, Dukes EM. The validity and reproducibility of a work productivity and activity impairment instrument. *Pharmacoeconomics* 1993;4:353–65.
- [53] Lerner D, Amick BC III, Rogers WH, et al. The work limitations questionnaire. *Med Care* 2001;39:72–85.
- [54] Prasad M, Wahlquist P, Shikier R, et al. A review of self-report instruments measuring health-related work productivity. *Pharmacoeconomics* 2004;22:225–44.
- [55] Dolan P. Modelling valuations in EuroQol health states. *Med Care* 1997;35:1095–108.
- [56] Janssen MF, Pickard AS, Golicki D, et al. Measurement properties of the EQ-5D-5L compared to the EQ-5D-3L across eight patient groups: a multi-country study. *Qual Life Res* 2013;22:1717–27.
- [57] Herdman M, Gudex C, Lloyd A, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Qual Life Res* 2011;20:1727–36.
- [58] Torrance GW, Feeny DH, Furlong WJ, et al. Multi-attribute utility function for a comprehensive health status classification system. *Health Utilities Index Mark 2*. *Med Care* 1996;34:702–22.
- [59] Torrance GW, Boyle MH, Horwood SP. Application of multi-attribute utility theory to measure social preferences for health states. *Oper Res* 1982;30:1043–69.
- [60] Feeny D, Furlong W, Torrance GW, et al. Multi-attribute and single-attribute utility functions for the health utilities index mark 3 system. *Med Care* 2002;40:113–28.
- [61] Kaplan RM, Anderson JP. A general health policy model: update and applications. *Health Serv Res* 1988;23:203–35.
- [62] Kaplan RM, Anderson JP. The general health policy model: an integrated approach. In: Spiker B, ed. *Quality of Life and Pharmacoeconomics in Clinical Trials*. Philadelphia: Lippincott-Raven, 1996.
- [63] Brazier J, Roberts J, Deverill M. The estimation of a preference-based measure of health from the SF-36. *J Health Econ* 2002;21:271–92.
- [64] Craig BM, Pickard AS, Stolk E, et al. US valuation of the SF-6D. *Med Decis Making* 2013;33:793–803.
- [65] Richardson J, Iezzi A, Khan MA, Maxwell A. Validity and reliability of the Assessment of Quality of Life (AQoL)-8D multi-attribute utility instrument. *Patient* 2014;7:85–96.
- [66] Brazier JE, Yang Y, Tsuchiya A, et al. A review of studies mapping (or cross walking) non-preference based measures of health to generic preference-based measures. *Eur J Health Econ* 2010;11:215–25.
- [67] Mortimer D, Segal L. Comparing the incomparable? A systematic review of competing techniques for converting descriptive measures of health status into QALY-weights. *Med Decis Making* 2008;28:66–89.
- [68] Hernández Alava M, Wailoo A, Wolfe F, et al. A comparison of direct and indirect methods for the estimation of health utilities from clinical outcomes. *Med Decis Making* 2013;34:919–30.

- [69] Fayers PM, Hays RD. Should linking replace regression when mapping from profile-based measures to preference-based measures? *Value Health* 2014;17:261–5.
- [70] Chan KK, Willan AR, Gupta M, et al. Underestimation of uncertainties in health utilities derived from mapping algorithms involving health-related quality-of-life measures: statistical explanations and potential remedies. *Med Decis Making* 2014;34:863–72.
- [71] Chuang L-H, Whitehead SJ. Mapping for economic evaluation. *Br Med Bull* 2012;101:1–15.
- [72] McTaggart-Cowan H, Teckle P, Peacock S. Mapping utilities from cancer-specific health-related quality of life instruments: a review of the literature. *Expert Rev Pharmacoecon Outcomes Res* 2013;13:753–65.
- [73] Norman R, Viney R, Brazier J, et al. Valuing SF-6D health states using a discrete choice experiment. *Med Decis Making* 2013;34:773–86.
- [74] Bansback N, Brazier J, Tsuchiya A, et al. Using a discrete choice experiment to estimate health state utility values. *J Health Econ* 2012;31:306–18.
- [75] Barbieri M, Drummond M, Rutten F, et al. ISPOR Good Research Practices Economic Data Transferability Task Force. What do international pharmacoeconomic guidelines say about economic data transferability? *Value Health* 2010;13:1028–37.
- [76] Marcano Belisario JS, Huckvale K, Saje A, et al. Comparison of self-administered survey questionnaire responses collected using mobile apps versus other methods. *The Cochrane Library* 2014;Issue 4. Available from: <http://onlinelibrary.wiley.com/doi/10.1002/14651858.MR000042/abstract>. [Accessed July 31, 2014].
- [77] Coons SJ, Gwaltney CJ, Hays RD, et al. Recommendations on evidence needed to support measurement equivalence between electronic and paper based patient-reported outcome (PRO) measures: ISPOR ePRO Good Research Practices Task Force report. *Value Health* 2009;12:419–29.
- [78] Bowling A. Mode of questionnaire administration can have serious effects on data quality. *J Pub Health* 2005;27:281–91.
- [79] Vemer P, Rutten-van Mölken MP. The road not taken: transferability issues in multinational trials. *Pharmacoeconomics* 2013;31:863–76.
- [80] Hay JW, Smeeding J, Carroll NV, et al. Good research practices for measuring drug costs in cost effectiveness analyses: issues and recommendations: the ISPOR Drug Cost Task Force report—part I. *Value Health* 2010;13:3–7.
- [81] Johnston K, Buxton MJ, Jones DR, Fitzpatrick R. Assessing the costs of healthcare technologies in clinical trials. *Health Technol Assess* 1999;3:1–76.
- [82] Cots F, Elvira D, Castells X, et al. Medicare's DRG-weights in a European environment: the Spanish experience. *Health Policy* 2000;51:31–47.
- [83] Quentin W, Rätto H, Peltola M, et al. EuroDRG group. Acute myocardial infarction and diagnosis-related groups: patient classification and hospital reimbursement in 11 European countries. *Eur Heart J* 2013;34:1972–81.
- [84] Luce BR, Manning WG, Siegel JE, et al. Report of the Panel on Cost-Effectiveness in Health and Medicine. In: Gold MR, Siegel JE, Russell LB, Weinstein MC, eds. *Cost-Effectiveness in Health and Medicine*. New York, NY: Oxford University Press, 1996.
- [85] Curtis LH, Brown J, Platt R. Four health data networks illustrate the potential for a shared national multipurpose big-data network. *Health Aff (Millwood)* 2014;33:1178–86.
- [86] Maro JC, Platt R, Holmes JH, et al. Design of a national distributed health data network. *Ann Intern Med* 2009;151:341–4.
- [87] Ramsey SD, Sullivan SD, Kaplan RM, et al. Economic analysis of lung volume reduction surgery as part of the National Emphysema Treatment Trial. NETT Research Group. *Ann Thorac Surg* 2001;71:995–1002.
- [88] Sullivan SD, Liljas B, Buxton M, et al. START Steering Committee. Design and analytic considerations in determining the cost-effectiveness of early intervention in asthma from a multinational clinical trial. *Control Clin Trials* 2001;22:420–37.
- [89] Spertus JA, Tooley J, Jones P, et al. Expanding the outcomes in clinical trials of heart failure: the quality of life and economic components of EPHEUS (EPlerenone's neuroHormonal Efficacy and SURvival Study). *Am Heart J* 2002;143:637–42.
- [90] Barber JA, Thompson SG. Analysis and interpretation of cost data in randomised controlled trials: review of published studies. *BMJ* 1998;317:1195–200.
- [91] O'Hagan A, Stevens JW. Assessing and comparing costs: how robust are the bootstrap and methods based on asymptotic normality? *Health Econ* 2003;12:33–49.
- [92] Thompson SG, Barber JA. How should cost data in pragmatic randomised trials be analysed? *BMJ* 2000;320:1197–9.
- [93] Briggs A, Gray A. The distribution of health care costs and their statistical analysis for economic evaluation. *J Health Serv Res Policy* 1998;3:233–45.
- [94] Efron B, Tibshirani R. *An Introduction to the Bootstrap*. New York: Chapman and Hall, 1993.
- [95] Desgagne A, Castilloux AM, Angers JF, et al. The use of the bootstrap statistical method for the pharmacoeconomic cost analysis of skewed data. *Pharmacoeconomics* 1998;13:487–97.
- [96] Barber JA, Thomson SG. Analysis of cost data in randomized trials: an application of the nonparametric bootstrap. *Stat Med* 2000;19:3219–36.
- [97] Gomes M, Ng ESW, Grieve RD, et al. Developing appropriate methods for cost-effectiveness analysis of cluster randomized trials. *Med Decis Making* 2012;32:350–61.
- [98] Nixon RM, Wonderling D, Grieve RD. Non-parametric methods for cost-effectiveness analysis: the central limit theorem and the bootstrap compared. *Health Econ* 2010;19:316–33.
- [99] Briggs AH, Gray AM. Handling uncertainty when performing economic evaluation of health care interventions. *Health Technol Assess* 1999;3:1–134.
- [100] Barber JA, Thompson SG. Analysis of cost data in randomized trials: an application of the nonparametric bootstrap. *Stat Med* 2000;19:3219–36.
- [101] Briggs A, Clarke P, Polsky D, et al. Modelling the cost of health care interventions. Paper prepared for DEEM III: Costing Methods for Economic Evaluation. University of Aberdeen, April 15–16, 2003.
- [102] Manning WG. The logged dependent variable, heteroscedasticity, and the retransformation problem. *J Health Econ* 1998;17:283–95.
- [103] Manning WG, Mullahy J. Estimating log models: to transform or not to transform? *J Health Econ* 2001;20:461–94.
- [104] Bland JM, Altman DG. The use of transformation when comparing two means. *BMJ* 1996;312:1153.
- [105] Bland JM, Altman DG. Transformations, means, and confidence intervals. *BMJ* 1996;312:1079.
- [106] Bland JM, Altman DG. Transforming data. *BMJ* 1996;312:770.
- [107] White IR, Thompson SG. Choice of test for comparing two groups, with particular application to skewed outcomes. *Stat Med* 2003;22:1205–15.
- [108] Blough DK, Madden CW, Hornbrook MC. Modeling risk using generalized linear models. *J Health Econ* 1999;18:153–71.
- [109] Patrick DL, Erickson P. Ranking costs and outcomes of health care alternatives. In: *Health Status and Health Policy: Quality of Life in Health Care Evaluation and Resource Allocation*. New York: Oxford University Press, 1993.
- [110] Basu A, Rathouz PJ. Estimating marginal and incremental effects on health outcomes using flexible link and variance function models. *Biostat* 2005;6:93–109.
- [111] Hernandez Alava M, Wailoo AJ, Ara R. Tails from the peak district: adjusted limited dependent variable mixture models of EQ-5D questionnaire health state utility values. *Value Health* 2012;15:550–61.
- [112] Austin PC. A comparison of methods for analyzing health-related quality of life measures. *Value Health* 2002;4:329–37.
- [113] Gray SM, Brookmeyer R. Estimating a treatment effect from multidimensional longitudinal data. *Biometrics* 1998;54:976–88.
- [114] Hedeker D, Gibbons RD. Application of random effects pattern-mixture models for missing data in longitudinal studies. *Psychol Methods* 1997;2:64–78.
- [115] Briggs AH, O'Brien BJ. The death of cost-minimization analysis? *Health Econ* 2001;10:179–84.
- [116] Dakin H, Wordsworth S. Cost-minimisation analysis versus cost-effectiveness analysis, revisited. *Health Econ* 2013;22:22–34.
- [117] Little RJA, Rubin DB. *Statistical Analysis with Missing Data*. New York: John Wiley & Sons, 1987.
- [118] Rubin DB. *Multiple Imputation for Nonresponse in Surveys*. New York: John Wiley, 1987.
- [119] Briggs A, Clark T, Wolstenholme J, et al. Missing ... presumed at random: cost-analysis of incomplete data. *Health Econ* 2003;12:377–92.
- [120] Schafer JL. Multiple imputation: a primer. *Stat Methods Med Res* 1999;8:3–15.
- [121] Horton NJ, Lipsitz SR. Multiple imputation in practice: comparison of software packages for regression models with missing variables. *Am Statistician* 2001;55:244–54.
- [122] Raikou M, McGuire A. Estimating costs for economic evaluation. In: Jones AM, ed. *The Elgar Companion to Health Economics*. Cheltenham, UK: Edward Elgar, 2006; chap 40.
- [123] O'Hagan A, Stevens JW. On estimators of medical costs with censored data. *J Health Econ* 2004;23:615–25.
- [124] Lin DY, Feuer EJ, Etzioni R, et al. Estimating medical costs from incomplete follow-up data. *Biometrics* 1997;53:419–34.
- [125] Bang H, Tsiatis AA. Median regression with censored cost data. *Biometrics* 2002;58:643–9.
- [126] Etzioni RD, Feuer EJ, Sullivan SD, et al. On the use of survival analysis techniques to estimate medical care costs. *J Health Econ* 1999;18:365–80.
- [127] Carides GW, Heyse JF, Iglewicz B. A regression based method for estimating mean treatment cost in the presence of right-censoring. *Biostatistics* 2000;1:299–313.
- [128] Raikou M, McGuire A. Estimating medical care costs under conditions of censoring. *J Health Econ* 2004;23:443–70.



- [129] Stinnett AA, Mullahy J. Net health benefits: a new framework for the analysis of uncertainty in cost effectiveness analysis. *Med Decis Making* 1998;18(Suppl. 2):S68–80.
- [130] Tambour M, Zethraeus N, Johannesson M. A note on confidence intervals in cost-effectiveness analysis. *Int J Technol Assess Health Care* 1998;14:467–71.
- [131] Willan AR, Lin DY. Incremental net benefit in randomized clinical trials. *Stat Med* 2001;20:1563–74.
- [132] Van Hout BA, Al MJ, Gordon GS, et al. Costs, effects and C/E-ratios alongside a clinical trial. *Health Econ* 1994;3:309–19.
- [133] Fenwick E, Claxton K, Sculpher M. Representing uncertainty: the role of cost-effectiveness acceptability curves. *Health Econ* 2001;10:779–87.
- [134] Fieller EC. Some problems in interval estimation with discussion. *J Roy Stat Soc Series B* 1954;16:175–88.
- [135] Willan AR, O'Brien BJ. Confidence intervals for cost-effectiveness ratios: an application of Fieller's theorem. *Health Econ* 1996;5:297–305.
- [136] Heitjan DF. Fieller's theorem and net health benefits. *Health Econ* 2000;9:327–35.
- [137] Claxton K. Bayesian approaches to the value-of-information: implications for the regulation of new pharmaceuticals. *Health Econ* 1999;8:269–74.
- [138] Eckerman S, Willan AR. Expected value of information and decision making in HTA. *Health Econ* 2007;16:195–209.
- [139] Carlson JJ, Thariani R, Roth J, et al. Value-of-information analysis within a stakeholder-driven research prioritization process in a US setting: an application in cancer genomics. *Med Decis Making* 2013;33:463–71.
- [140] Griffin S, Welton NJ, Claxton K. Exploring the research decision space: the expected value of information for sequential research designs. *Med Decis Making* 2009;30:155–62.
- [141] Willke RJ, Glick HA, Polsky D, et al. Estimating country-specific cost-effectiveness from multinational clinical trials. *Health Econ* 1998;7:481–93.
- [142] Barbieri M, Drummond M, Willke RJ, et al. Variability of cost-effectiveness estimates for pharmaceuticals in Western Europe: lessons for inferring generalizability. *Value Health* 2005;8:10–23.
- [143] Jönsson B, Weinstein M. Economic evaluation alongside multinational clinical trials: study considerations for GUSTO IIb. *Int J Technol Assess Health Care* 1997;13:49–58.
- [144] Cook JR, Drummond M, Glick H, et al. Assessing the appropriateness of combining economic data from multinational clinical trials. *Stat Med* 2003;22:1955–76.
- [145] Wordsworth S, Ludbrook A. Comparing costing results in across country evaluations: the use of technology specific purchasing power parity estimates. *Health Econ* 2005;14:93–9.
- [146] Bagust A, Beale S. Survival analysis and extrapolation modeling of time-to-event clinical trial data for economic evaluation: an alternative approach. *Med Decis Making* 2014;34:343–51.
- [147] Anderson KM, Wilson PWF, Odell PM, et al. An updated coronary risk profile. *Circulation* 1991;83:356–62.
- [148] Latimer NR. Survival analysis for economic evaluations alongside clinical trials—extrapolation with patient-level data: inconsistencies, limitations, and practical guide. *Med Decis Making* 2013;33:743–54.
- [149] Schleinitz MD, Weiss JP, Owens DK. Clopidogrel versus aspirin for secondary prophylaxis of vascular events: a cost-effectiveness analysis. *Am J Med* 2004;116:797–806.
- [150] Mark DB, Hlatky MA, Califf RM, et al. Cost effectiveness of thrombolytic therapy with tissue plasminogen activator as compared with streptokinase for acute myocardial infarction. *N Engl J Med* 1995;332:1418–24.
- [151] Mark DB, Nelson CL, Tsiatis AA, et al. Cost-effectiveness of defibrillator therapy or amiodarone in chronic stable heart failure. *Circulation* 2006;114:135–42.
- [152] Ramsey SD, Berry K, Etzioni R, et al. Cost effectiveness of lung-volume-reduction surgery for patients with severe emphysema. *N Engl J Med* 2003;348:2092–102.
- [153] Caro JJ, Briggs AH, Siebert U, et al. Modeling good research practices—overview: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force-1. *Value Health* 2012;15:796–803.
- [154] Roberts M, Russell LB, Paltiel AD, et al. Conceptualizing a model: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force-2. *Value Health* 2012;15:804–11.
- [155] Siebert U, Alagoz O, Bayoumi AM, et al. State-transition modeling: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force-3. *Value Health* 2012;15:812–20.
- [156] Karnon J, Stahl JE, Brennan A, et al. Modeling using discrete event simulation: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force-4. *Value Health* 2012;15:821–7.
- [157] Pitman RJ, Fisman D, Zanic GS, et al. Dynamic transmission modeling: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force-5. *Value Health* 2012;15:828–34.
- [158] Briggs AH, Weinstein MC, Fenwick E, et al. Model parameter estimation and uncertainty analysis: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force-6. *Value Health* 2012;15:835–42.
- [159] Eddy DM, Hollingworth W, Caro JJ, et al. Model transparency and validation: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force-4. *Value Health* 2012;15:843–50.
- [160] Meltzer D. Accounting for future costs in medical cost-effectiveness analysis. *J Health Econ* 1997;16:33–64.
- [161] O'Neill P, Devlin NJ. An analysis of NICE's 'restricted' (or 'optimized') decisions. *Pharmacoeconomics* 2010;28:987–93.
- [162] Willke RJ, Zheng Z, Subedi P, et al. From concepts, theory and evidence of heterogeneity of treatment effects to analytic approaches: a primer. *BMC Med Res Methodol* 2012;12:185.
- [163] European Medicines Agency Committee for Medicinal Products of Human Use. Guideline on the investigation of subgroups in confirmatory clinical trials (DRAFT). 2014. Available from: [http://www.ema.europa.eu/docs/en\\_GB/document\\_library/Scientific\\_guideline/2014/02/WC500160523.pdf](http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2014/02/WC500160523.pdf). [Accessed December 19, 2014].
- [164] Scuffham PA, Yelland MJ, Nickles J, et al. Are n-of-1 trials economically viable options to improve access to selected high-cost medications? *Value Health* 2008;11:97–109.
- [165] Deb P, Trivedi PK. The structure of demand for health care: latent class vs. two-part models. *J Health Econ* 2002;21:601–25.
- [166] Frühwirth-Schatter S, Tüchler R, Otter T. Bayesian analysis of the heterogeneity model. *J Bus Econ Stat* 2004;22:2–15.
- [167] Drummond MF, Jefferson TO, for the BMJ Working Party on Guidelines for Authors and Peer-Reviewers of Economic Submissions to the *British Medical Journal*. Guidelines for authors and peer-reviewers of economic submissions to the *British Medical Journal*. *BMJ* 1996;313:275–383.
- [168] Siegel JE, Weinstein MC, Russell LB, Gold MR. Recommendations for reporting cost-effectiveness analyses. Panel on Cost-Effectiveness in Health and Medicine. *JAMA* 1996;276:1339–41.
- [169] Husereau D, Drummond M, Petrou S, et al. Consolidated Health Economic Evaluation Reporting Standards (CHEERS) statement. *Value Health* 2013;16:e1–5.
- [170] Husereau D, Drummond M, Petrou S, et al. Consolidated Health Economic Evaluation Reporting Standards (CHEERS)—explanation and elaboration: a report of the ISPOR Health Economic Evaluations Publication Guidelines Good Reporting Practices Task Force. *Value Health* 2013;16:231–50.
- [171] Schulz KF, Altman DG, Moher D. CONSORT 2010 statement: updated guidelines for reporting parallel group randomized trials. *Ann Intern Med* 2010;152:726–32.