BEST PRACTICES FOR
ECONOMIC ANALYSIS ALONGSIDE CLINICAL TRIALS:
AN ISPOR RCT-CEA TASK FORCE REPORT

Scott Ramsey¹, Richard Willke², Andrew Briggs³, Ruth Brown⁴, Martin Buxton⁵, Anita Chawla⁶, John Cook⁷, Henry Glick⁸, Bengt Liljas⁹, Diana Petitti¹⁰, Shelby Reed¹¹

DRAFT FOR REFERENCE GROUP REVIEW
November 12, 2004

¹Fred Hutchinson Cancer Research Center, Seattle, USA
²Pfizer, Inc., Bridgewater, USA
³University of Oxford, Oxford, UK
⁴MEDTAP International, London, UK
⁵Brunel University, Uxbridge, Middlesex, UK
⁶Genentech, San Francisco, USA
⁷Merck, Blue Bell, USA
⁸University of Pennsylvania, Philadelphia, USA
⁹AstraZeneca, Lund, Sweden
¹⁰Kaiser Permanente, Pasadena, USA
¹¹Duke Clinical Research Institute, Durham, USA

Address for correspondence: Scott Ramsey, Fred Hutchinson Cancer Research Center, 1100 Fairview Avenue North, PO Box 19024, Seattle, WA 98109 Phone (206) 667-7846, Fax (206) 667-7850, E-mail: sramsey@fhcrc.org
I. INTRODUCTION

Clinical trials evaluating medicines, medical devices and procedures now commonly assess 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. In recent years, research has also improved the methods for design, conduct, and analysis of economic data collected alongside clinical trials. Despite these advances, the literature reveals a great deal of variation in methodology and reporting of these studies. Improving the quality of these studies will enhance the credibility and usefulness of cost-effectiveness analyses to decision-makers worldwide.

To foster improvements in the conduct and reporting of trial-based economic analysis, the International Society for Pharmacoeconomics and Outcomes Research (ISPOR) chartered the ISPOR Task Force on Good Research Practices: Randomized Clinical Trials - Cost-Effectiveness Analysis (RCT-CEA). The Task Force co-chairs were selected by the ISPOR Board of Directors, and the co-chairs invited the other panel members to participate. The panel was first assembled in January 2004, met monthly, and agreed on an outline and preliminary content that was presented for comment by ISPOR membership at the May 2004 Annual Meeting. The draft report was then written and presented to the ISPOR membership at the October 2004 European Meeting. A volunteer Reference Group of ISPOR members provided valuable comments on the draft report which supported the completion of the final report in February 2005.

The purpose of this report is to state the consensus position of this Task Force. The goal for the panel was to develop a guidance document for the design, conduct, and
reporting of economic analyses alongside clinical trials. The intended audiences are researchers in academics, industry, and government who design and implement these studies, decision makers who evaluate clinical and economic evidence for formulary and insurance coverage policies, and students of the area. The panel recognizes that advances in methodology for joint clinical/economic analyses will continue, and that clinical/economic trials are heterogeneous in nature. Therefore, the report highlights areas of consensus, emerging methodologies with a diversity of professional opinions, and issues where further research and development are needed. In addition, clinical trials are artificial treatment environments and usually do not reflect the duration of impact of a test or treatment on persons over their lifetime. In this sense, modeling may be a necessary component of an economic evaluation of an intervention.

There are common issues in economic analysis that are fundamental to all studies of this nature that will not be addressed in this manuscript. These include study perspective, choice of discount rate for costs and outcomes, type of analysis (e.g., cost-utility, cost-benefit), types of costs that will be included (direct medical, non-medical, etc), and marginal versus average costing. These issues apply to all economic analyses, not just economic studies within clinical trials, and are well described in the literature.

II. INITIAL TRIAL DESIGN ISSUES

The quality of economic information that is derived from trials depends on the attributes of the trial’s design. That economic analyses are often described as being conducted ‘alongside’ clinical trials is indicative of an important practical design issue. Economic analysis is rarely the primary purpose of an experimental study. Nevertheless,
it is important that the analyst contributes to the design of the study, to ensure that the structure of the trial will provide the data necessary for a high quality economic study.

**Appropriate trial design**

The distinction between pragmatic and exploratory trials and the corresponding distinction between effectiveness and efficacy is well understood. 1,2 It is generally acknowledged that pragmatic effectiveness trials are the best vehicle for economic studies; however, it is usually necessary to undertake economic evaluations earlier in the development cycle where the focus is on efficacy (including phase III or even phase II drug trials), in order to provide timely information for pricing and reimbursement.

Large simple trials 3 are efficient for addressing clinical questions because they capture the main effects of treatments that have small to moderate impacts on large potential populations. They will also be efficient for answering economic questions for diseases or treatments where the bulk of costs derive from primary outcomes that are measured in the trial and for which the quality of life impacts are persistent, and thus can be measured infrequently.

An ideal follow-up period for an economic study is independent of the occurrence of clinical events, whether study related or not. All patients should be followed for a common length of time or the full duration of the trial. Discontinuation of data collection due to a clinical event will fail to capture the important aspects of the disease under study: the adverse effects of the event on quality of life, resource use and cost.
Sample size and power

In an ideal world the economic appraisal would be factored into sample size calculations using standard methods\textsuperscript{4,5} based on asymptotic normality, or by simulation.\textsuperscript{6} However, it is common for the sample size of the trial to be based on the primary clinical outcomes alone. As a consequence, it is possible that the economic comparisons will be underpowered. Analysts should calculate the likely power of the study at the design stage in order to ascertain whether, given the proposed sample size, it makes sense to undertake an economic appraisal. In many cases, sample size restrictions will necessitate focus on estimation rather than hypothesis testing of the economic outcomes. In cases where researchers wish to set up formal hypotheses for the economic analyses, these should be stated \textit{a priori} including the thresholds (e.g., $50,000 or $100,000/QALY) and the power to detect when incremental analysis meets or exceeds those thresholds\textsuperscript{7}.

Study endpoints

The choice of primary endpoint in a clinical study may not correspond with the ideal endpoint for economic evaluation. For example, the use of composite clinical endpoints is common in clinical trials (e.g., fatal events and non fatal events combined) in order to provide greater statistical power. However, cost per composite clinical endpoint is often an unsatisfactory summary measure for an economic analysis, in part because the different outcomes are rarely of equal importance. It is recommended that clinical endpoints used in economic evaluations be presented in disaggregated form. We recommend weighting endpoints (e.g., by utilities) so they yield a measure of quality-adjusted life-years (QALYs) in the case of cost-effectiveness analysis, or a monetary
benefit measure in the case of cost-benefit analysis. Alternatively, quality of life values may be obtained within the trial at regular intervals and the QALYs estimated as one of the outcomes of the trial.

If possible, one should avoid using intermediate or surrogate endpoints (e.g., percent LDL reduction) as the measure of benefit; however, intermediate outcome measures are often employed when the costs of conducting a long term trial are prohibitive. When use of intermediate outcomes 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 analyst should argue for follow-up sufficient to include clinically meaningful disease endpoints.

**Appropriate follow-up**

Economic analyses ideally include lifetime costs and outcomes of treatment. However, 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 endpoints gathered in the short run and long term disease outcomes – the stronger that relationship, the more a reliance on intermediate endpoints can be justified.

**III. DATA ELEMENTS**

The design issues discussed above will impact decisions about which resource use and outcome measures to collect, how to collect them, and how to value them. To begin, we recommend developing a description of the production function for the intervention
and how the intervention may impact resource use in the short-term and long-term. In this process, study perspective affects the types of resource use -- both medical and non-medical -- that should be considered for inclusion in the study. For example, the societal perspective might include patients’ costs for transportation, time spent undergoing treatment, caregiver time, and non-medical goods and services attributable to the disease or treatment.

After resources have been identified, logistical and cost considerations often require prioritization as to which data elements will be collected. We recommend that analysts focus on ‘big ticket’ items as well as resources that are expected to differ between treatment arms. For the items chosen, the study should collect information on all resource use, not just that considered to be disease- or intervention-related. If necessary, the distinction between costs related and unrelated to the disease can be attempted at the analysis phase.

For each resource, prospectively determine the level of aggregation desired. As an example, hospitalizations could be considered in disaggregated units, such as nursing time, operating room time, and supplies, or in highly aggregated units such as numbers of hospitalizations or days in the hospital. The decision is typically driven by the characteristics of the intervention under study, resource use patterns expected, and availability of unit costs (also called unit prices or price weights). For practical reasons, the level of aggregation may be varied by whether the resource use is thought to be disease or intervention related.
In some trial settings, secondary data such as hospital bills or claims data are available. These data sources can provide an inexpensive, detailed accounting of some resources consumed by patients, and should be used when available.\textsuperscript{12}

**Valuing Resource Use**

Unit costs should be consistent with measured resource use, the study’s perspective, and its time horizon. For example, if the level of resource aggregation is hospital days that include ICU and non ICU days, the unit costs should reflect the costs of each type of service;\textsuperscript{12,13} if the study is conducted from the societal perspective, unit costs should reflect social opportunity cost. In selecting a costing approach, analysts should weigh issues of accuracy/bias, cost, feasibility, and generalizability.\textsuperscript{14} For more thorough discussion on costing issues, we refer you to Drummond et al.\textsuperscript{15} and Luce et al.\textsuperscript{8}

Unit cost estimates are rarely derived via direct observation of patients in trials. Most often they are derived from substudies that are divorced from the trial itself. Sometimes, unit costs will be estimated from trial centers, but more commonly they are derived from national data.\textsuperscript{16-19} If a reliable method for cost imputation exists (e.g., diagnostic group weights), one can combine the two methods by collecting a limited set of unit costs in a number of countries and imputing the remainder of costs.\textsuperscript{20}

Because relative costs can affect resource use, in general one should use unit cost estimates that are specific to the precise intervention under study and specific events of interest in the trial, but which are generalizable to the population that the study is intended to inform.
When unit costs from more than one country are used in an analysis (e.g., the pooled analysis, when resource use in a country is multiplied by unit costs from that country), results have to be converted to a common currency if they are to be compared appropriately. Purchasing power parity adjustments are recommended for such conversion.\textsuperscript{21,22}

**Selecting and Tracking Measures of Outcomes**

Because cost-utility analyses are widely accepted, we recommend that analysts collect preference weights as part of clinical trials. The most common method of assessing preferences is the use of a preference-weighted health state classification system such as the EuroQol-5D\textsuperscript{23,24}, one of the 3 versions of the Health Utilities Index (HUI)\textsuperscript{25-28}, or the Quality of Well-Being Scale (QWB)\textsuperscript{29}. Analysts may also consider the inclusion of a rating scale to measure patient-based preferences.\textsuperscript{30} Frequency and timing of these assessments should capture changes in patients’ quality of life that may be affected by the treatment, but will be influenced by the disease severity of the study population, the study duration, the timing of trial visits, and patient burden.\textsuperscript{31}

Other options for collecting preference data include direct-elicitation methods such as standard gamble or time-tradeoff exercises. These methods have certain theoretical advantages; however, their use in clinical trials is often difficult. Trained interviewers or computerized applications are routinely used to conduct such exercises.\textsuperscript{15,32} Also, many respondents have difficulty understanding and completing the exercises.\textsuperscript{33-35} Finally, there is some evidence that these measures are generally unresponsive to changes in health status.\textsuperscript{36-39} At present, the balance between the
feasibility and desirability of using direct elicitation methods in clinical trials remains an issue to be decided on a case-by-case basis.

IV. DATABASE DESIGN AND MANAGEMENT

Ideally, collection and management of the economic data should be fully integrated into the clinical data. As such, there should be no distinction between the clinical and economic data elements. As with any prospective study, there should be a plan for ongoing data quality monitoring to address missing and poor quality data issues immediately. Queries should be managed on an ongoing basis rather than at the end of the trial to maximize data completeness and quality, and the timeliness of final analysis.

Informed consent for clinical studies does not routinely include provisions for collection economic data, particularly from third party databases. Therefore, explicit language should be included in trial consent documents. The consent forms should allow for capture of pre- and post-trial economic data if such data are necessary for the economic analysis.

Collection of economic data may reveal adverse events, such as hospitalization, not otherwise found in the clinical data. Data handling procedures are necessary to maintain consistency between economic and safety databases.

In reality, the clinical analysts are usually separate from the economic analysts. The clinical data elements and data formatting procedures needed for the economic analysis should be pre-specified such that transfer of all necessary data for the economics study is timely and complete.
V. ANALYSIS

Guiding Principles

The analysis of economic measures should be guided by a data analysis plan. A pre-specified plan is particularly important if formal tests of hypotheses are planned. Any tests of hypotheses that are not specified within the plan should be reported as exploratory. In writing the plan, specify whether regression or other multivariable analyses will be used to improve precision and to adjust for treatment group imbalances. The plan should also identify any selected subgroups and state whether a non-intention to treat analysis will be conducted.

While the specific analytic methods used in the analysis of resource use, cost and cost-effectiveness are likely to differ, there are several analysis features that should be common to all economic analyses alongside clinical trials:

- The *intention to treat* population should be used for the primary analysis
- 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 modeling or projecting beyond the time horizon of the trial;
- An *assessment of uncertainty* is necessary for each measure (standard errors or confidence intervals for point estimates; p-values for hypothesis tests);
- A (common) *discount rate* should be applied to future costs and, when used in a cost-effectiveness analysis, to future outcomes; and
• 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 whether the differences occurred by chance.

Once resources have been identified and valued, differences between groups must be summarized. Arithmetic mean cost differences are generally considered the most appropriate measure.\(^\text{40}\) However, cost data often do not conform to the assumptions for standard statistical tests for comparing differences in arithmetic means.\(^\text{41-43}\) They are usually right skewed and truncated at zero due to small numbers of high-resource-use patients, patients who incur no costs, and the impossibility of incurring costs less than zero. In most cases, the nonparametric bootstrap is an appropriate method to compare means and calculate confidence intervals.\(^\text{44,45}\) Other common nonparametric tests (e.g. Wilcoxon) compare medians and not means, and thus are not appropriate.\(^\text{46-48}\) Transformations to normalize the distribution (e.g. natural log transformation) are not straightforward and the test is not often on the proper scale (e.g., for the natural log, the test is of the geometric mean).\(^\text{49-53}\)

The same distributional issues that affect univariate tests of costs also affect use of costs as a dependent variable in a multivariate regression analyses. The underlying distribution of costs should be carefully assessed to determine the most appropriate
approach to conduct statistical inference on the costs between treatment groups.\textsuperscript{54} The choice of the multivariate model requires careful consideration: ordinary least squares (OLS) and generalized linear models (GLM) perform differently in terms of bias and efficiency of estimation, depending upon the underlying data distribution.\textsuperscript{50} If differences in resource use or subsets of costs are to be estimated, similar considerations regarding the appropriateness of statistical tests 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 to be related to the disease or treatment under study. If such analyses are performed, it is important to identify services that were deemed ‘disease related’ versus those deemed ‘unrelated,’ and to display costs for each component in the treatment and control arms.

**Outcomes**

When one of the trial’s clinical endpoints is also used as the outcome for the cost-effectiveness analysis (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. use of relative treatment differences rather than absolute treatment differences); if other outcomes are used for the economic analysis, the linkage between the clinical and economic measures should be
clearly specified. Economic analyses may employ multivariable or other methods that are consistent with the cost analysis or otherwise appropriate for the data.55-59

Using non-clinical effectiveness endpoints, such as QALY’s, involves both construction and analysis. Health state utilities, either collected directly from trial patients or imputed based on observed health states, can be transformed into QALYs using standard area-under-the-curve methods15; 60; a recent consideration involves adjusting for changes in health61. Simple analysis of means is the usual starting point; refinements may include adjusting for ceiling effects62 and modeling of longitudinal effects.63,64

**Missing and Censored Data**

Missing data are inevitable in economic analyses conducted alongside trials. Such data can include item level missingness and missingness due to censoring. In analyzing datasets 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); have a relationship to observed variables (missing at random), or be related to unobserved factors (not missing at random).65 Eliminating cases with missing data is not recommended as it may introduce bias or severely reduce power to test hypotheses. In general, if missing data comprise less than 10% of all data for the study and cannot be predicted, it can safely be ignored.

Imputation refers to replacing missing fields with estimates. If one chooses to impute missing data, most experts recommend multiple imputation approaches, as they
reflect the uncertainty that is inherent when replacing missing data.\textsuperscript{66-68} Most commonly used statistical software packages include programs for imputation of missing data. A review of these programs can be found at www.multiple-imputation.com.\textsuperscript{69}

Censoring can be addressed with a number of approaches. Most assume that censoring is either completely at random\textsuperscript{70} or at random.\textsuperscript{71-75}

**Summary Measures**

One or more summary measures should be used to characterize the relative value of the treatments in the clinical trial. Three general classes of summary measures are available that differ in how the incremental costs and outcomes are combined into a single metric:

- **Ratio measures** (e.g., incremental cost-effectiveness ratios) are obtained by dividing the incremental cost by the incremental health benefit
- **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\textsuperscript{76-78}
- **Probability measures** (e.g., acceptability curves) characterize the likelihood the new treatment will be deemed cost-effective based on the incremental costs and outcomes\textsuperscript{79,80}

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 of the most common measures of this variability is the confidence interval.

Policy inferences about adoption of a therapy should be based on one's level of confidence that its cost for a unit of outcome, for example, a quality-adjusted life year, is less than one's maximum willingness to pay. Thus, one should report ranges of ceiling ratios for which one 1) is confident that the therapy is good value for the cost, 2) is confident that the therapy is not good value, and 3) cannot be confident that the therapies differ from one another. Policy makers can then draw inferences by identifying their maximum willingness to pay and determining into which of the ranges it falls.

These ranges of ceiling ratios where one can and cannot be confident about a therapy's value can be calculated by use of confidence intervals for cost-effectiveness ratios, confidence intervals for net monetary benefit, or the acceptability curve. One advantage of the confidence interval for the cost-effectiveness ratio is that its limits
directly define the boundaries between these ranges. One advantage of the acceptability curve is that it defines the boundaries between these ranges for varying levels of confidence that range from 0 to 100%.

Parameter Uncertainty: Uncertainty related to parameter estimates such as unit costs and the discount rate should be assessed by use of 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% or a 5% rate. Analysts should evaluate all parameters which, when varied, have the potential to influence policy decisions. Measures of stochastic 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 employed to address missing or censored data (e.g., use of an imputed mean) may artificially reduce estimates of stochastic uncertainty. One should make efforts to address this shrinkage when reporting stochastic uncertainty, for example, by bootstrapping the entire imputation and estimation process.

Identifying and addressing threats to external validity/generalizability

Due to the “artificiality” of most clinical trials, they have high internal, but may have low external validity. The threats to external validity come from:
• 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)
• unrepresentative recruiting centers (e.g., large, urban, academic hospitals),
• inclusion of study sites from countries with varying access and availability of healthcare services (e.g., rehabilitation, home care, or emergency services),
• restrictive inclusion and exclusion criteria (patient population, disease severity, comorbidities), and
• artificially enhanced compliance.

The external validity can best be increased by making the trial more naturalistic during the design phase of the trial. 15,83

Additional threats arise with international trials, as treatment pathways, patient and health care provider behavior, supply and financing of health care, and unit costs (prices) can differ tremendously between countries.84-91 Pooled results may not be representative of any one country, but the sample size is usually not large enough to analyze countries separately.

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 within countries of similar levels of economic development and may not adjust for important country-specific differences.92,93 Rather, inter-country differences in population characteristics and
treatment patterns are more likely to influence summary measures between countries. Recommended approaches to address this issue include: \(^{92,94-96}\)

- hypothesis tests of homogeneity of results across countries (and adjusting the resource use in other countries to better match those seen in country X),
- multivariate cost or outcome regressions to adjust for country effects (eg. include country dummies or adjusted GNP per capita as covariates), and
- multi-level random effects model with shrinkage estimators.

**Modeling 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. Modeling is used to estimate costs and outcomes that one projects would 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. Parametric survival models estimated on trial data are recommended for such projections. In cases where 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. The reader is referred to the consensus position of the ISPOR Task Force on Good Research Practices—Modeling Studies for discussion of modeling issues. \(^{97}\)
Cost-effectiveness ratios should be calculated at various time horizons (e.g., 2, 5, 10 yrs), 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 not directly related to treatment should be taken into account as well as possible. As always, assumptions used must be described and justified, and the uncertainty associated with projections must be taken into account.

**Subgroup analysis**

The dangers of spurious sub-group effects are well known. For example, the probability of finding a difference due solely to random variation increases with the number of differences examined unless the alpha-level is scrupulously adjusted. Yet economics requires a marginal approach, so proper subgroup analysis can be vital to decision-makers. The focus should be on testing treatment interactions on the absolute scale, with a justification for choice of scale used. In cases where pre-specified clinical interactions are significant, subgroup analyses may be justified. Subgroup analysis based on pre-specified economic interactions should also be reported.

**VI. REPORTING THE METHODS AND RESULTS**

We anticipate that the results of an economic analysis will have a variety of audiences. Correspondingly, detailed and comprehensive information on the methods and results should be readily available to any interested reader. Journal word limits often necessitate parsimony in reporting; therefore, we recommend that detailed technical reports be made available on the world wide web.
A number of organizations have developed minimum reporting standards for economic analyses (e.g., study perspective, discount rate, marginal versus average outcomes and analyses). These should be adhered to in all economic studies. Here, we highlight issues particular to economic studies conducted alongside clinical trials.

The report should include these elements:

**Trial Related Issues**

- General description of the clinical trial, including patient demographics, trial setting (e.g., country, tertiary care hospital, etc), inclusion and exclusion criteria and protocol-driven procedures that influence external validity, intervention and control arms, and time horizon for the intervention and follow-up.

- Key clinical findings

**Data for the Economic Study**

- Clear delineation between patient level data collected as part of the trial versus data not collected as part of the trial:
  - Trial: health related quality of life survey instruments, data sources, collection schedule (including the follow-up period), etc.
  - Nontrial: unit costs, published utility weights, etc.

- Amount of missing and censored data

**Methods of Analysis**

- Construction of costs and outcomes
• In case the main clinical endpoint is used as the denominator in the incremental cost-effectiveness ratio, and different methods of how to analyze this endpoint have been used in the clinical and the economic analysis, any difference in the point estimate should be explained

• Methods for addressing missing and censored data

• Statistical methods used to compare resource use, costs, and outcomes

• Methods and assumptions used to project costs and outcomes beyond the trial period

Results

• Resource use, costs, outcome measures, including point estimates and measures of uncertainty

• Results within the time horizon of the trial

• Results with projections beyond the trial (if conducted)

• Graphical displays of results not easily reported in tabular form (e.g., cost-effectiveness acceptability curves, joint density of incremental costs and outcomes)

When there are economic analyses alongside several clinical trials for a given intervention, attempts may be made to estimate a summary cost-effectiveness ratio across trials (although the methods for this are not perfectly straightforward, i.e., clearly not a simple average of the individual incremental cost-effectiveness ratios). Data from economic analyses done in the context of trials may also be used in independent cost-
effectiveness models based on decision-analysis. To facilitate synthesis, report means and standard errors for the incremental costs and outcomes and their correlation.

VII. CONCLUSIONS

When feasible, conducting high quality economic analyses alongside clinical studies is desirable because they provide timely information with high internal validity. This ISPOR RCT-CEA Task Force Report is intended to provide guidance to improve their quality and consistency. The Task Force recognizes that there are areas where future methodological research could further improve the quality and usefulness of these studies. Examples here include (among many): new approaches for pooling and analyzing data from multinational trials; methods for projecting trial findings; appropriate methods for a priori selection of items of resource use to be included in trial protocols (e.g., whether to include outpatient services, non-study drugs), and; selecting levels of aggregation of resources necessary for discriminating between intervention and control (e.g., counts of hospitalizations versus length of stay, etc.). As these methods are identified and validated, they will be included in future versions of this guidance document.
Table 1. Core Recommendations for Conducting Economic Analyses Alongside Clinical Trials

<table>
<thead>
<tr>
<th>TRIAL DESIGN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trial design should reflect effectiveness rather than efficacy when possible.</td>
</tr>
<tr>
<td>Full follow-up of all patients is encouraged.</td>
</tr>
<tr>
<td>Describe power and ability to test hypotheses, given the trial sample size.</td>
</tr>
<tr>
<td>Clinical endpoints used in economic evaluations should be disaggregated.</td>
</tr>
<tr>
<td>Direct measures of outcome are preferred to use of intermediate endpoints.</td>
</tr>
<tr>
<td>Obtain information to derive health state utilities directly from the study population.</td>
</tr>
<tr>
<td>Collect all resources that may substantially influence overall costs; these include those related and unrelated to the intervention</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>DATA ELEMENTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Collection and management of the economic data should be fully integrated into the clinical data.</td>
</tr>
<tr>
<td>Consent forms should include wording permitting the collection of economic data, particularly when it will be gathered from third party databases and may include pre- and/or post-trial records.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>ANALYSIS</th>
</tr>
</thead>
<tbody>
<tr>
<td>The analysis of economic measures should be guided by a data analysis plan that is drafted prior to the onset of the study.</td>
</tr>
<tr>
<td>All economic analyses should include the following: an intention to treat analysis, common time horizon(s) for accumulating costs and outcomes; a within trial</td>
</tr>
<tr>
<td>assessment of costs and outcomes; an assessment of uncertainty; a common discount rate applied to future costs and outcomes; an accounting for missing and/or censored data.</td>
</tr>
<tr>
<td>---</td>
</tr>
<tr>
<td>Incremental costs and outcomes should be measured as differences in arithmetic means, with statistical testing accounting for issues specific to these data (e.g., skewness, mass at zero, censoring, construction of QALYs).</td>
</tr>
<tr>
<td>Imputation is desirable if there is a substantial amount of missing data. Censoring, if present, should also be addressed.</td>
</tr>
<tr>
<td>One or more summary measures should be used to characterize the relative value of the intervention. Examples include ratio measures, difference measures, and probability measures (e.g., cost-effectiveness acceptability curves).</td>
</tr>
<tr>
<td>Uncertainty should be characterized. Account for uncertainty that stems from sampling, fixed parameters such as unit costs and the discount rate, and methods to address missing data.</td>
</tr>
<tr>
<td>Threats to external validity—including protocol-driven resource use, unrepresentative recruiting centers, restrictive inclusion and exclusion criteria, and artificially enhanced compliance—are best addressed at the design phase.</td>
</tr>
<tr>
<td>Multinational trials require special consideration to address inter-country differences in population characteristics and treatment patterns.</td>
</tr>
<tr>
<td>When models are used to estimate costs and outcomes beyond the time horizon of the trial, good modeling practices should be followed. Models should reflect the expected duration of the intervention on costs and outcomes.</td>
</tr>
<tr>
<td>Subgroup analyses based on pre-specified clinical and economic interactions,</td>
</tr>
</tbody>
</table>
when found to be significant *ex post*, are appropriate. *Ad hoc* subgroup analysis is discouraged.

**REPORTING THE RESULTS**

Minimum reporting standards for economic analyses should be adhered to for those conducted alongside clinical trials.

The economic report should include a general description of the clinical trial and key clinical findings.

Reporting should distinguish economic data collected as part of the trial versus data not collected as part of the trial.

The amount of missing data should be reported. If imputation methods are used, the method should be described.

Methods used to construct and compare costs and outcomes, and to project costs and outcomes beyond the trial period should be described.

The results section should include summaries of resource use, costs, and outcome measures, including point estimates and measures of uncertainty. Results should be reported for the time horizon of the trial, and for projections beyond the trial (if conducted).

Graphical displays are recommended for results not easily reported in tabular form (e.g., cost-effectiveness acceptability curves, joint density of incremental costs and outcomes).
REFERENCES


U.S Diagnosis Related Group Weights. 
http://www.cms.hhs.gov/providers/hipps/ippspufs.asp

U.S. National Physician Fee Schedule.
http://cms.hhs.gov/providers/pufdownload/rvudown.asp

NHS Reference Costs 2003 and National Tariff 2004 (‘Payment by Results Core Tools 2004’).

Australian Hospital Information, Performance Information Program. Round 6 National Hospital Cost Data Collection Cost Weights.


White IR, Thompson SG. Choice of test for comparing two groups, with particular application to skewed outcomes. Statistics in Medicine 2003;22:1205-1215


O'Hagan A, Stevens JW. On estimators of medical costs with censored data. J Health Econ. 2004 May;23(3):615-25;


96 Wordsworth S, Ludbrook A. Comparing costing results in across country evaluations:
the use of technology specific purchasing power parity estimates. Health Economics
2004 (in press).

BR. Principles of good practice for decision-analytic modeling in health-care
evaluation: Report of the ISPOR Task Force on Good Research Practices – Modeling

98 Meltzer D. Accounting for future costs in medical cost-effectiveness analysis. Journal
of Health Economics 1997;16:33-64.

99 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

100 Siegel JE, Weinstein MC, Russell LB, Gold MR. Recommendations for reporting