Designing Economic Evaluations in Randomized Trials

ISPOR Education Teleconference
January 23, 2008

Henry Glick, PhD
University of Pennsylvania

Outline

• Steps in economic evaluation
• The gold standard and its tensions
• Strategic issues

Good Value for the Cost

• Economic data collected as primary or secondary endpoints in randomized trials are commonly used in the evaluation of the value for the cost of medical therapies
  – Short-term economic impacts directly observed; longer term impacts potentially projected by use of decision analysis
  – Reported results: point estimates and confidence intervals for estimates of incremental costs, outcomes, and the comparison of costs and effects
  – Impact of sensitivity analysis on the comparison of costs and effects judged by its impact on both the point estimates and the confidence intervals of the ratios

Example

<table>
<thead>
<tr>
<th>Analysis</th>
<th>Point Estimate</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Incremental Cost</td>
<td>-713</td>
<td>-2123 to 783</td>
</tr>
<tr>
<td>Incremental QALYs</td>
<td>0.13</td>
<td>0.07 to 0.18</td>
</tr>
<tr>
<td>Cost-Effectiveness Ratios</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Principal Analysis</td>
<td>Dominates</td>
<td>Dom to 6650</td>
</tr>
<tr>
<td>Survival Benefit</td>
<td>Dominates</td>
<td>Dom to 9050</td>
</tr>
<tr>
<td>-33%</td>
<td>Dominates</td>
<td>Dom to 5600</td>
</tr>
<tr>
<td>33%</td>
<td>Dominates</td>
<td></td>
</tr>
<tr>
<td>Hospitalization Cost</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-50%</td>
<td>Dominates</td>
<td>Dom to 5300</td>
</tr>
<tr>
<td>+50%</td>
<td>Dominates</td>
<td>Dom to 8400</td>
</tr>
<tr>
<td>Drug Cost</td>
<td>Dominates</td>
<td>Dom to 4850</td>
</tr>
<tr>
<td>-50%</td>
<td>Dominates</td>
<td>Dom to 8750</td>
</tr>
</tbody>
</table>

Steps in Economic Evaluation

Step 1: Quantify the costs of care
Step 2: Quantify outcomes
Step 3: Assess whether and by how much average costs and outcomes differ among the treatment groups
Step 4: Compare magnitude of difference in costs and outcomes and evaluate “value for costs” (e.g. by reporting a cost effectiveness ratio or the probability that the ratio is acceptable)
  – A hypothesis that might be tested in such a study is that the ratio of the cost per quality-adjusted life year saved is significantly less than $60,000
Step 5: Perform sensitivity analysis

Ideal Economic Evaluation Within a Trial

• An ideal economic evaluation within a trial is:
  – Conducted in naturalistic settings; compares the therapy with other commonly used therapies; and studies the therapy as it would be used in usual care
  – Performed with adequate power to assess the homogeneity of results in the wide range of clinical settings and among the wide range of clinical indications in which the therapy will be used
  – Designed with an adequate length of follow-up to assess the full impact of the therapy
  – Conducted within a time frame that allows the resulting information to inform important decisions in the adoption and dissemination of the therapy
Ideal Economic Evaluation Within a Trial (II)

• Measure all costs of all participants prior to randomization and for the duration of follow-up
  – Costs after randomization—cost outcome
  – Costs prior to randomization—potential predictor
• Independent of the reasons for the costs
• Most feasible when:
  – Easy to identify when services are provided
  – Service/cost data already being collected
  – Ready access to data

Difficulties Achieving an Ideal Evaluation

• Potential difficulties in meeting these goals within trials
  – Settings often controlled; comparator isn’t always the most commonly used therapy; investigators haven’t always learned fully how to use the new therapy under study
  – In some cases, sample size required to answer economic questions is greater than sample size required for clinical questions
  – In some cases, ideal length of follow-up required to answer economic questions is longer than follow-up necessary to answer clinical questions

Trade-off

• These trials may be the only source of information needed for important early decisions about the adoption and diffusion of the therapy
• TRADE-OFF: Ideal vs best feasible

Issue #1: What Medical Service Use Should One Collect?

• Real/perceived problems
  1. Don’t have sufficient resources to track all medical service use
  2. (In some cases), Don’t expect to affect all medical service use, just that related to the disease in question
– Implication: given sample size in trial, collection of all medical services, independent of the reason for these services, may swamp the “signal” with “noise”
→ Why not limit data to disease-related services?

What Medical Service Use? (2)

• Measure services that make up a large portion of the difference in treatment between patients randomized to the different therapies under study
  – Provides an estimate of the cost impact of the therapy
• Measure services that make up a large portion of the total bill
  – Minimizing unmeasured services reduces the likelihood that differences among them will lead to biased estimates
  – Provides a measure of overall variability

Measure as Much as Possible

• The best approach is to measure as many services as possible
• General Strategy: Identify a set of medical services one will collect, and assess them any time they are used, independent of the reason for their use
• There are no a priori guidelines about how much data are enough, nor are there data on the incremental value of specific items in the economic case report form
• Decisions about the services to measure should take into account the expense of collecting particular data items
Document Likely Service Use During Trial Design

- Decisions improved if one documents the types of services used by patients who are similar to ones who will be enrolled in the trial
  - Review medical charts or administrative data sets
  - Survey patients and experts about the kinds of care received
  - Have patients keep logs of their health care resource use
- Guard against possibility that new therapy will induce medical service use that differs from current medical service use

Limit Data to Disease-Related Services?

- Little if any evidence exists about the accuracy, reliability, or validity of such judgments
- Investigators routinely attributes AEs to the intervention, even when participants received vehicle/placebo
- Much of medical practice is multifactorial: modifying disease in one body system may affect disease in another body system
  - In the Studies of Left Ventricular Dysfunction, hospitalizations "for heart failure" were reduced by 30% (p<0.0001), and hospitalizations for noncardiovascular reasons were reduced 14% (p = 0.006)

Other Types of Costs?

- Other types of costs that sometimes are documented within economic evaluations include:
  - Time costs: Lost due to illness or to treatment
  - Intangible costs
- Types of costs that should be included in an analysis depend on what is affected by illness and its treatment and what is of interest to decision makers
  - e.g., the National Institute for Clinical Excellence (U.K.) and the Australian Pharmaceutical Benefits Scheme has indicated they are not interested in time costs

Specific Recommendations, Which Services?

- Identify common patterns of medical service use in centers/countries that will participate in the trials
  - Speak with experts in multiple centers
  - Focus groups, etc.
- Design case report forms to collect important, common medical service use
- Collect the services independent of the reason for their use
- Pilot test the forms
- Consider collecting costs other than medical service use

Issue #2. How Naturalistic Should The Study Design Be?

- The primary purpose of cost-effectiveness analysis is to inform real-world decision-makers about how to respond to real-world health care needs
- Thus, the more naturalistic the trial, in terms of participants, analysis based on the intention to treat, and limitation of loss to follow-up, the more likely the data developed within the trial will speak directly to the decision question

Naturalism: Intention-To-Treat Analysis

- Economic questions relate to treatment decisions (e.g., whether to prescribe a therapy), not whether the patient received the drug prescribed nor whether, once they started the prescribed drug, they were switched to other drugs
  - Implication: costs and effects associated with these later decisions should be attributed to the initial treatment decision
- Thus, trial-based cost-effectiveness analyses should adopt an intention-to-treat design
Naturalism: Loss to Follow-up

- Trials should be designed in such a way that they minimize the occurrence of missing data
  - For example, study designs should include plans to aggressively pursue subjects and data throughout the trial
  - One recent long-term study of treatment for bipolar disorder was designed from the outset to respond to missed interviews by:
    1. intensive outreach to reschedule the assessment, followed by
    2. telephone assessment, followed by
    3. interview of a proxy who had been identified and consented at the time of randomization

- Follow-up should continue until the end of the study period
  - Data collection should not be discontinued simply because a subject reaches a clinical or treatment stage such as failure to respond (as, for example, happens in some antibiotic, cancer chemotherapy, and psychiatric drug trials)
  - Given that failure often is associated with a change in the pattern of costs, discontinuation of these patients from the economic study is likely to bias the results of an economic evaluation that is conducted as part of the trial

Naturalism: Protocol-Induced Costs and Effects

- Clinical trial protocols often try to standardize the care of patients in the trial
  - They may require substantial number of investigations and diagnostic tests that would not be performed under normal clinical practice
- Trials also tend to prescribe aggressive documentation and treatment of potential adverse effects observed in the trial

- Requirements for “extra” diagnostic testing may bias the evaluation:
  - Use and cost of tests may be biased towards the null hypothesis of no difference
  - Diagnosis and treatment cost may be increased because of detection in the trial of outcomes that in usual care would not have been detected
    - e.g., in trials of prophylaxis for DVTs in elective hip replacement surgery, repeated testing for DVTs may identify a number of cases that never would have been detected or treated in usual practice

Naturalism: Protocol-Induced Costs and Effects (III)

- Adjustment for this extra detection -- potentially by use of decision analytic models -- may be difficult, because information usually is not available from the trial about whether active therapy avoided complications that would have been detected and treated in routine practice or whether it avoided those that never would have been detected

Specific Recommendations, Naturalism

- Use the intention to treat sample for the economic analysis
- Be aggressive in maintaining follow-up, including continuing to collect data on those who fail or switch therapy
- Use appropriate analytic to address missing data if and when they occur
- To the extent possible in a registration trial, minimize the effect of the protocol on patient care
Issue #3. What is the Appropriate Sample Size to Address Economic Questions?

- Economic sample size calculations are based on the number of study subjects needed to rule out that the therapy is unacceptable (equivalently, to ruling out that the net monetary benefits of the intervention are less than 0)

\[
 n = \left(\frac{\alpha + \beta}{2} + 2 \cdot \text{sd}^2 + 2 \cdot W \cdot (2 \cdot \text{sd}^2)^{1.5} \cdot (2 \cdot W^2)^{1.5} \right)^2
\]

where \( n \) equals n/group; \( \text{sd} \) = the standard deviation for costs (c) and effects (q); \( W \) equals the maximum willingness to pay one wishes to rule out; and \( \rho \) equals the correlation of the difference in cost and effect

Correlation Between Costs and Effects

- All else equal, the required sample size is less when the therapies have a Win/Lose (positive) correlation
  - As the effectiveness increases, the cost increases (e.g., stroke care)
- All else equal, the required sample size is greater when the therapies have a Win/Win (negative) correlation
  - As the effectiveness increases, the cost decreases (e.g., asthma care)
- Extreme values of correlation between costs and effects can have dramatic effects on the confidence interval for the cost effectiveness ratio/NMB and thus on the sample size required to demonstrate value for the cost

Where to Obtain the Necessary Data?

- For therapies already in use, expected differences in outcomes and standard deviations can be derived from review of medical charts and administrative data sets; patient logs of their health care resource use; or by asking patients and experts about the kinds of care received by those with the condition under study
  - Simple correlation between observed costs and effects may be an adequate proxy for the measure of correlation used for estimating sample size
- For novel therapies, expected differences may need to be generated by assumption
  - Data on the standard deviations and correlation for those who receive usual care/placebo may be obtained from feasibility studies or patient records

Maximum Willingness to Pay and Identification of an Appropriate Outcome Measure

- The sample size calculations described above assume that we have an idea about what we would like to pay to obtain a unit of outcome
- In many medical specialties, researchers use disease specific outcomes
- While one can calculate a cost-effectiveness ratio for any outcome one wants (e.g., cost/case detected or cost/additional abstinence day), to be convincing that a new, more costly and more effective therapy is good value, the outcome must be one for which we have recognized benchmarks of cost effectiveness
  - Argues against use of too disease-specific an outcome for economic assessment

Specific Recommendations, Sample Size

- If one therapy is both clinically equivalent to and less expensive than the alternative, a “disease-specific” outcome such as cases detected may be sufficient (depending on “how equivalent” it is)
- If the therapy is either less effective and less expensive or more effective and more expensive, one needs to know the value of the effectiveness so that it can be compared with the cost
  - Requires that we either know the value of the disease-specific outcome or that we use a more general health outcome such as QALYs
Issue #4. What Is The Likelihood That The Cost-effectiveness Ratio Observed In The Trial Describes Longer Term Therapy?

- When the trial observes cost-effectiveness for a time-limited period (e.g., 2 or 3 years), but the therapy will be taken for lifetime, one should consider the likelihood that the cost-effectiveness ratio observed in the trial will describe longer term therapy
  - Referred to as a time by treatment interaction

Likelihood of a Time By Treatment Interaction

- Time by treatment interaction less likely to be substantial when the intervention’s cost and outcome begin at approximately the same time and continue to be incurred together over time (e.g., drug therapy for heart failure)
- Interaction more likely to be substantial when:
  - Treatment cost and outcome incurred over time, but outcome delayed for a number of years (e.g., risk reduction from cholesterol-modifying therapy) OR increasing with time
  - Treatment cost incurred initially (e.g., surgical removal of tumor) and outcome (e.g., survival) accrued over time

Addressing a Time By Treatment Interaction

- Evaluate what was observed during the trial (within-trial analysis)
- Develop decision analytic models to make projections beyond the period of observation (projection)

Strengths and Weaknesses

- Within-trial analysis and longer term projections have opposing strengths/weaknesses:
  - We are more certain of what was observed during the trial, but follow-up may be too short to capture the most important impacts of the therapy
  - We are less certain about the projection beyond the trial, but this projection attempts to quantify what may be the most important impacts of the therapy

Within-Trial Analysis

- Even if one decides that the primary analysis will be a projection beyond the period of observation, one should still evaluate the costs and outcomes that were observed during the trial
- In such a within-trial evaluation, one should maintain the same time horizons for costs and outcomes observed in the trial (e.g., if follow-up for the trial was for one year, then costs and effects should be measured for one year)
- Not always easy to demonstrate cost-effectiveness in a within-trial analysis
  - e.g., no within-trial analysis of cholesterol-modifying therapy has demonstrated reasonable cost-effectiveness

Longer-Term Projection

- To investigate whether the cost-effectiveness ratio is homogeneous with respect to time, one should also project the results for longer periods
- For projection: Maintain a common time horizon for both costs and effects
  - some studies have used the cost difference observed within the trial; argued that the benefits of the therapy extend beyond the trial; and incorporated the benefits but not costs from beyond the trial
  - E.g., West of Scotland Coronary Prevention Study
  - If the therapy has downstream benefits that have not been adequately captured during the trial, it most likely has downstream costs that also have not been adequately captured
Time Horizon for Projection

• Given that the longer the projection, the less certain the results, one should make projections for different time horizons
  – Even if the longest time horizon in a lifetime projection is 30-40 years, one may observe that projected long-term cost-effectiveness reaches equilibrium after only 5 or 10 years of projection
• To add face validity to the trajectory of the projected cost-effectiveness ratios:
  – If there is sufficient follow-up during the trial, make estimates for differing lengths of follow-up during the trial (e.g., the first year, the first 2 years, etc.)

CER And CI Within the Trial and Projected

<table>
<thead>
<tr>
<th>Years of Follow-up</th>
<th>Point Estimate</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Within the trial</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Dominated</td>
<td>168,884 to Dominated</td>
</tr>
<tr>
<td>2</td>
<td>282,857</td>
<td>45,577 to Dominated</td>
</tr>
<tr>
<td>3</td>
<td>73,529</td>
<td>Dominates to Dominated</td>
</tr>
<tr>
<td>4</td>
<td>12,074</td>
<td>Dominates to Dominated</td>
</tr>
<tr>
<td>5</td>
<td>15,258</td>
<td>Dominates to 122,772</td>
</tr>
<tr>
<td>Longer term projection</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>12,246</td>
<td>Dominates to 42,263</td>
</tr>
<tr>
<td>15</td>
<td>8,578</td>
<td>Dominates to 26,721</td>
</tr>
<tr>
<td>20</td>
<td>7,320</td>
<td>681 to 21,841</td>
</tr>
</tbody>
</table>

Specific Recommendations, Treatment by Time Interaction

• Evaluate whether a strong treatment by time interaction is expected
• If yes, consider development of a decision model to assess the potential magnitude of the interaction
  – Substantial amounts of the data used for the decision model should be derived from the trial
  – Where necessary, augment data from the trial with epidemiologic data on long term outcomes, etc.

Summary

• Clinical trials may provide one of the best opportunities for developing information about a medical therapy’s value for the cost
• When appropriate types of data are collected and when they are analyzed appropriately, these evaluations can provide data about uncertainties related to the assessment of the value for the cost of new therapies that may be used by policy makers, drug manufacturers, health care providers and patients when the therapy is first introduced in the market

Further reading: