BACKGROUND

1. In response to increasing health care costs, government regulators, health care providers, members of managed care formulary committees, payers, and patients have begun to evaluate the value for the cost of individual medical therapies.

2. One of the growing trends in this evaluation has been the incorporation of economic evaluations within randomized controlled trials of medical therapies.

3. Most frequently these evaluations are incorporated into the drug development process.
   - Prior to approval in phases II and III: during which a drug’s safety and efficacy are evaluated prior to regulatory approval.
   - After approval in phase IV.

4. To a lesser extent, they are conducted within trials of other medical therapies (e.g., surgical procedures, behavioral interventions, etc.)

5. This approach to gathering economic data has been proposed by the U.S. Food and Drug Administration (FDA) to pharmaceutical companies wishing to make economic claims about their products.

ECONOMIC EVALUATION IN RANDOMIZED TRIALS

Henry Glick
University of Pennsylvania

ISPOR Shortcourse
May 17, 2003

www.uphs.upenn.edu/dgimhsr

DEMONSTRATING THAT A TECHNOLOGY IS “GOOD VALUE FOR THE COST” (I)

10-15 years ago, most likely would have supported this message by use of a decision analytic model such as a decision tree or a Markov model

- Little or none of the economic results would have been directly observed
  * e.g., the clinical evidence about the therapy would be that it reduced blood pressure
  * These data often would be combined with epidemiologic data relating blood pressure to death and disability to project the likely economic impact of therapy

- Reported results would have included point estimates of incremental costs, outcomes, and comparison of costs and effects from a “principal” or “base-case” analysis as well as the results of sensitivity analysis

“GOOD VALUE FOR THE COST” (II)

In the mid-90’s, for a cutting edge evaluation, the message would have been supported by use of an data from a randomized trial that included economic outcomes as primary or secondary endpoints of the trial

- Short-term economic impacts of the therapy would be directly observed; longer term impacts potentially would be projected by use of decision analysis

- Reported results would include point estimates and confidence intervals for estimates of incremental costs, outcomes, and the comparison of costs and effects

- The impact of sensitivity analysis on the comparison of costs and effects would be judged by its impact on both the point estimates and the confidence intervals of the ratios

“GOOD VALUE FOR THE COST” (III)

By the end of the 90’s, for a cutting edge evaluation, the message would have been supported by use of an evaluation from a randomized trial that included economic outcomes as primary or secondary endpoints of the trial

- Short-term economic impacts of the therapy would be directly observed; longer term impacts potentially would be projected by use of decision analysis

- Reported results would include point estimates and confidence intervals for estimates of incremental costs, outcomes, and the comparison of costs and effects

- The impact of sensitivity analysis on the comparison of costs and effects would be judged by its impact on both the point estimates and the confidence intervals of the ratios
EXAMPLE OF TYPICAL LATE '90'S RESULTS

<table>
<thead>
<tr>
<th>Analysis</th>
<th>Point Estimate</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Incremental Costs</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>
</tbody>
</table>

Cost-Effectiveness Ratios

Principal Analysis
- Dominates Dom to 6650

Survival benefit
-33% - Dominates Dom to 9050
+33%  - Dominates Dom to 5800

Hospitalization costs
-50% - Dominates Dom to 5300
+50%  - Dominates Dom to 8400

Drug costs
-50% - Dominates Dom to 4850
+50%  - Dominates Dom to 8750

Discount rate
- 0%   - Dominates Dom to 6350
    7%   - Dominates Dom to 7000

OUTLINE

- Steps in economic evaluation
- The gold standard and its tensions
- 4 strategic issues

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 differences in costs and outcomes and evaluate "value for cost" (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 analyses
IDEAL ECONOMIC EVALUATION WITHIN A TRIAL

- An ideal economic evaluation within a clinical 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 reason for the costs

- Most feasible when:
  - Easy to identify when services are provided
  - Service / cost data already being collected
  - Ready access to data

IDEAL ECONOMIC EVALUATION WITHIN A TRIAL (III)

- A number of design issues apply equally to economic evaluations that are incorporated within clinical trials as well as to other economic evaluations:
  - The type of analysis that will be conducted (e.g., cost-benefit, cost-effectiveness, or cost minimization analysis)
  - The types of costs that will be included (e.g., direct medical, direct nonmedical, productivity, and intangible)
  - The perspective from which the study will be conducted

- These issues have been well addressed in the literature

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

- However, these trials may be the only source of information needed for important early decisions about the adoption and diffusion of the therapy

- TRADEOFF: ideal vs best feasible study
4 STRATEGIC ISSUES

! What is an appropriate length of follow-up for economic outcomes?
! What proportion of the total costs should be collected in the trial?
! What is the appropriate sample size for economic questions?
! Will the relative magnitude of costs and effects of lifetime therapies observed within time-limited trials be descriptive of the relative magnitudes that would be observed had the trials continued for the patients’ lifetime?

STRATEGIC ISSUE #1. IDENTIFYING APPROPRIATE LENGTH OF FOLLOW-UP

! Economic assessments conducted as part of randomized trials are meant to allow decision makers to use the results of the trial to reach conclusions about the economic benefits of the therapy under investigation

! One design issue that may limit the interpretability of the economic data collected within the trial is the study time horizon

! Although clinical efficacy may be demonstrated when a difference in clinical endpoints is observed between study arms, from an economic perspective the appropriate time horizon for a trial would include all (or a substantial portion) of the time when there is resource use related to the illness under study

! The economic time horizon that would best inform decision makers about the value for the cost of a therapy thus need not be the same as the one adopted for answering the clinical question

STRATEGIC ISSUE #2. PROPORTION OF THE TOTAL RESOURCES

! Access to billing data may obviate need for tradeoffs in collection of direct medical costs

! If administrative data unavailable:

- 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 the services that go unmeasured reduces the likelihood that differences among them will lead to study artifacts
  * Provides a measure of overall variability of costs
**CRF: SITE OF CARE IN HOSPITAL**

<table>
<thead>
<tr>
<th>Admission to Hospital or Unit</th>
<th>Unit (*)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ICU</td>
</tr>
<tr>
<td></td>
<td>CCU</td>
</tr>
<tr>
<td></td>
<td>Step-down</td>
</tr>
<tr>
<td></td>
<td>General care</td>
</tr>
<tr>
<td></td>
<td>OTHER</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Pretreatment</th>
<th>Posttreatment</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Day -2</td>
<td>Day -1</td>
</tr>
<tr>
<td>Type of Bed</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(at noon)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intensive care unit</td>
<td></td>
<td></td>
</tr>
<tr>
<td>High care unit</td>
<td></td>
<td></td>
</tr>
<tr>
<td>General care unit</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other care unit</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Discharged</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**PROPORTION OF THE TOTAL RESOURCES (II)**

- The best approach is to measure as many services as possible, but 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 those procedures that will be recorded within the case report form and those that will not should take into account the expense of collecting particular data items.

- Identification of the types of medical services used by study subjects during the preplanning activities for the trial can provide information to guide these decisions, by documenting the types of services used by study subjects with the disease that is being studied.

  - Review of medical charts or administrative data sets
  - Have patients keep logs of their health care resource use
  - Ask patients and experts about the kinds of care received by those with the condition under study.
**CRF: IMAGING STUDIES IN THE HOSPITAL**

<table>
<thead>
<tr>
<th></th>
<th>Day -2</th>
<th>Day -1</th>
<th>Day 1</th>
<th>Day 2</th>
<th>Day 3</th>
<th>Day 4</th>
<th>Day 5</th>
<th>...</th>
</tr>
</thead>
<tbody>
<tr>
<td>Imaging studies</td>
<td>Enter number of images per day</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>X-rays (#)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CT Scans (#)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MRI scans (#)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ultrasound (#)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Specify (#)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**CRF: PHYSICIAN SERVICES IN THE HOSPITAL**

<table>
<thead>
<tr>
<th></th>
<th>Day -2</th>
<th>Day -1</th>
<th>Day 1</th>
<th>Day 2</th>
<th>Day 3</th>
<th>Day 4</th>
<th>Day 5</th>
<th>...</th>
</tr>
</thead>
<tbody>
<tr>
<td>Physician Services</td>
<td>Enter # of minutes or use (*) to indicate a visit when number of minutes are unknown</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cardiologist</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neurologist</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Infectious disease</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nephrologist</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Surgeon</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**PROPORTION OF THE TOTAL RESOURCES (III)**

! Collect services physicians indicate were related to disease / therapy?
  
  - Little if any evidence about the accuracy, reliability, or validity of such judgments
  
  - Much of medical practice is multifactorial
    
    * In a trial for heart failure, a patient with more severe heart failure may be hospitalized for a comorbid condition, whereas a patient with the same comorbidity, but with milder heart failure may not be hospitalized
  
  ! Frequently performed, low cost items?
    
    - 6,700 blood gas tests equaled 1.8% of procedure and diagnostic test costs
    
    - 420 angiocardio pneumographies equaled 4.3%
STRATEGIC ISSUE #3. SAMPLE SIZE FOR ECONOMIC QUESTIONS

- Recent changes in the methods for estimating sample size for economic assessments in trials reflect changes in the economic question being asked.

- Prior to the development of the literature that described confidence intervals for cost-effectiveness ratios, a common approach was to base sample size on the larger of the sample sizes needed for estimating pre-specified cost and effect differences.

- i.e., what sample size was required to identify a $1000 difference in costs, and what was required to identify a 10% reduction in mortality.

- Formula for continuous variable (assuming a normal distribution):

\[ n = \frac{2\sigma (1-R^2)(z + z')}{\Delta^2} \]

where \( n \) = sample size per group; \( \Delta^2 \) equals the difference in means between the therapies; \( \sigma^2 \) equals the variance; \( R^2 = \) variance explained in OLS (assuming OLS is used to predict the difference).

SAMPLE SIZE FOR ECONOMIC QUESTIONS (II)

- Once the literature on confidence intervals developed, however, it became clear that the goal of economic evaluations in trials was to determine the likelihood that the therapy represented good value for the cost.

- Current sample size methods base their calculations on the number of study subjects needed to rule out unacceptably high upper confidence limits for the cost-effectiveness ratio (equivalently, to rule out that the net monetary benefits of the intervention are less than 0).

INFORMATION REQUIRED TO ESTIMATE SAMPLE SIZE FOR ECONOMIC QUESTION

- The newer methods generally require more information than is needed for estimating sample sizes for clinical outcomes or for cost differences alone.

- Basic data for such calculations include the magnitude of the incremental costs and outcomes one expects to observe in the trial; the standard deviations for costs and outcomes in each of the treatment groups; and the correlation between costs and outcomes.

  - Win/lose (positive) correlation

  - Win/win or lose/lose (negative) correlation

CORRELATION BETWEEN COSTS AND EFFECTS

- Correlation between costs and effects can have dramatic effects on the confidence interval -- and thus the sample size -- required for estimation of the cost-effectiveness ratio.
WHERE TO OBTAIN THE NECESSARY DATA?

- If both therapies are already in use, expected differences in outcomes and standard deviations can be derived from feasibility studies or from records of patients like those who will be enrolled in the trial

- Potential sources
  - Medical charts of administrative data sets
  - Patient logs of their health care resource use
  - Asking patients and experts about the kinds of care received by those with the condition under study

- In addition, at least one study has suggested that the correlation between costs and effects observed in these data may be an adequate proxy for the measure of correlation used for estimating sample size

WHERE TO OBTAIN THE NECESSARY DATA? (II)

- For novel therapies, information about the magnitude of the incremental costs and outcomes may not be available

  - May need to be generated by assumption

  - Data on the standard deviations for those who receive usual care/placebo may be obtained from feasibility studies or patient records

    - One may assume that the standard deviation will apply equally to both treatment groups, or one may make alternative assumptions about the relative their magnitudes

    - The correlation also would be obtained from such data

CEILING RATIO 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 (e.g., $50,000 per quality-adjusted life year saved)

- In many medical specialties, researchers use disease specific outcomes

  - e.g., cases detected; in alcohol research, abstinence days; etc.

- 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
STRATEGIC ISSUE #4. UNCERTAINTY ABOUT PROJECTION OF RESULTS

When lifetime therapies are studied with time-limited trials, uncertainty exists about whether the relative magnitude of costs and effects observed within the trials will be descriptive of the relative magnitudes that would be observed had the trials continued for the patients’ lifetime.

! To address these uncertainties one should:

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

The two types of analyses have offsetting strengths/weaknesses:

- We are more certain of the data from the trial, but they may fail to capture the most important impacts of the therapy
- We are less certain about the projections beyond the trial, but they may attempt to quantify 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 also 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

Because the relative magnitude of incremental costs and outcomes observed during the trial may not be reflective of the relative magnitude that would have been observed had the trial been continued until all study subjects discontinued therapy or died, one should also project the results for longer periods.

For projection: Maintain a common time horizon for both costs and effects.

- A number of 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 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 during the trial.

Given that the longer the projection, the less certain the results, make projections for different time horizons.

<table>
<thead>
<tr>
<th>Years of Follow-up</th>
<th>Point Estimate</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<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>5.5</td>
<td>$15,258</td>
<td>Dominates to $122,772</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 $6,721</td>
</tr>
<tr>
<td>20</td>
<td>$7,320</td>
<td>$681 to $21,841</td>
</tr>
</tbody>
</table>
OTHER CONTENTIOUS STRATEGIC ISSUES

! Should one collect data on medical service use (which will then be multiplied by unit cost estimates) or should one collect data costs?

! Which unit cost estimates should be used for the study?

! In what delivery settings should data be collected, and how should utilization in nonstudy sites be identified?

! Among which patients should data be collected?

! Should follow-up be discontinued if participants fail therapy?

! Are the observed results transferable to the specific centers/countries that participated in the trial? To other centers/countries?

SUMMARY

! Many opportunities exist for incorporating economic assessments into randomized trials assessing medical therapies

! Clinical trials may provide the best opportunity for developing information about a medical therapy’s value for the cost early in its product life

! 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