“GOOD VALUE FOR THE COST”

- Cutting edge evaluation of the value for the cost is 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 are directly observed; longer term impacts potentially are be projected by use of decision analysis
  - Reported results 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

<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>
<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></td>
<td></td>
</tr>
<tr>
<td>-33%</td>
<td>Dominates</td>
<td>Dom to 9050</td>
</tr>
<tr>
<td>+33%</td>
<td>Dominates</td>
<td>Dom to 5800</td>
</tr>
<tr>
<td>Hospitalization costs</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 costs</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-50%</td>
<td>Dominates</td>
<td>Dom to 4850</td>
</tr>
<tr>
<td>+50%</td>
<td>Dominates</td>
<td>Dom to 8750</td>
</tr>
<tr>
<td>Discount rate</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0%</td>
<td>Dominates</td>
<td>Dom to 6350</td>
</tr>
<tr>
<td>7%</td>
<td>Dominates</td>
<td>Dom to 7000</td>
</tr>
</tbody>
</table>
OUTLINE

- Steps in economic evaluation
- The gold standard and its tensions
- 3 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: Evaluate stochastic uncertainty and perform sensitivity analysis

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

3 STRATEGIC ISSUES

- Could devote days to discussing issues related to design and analysis of economic assessments in trials
- Focus on 3 major sets of issues:
  1. What medical service use should one collect?
  2. What is the appropriate sample size to address economic questions?
  3. How naturalistic should the study design be?
    - Intention-to-treat analysis
    - Loss to follow-up
    - Protocol-induced Costs and Effects

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?
LIMITED DATA COLLECTION RESOURCES

- Access to billing data may obviate resource limitations associated with tracking all medical service use
- 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 biased estimates
    * Provides a measure of overall variability of costs

MEASURE AS MUCH AS POSSIBLE

- The best approach is to measure as many services as possible
  - 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 will be improved if, when designing a study, 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

- Must guard against possibility that new therapy will induce medical service use that differs from current medical service use

ACCOUNT FOR DATA COLLECTION EXPENSE

- Decisions about the services to measure should take into account the expense of collecting particular data items
  - e.g., frequently performed, low cost items?
    * 6,700 blood gas tests equaled 1.8% of procedure and diagnostic test costs
    * 420 angiocardiopneumographies equaled 4.3%
LIMIT DATA TO DISEASE-RELATED SERVICES?

- Little if any evidence exists about the accuracy, reliability, or validity of such judgments

- Easy for judgments to be flawed:
  - 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 (positively or negatively) affect disease in another body system
    * In the Studies of Left Ventricular Dysfunction, hospitalizations "for heart failure" were reduced by 30% (combined endpoint, death and HF hospitalization, p<0.0001); simultaneously, hospitalizations for noncardiovascular reasons were reduced 14% (p = 0.006)
  - If a patient has an automobile accident, how does the clinician determine whether or not it was due to a hypotensive event caused by therapy?

- Potential biases more of a problem in unblinded studies, but need not "balance out" in double-blinded studies

GENERAL RECOMMENDATIONS

- 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

- If data collection is limited to a single page in the CRF:
  - First impression: Collect big-ticket items, (e.g., hospitalization, long term care, etc); don't sweat smaller ticket items
    * Heart failure: hospitalization costs, number of outpatient visits
    * Hospitalized infections: ICU, stepdown, and routine care days; major procedures
    * Asthma: ER visits, Hospitalizations, comedications

BETTER APPROACH

- Prior to the study, invest in determining which services will likely make up a large portion of the difference in costs between the treatment groups
  - If the therapy is likely to affect the number of hospitalizations: Collect information that will provide a reliable estimate of the cost of these hospitalizations
  - If the therapy is likely to affect days in the hospital and location in the hospital, collect this information
  - If the therapy is principally likely to affect outpatient care, collect measures of outpatient care, etc.
ISSUE #2. WHAT IS THE APPROPRIATE SAMPLE SIZE TO ADDRESS ECONOMIC QUESTIONS?

- Prior to the development of the literature that described confidence intervals for cost-effectiveness ratios, a common approach for estimating sample size was to base it 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

- 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)

SAMPLE SIZE FORMULA

- Sample size for NMB uses the following formula:

\[ n = \frac{(\alpha + \beta)^2 (2 \text{ sd}_c^2) + (2 \text{ rc} \text{ sd}_c^2) - (2 \text{ rc} \rho (2 \text{ sd}_c^2)^{0.5} (2 \text{ sd}_q^2)^{0.5})}{\Delta \text{NMB}^2} \]

where \( n \) equals \( n/\text{group} \); \( \text{sd} \) = the standard deviation for costs (\( c \)) and effects (\( q \)); \( \text{rc} \) equals the ceiling ratio one wishes to be able to rule out; and \( \rho \) equals the correlation of the difference in cost and effect; and \( \Delta \text{NMB} \) equals \( (\text{rc} \Delta q) - \Delta c \)

http://www.uphs.upenn.edu/dgimhsr/stat%20samps.htm

- Basic data for such calculations include the magnitude of the incremental costs and outcomes (in the formula, they appear in the NMB term); the standard deviations for costs and outcomes; and the correlation between costs and outcomes

CORRELATION BETWEEN COSTS AND EFFECTS

- Win/Lose (positive) correlation: As the effectiveness (cost) increases, the cost (effectiveness) increases
- Win/Win (negative) correlation: As the effectiveness (cost) increases, the cost (effectiveness) decreases
- 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?

- 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 simple correlation between costs and effects observed in these data may be an adequate proxy for the measure of correlation used for estimating sample size

OBTAINING DATA FOR NOVEL THERAPIES

- 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 their relative magnitudes
  - The correlation also would be obtained from such data

RELATIVE SAMPLE SIZE FOR CLINICAL AND ECONOMIC QUESTIONS

- Generally accepted that economic results have less power than do clinical results
- Exceptions exist:

![Graph showing trade-offs between costs and QALYS saved](image-url)
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

ISSUE #3. HOW NATURALISTIC SHOULD THE STUDY DESIGN BE?

- Given that the primary purpose of cost-effectiveness analysis is to inform real-world decision-makers about how to respond to real-world health care needs, 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 benefits 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

- One should design studies 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

- Investigators should also ensure that follow-up continues until the end of the study period, and that data collection 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
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.

These requirements for 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.

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.

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?

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

What is the likelihood that the cost-effectiveness ratio observed in the trial describes longer term therapy?

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.