

Panel I: Methodological Issues in Pharmacoeconomic Evaluations—Clinical Studies

Co-chairs: Margaret Healey, PhD,¹ Patricia Deverka, MD, MS²

Panelists: Steven Fox, MD,³ Kathleen Gondek, PhD,⁴ Barbara Edelman Lewis, PhD,⁵ Eva Lydick, PhD,⁶ Gurkupal Singh, MD,⁷ Robert Temple, MD,⁸ Jeff Trotter, MBA⁹

¹The Institute for Research & Education, HealthSystems Minnesota, Minneapolis, MN, ²Merck Medco Managed Care, LLC, Montvale, NJ, ³Center for Outcomes & Effectiveness Research, Agency for Health Care Policy and Research, Rockville, MD, ⁴Boehringer Ingelheim Pharmaceuticals Inc., Ridgefield, CT, ⁵Focus Managed Research Inc., Worcester, MA, ⁶SmithKline Beecham, Collegeville, PA, ⁷Stanford University Medical Center, Palo Alto, CA, ⁸Food & Drug Administration, Rockville, MD, ⁹ClinTrials Ovation Inc., Highland Park, IL

The goal of this panel was to identify key contentious methodology issues in conducting healthcare pharmacoeconomic evaluations in the context of clinical studies. Its specific objectives were to:

- identify and prioritize the key issues associated with including pharmacoeconomic and outcomes research projects in clinical studies;
- identify a plan of action to resolve these issues;
- recommend next steps.

Background and Context

Consumers of pharmacoeconomic information include patients, clinicians, and, in particular, healthcare decision- and policy-makers in managed care and government. Evaluations of the costs and outcomes of various treatment options are in increasing demand to inform formulary and treatment guideline decision-making. In order to assess the value, and not just the price, of new interventions, consumers need data that are both valid and generalizable. Pharmacoeconomic data used to provide such evidence of the value of therapeutic interventions can be generated from a variety of study methodologies.

The strengths and limitations of various clinical study designs for economic evaluation, particularly randomized controlled trials (RCTs), have been described elsewhere [1–4]. Briefly, the use of RCTs to develop pharmacoeconomic data has several advantages: these data are based on a well-established and accepted methodology, are readily interpretable, have statistical rigor, and can be designed free of most types of bias. The process of randomization is particularly critical in decreasing the probability of selection bias. Especially when

studying small effects, other methodologies (e.g., observational studies) are less reliable.

The limitations of RCTs have also been recognized. These arise primarily from the costs associated with carrying out trials with multiple comparators, trials of long duration, and trials large enough to detect infrequent outcomes. In addition, the prospective nature of RCTs means that there is a substantial delay between posing a question and answering it. Furthermore, it is difficult to separate protocol-driven costs from actual costs of care. In many cases, the choice of clinical trial outcome measures may not be relevant to standard or usual care. Concern has also been expressed that when therapies with proven efficacy are translated to “usual care” settings, the results may fall short of those predicted by controlled studies. The outcomes achieved when new interventions are implemented in actual practice settings are often referred to as evidence of the effectiveness (as distinguished from the efficacy) of that treatment.

These limitations on the usefulness and the generalizability of the results from controlled clinical trials have led to the increased use of alternative designs, including observational studies, to address these problems. The advantages of using observational studies to collect pharmacoeconomic data include the ability to follow relatively large numbers of patients for extended periods of time. Whereas the outcomes in RCTs sometimes represent intermediate outcomes, prolonged follow-up in observational studies allows for capturing long-term outcomes, including mortality.

Observational designs are often used when studying the effectiveness of a treatment. Whereas choice of comparator is often precluded, comparators are generally those used most frequently in

practice settings, enhancing the generalizability of study results.

The most serious limitations of observational studies are methodological. Arguably the greatest weakness of the observational study is the inability to control for selection bias. In addition, strategies to control for confounding variables are limited.

Users of pharmacoeconomic data need information on both the efficacy and effectiveness of new treatments and how these compare for alternative treatments. Although few would argue that the RCT is the “gold standard” when evaluating the efficacy of an intervention, it may be possible to maximize the usefulness of other methods by controlling for as many confounding variables as possible. Our understanding of the relative value of interventions could be advanced by the adoption of study designs chosen to balance scientific exigencies with practical application.

Problem Statement

The study methods that are acceptable for demonstrating drug efficacy are largely agreed upon and are, in most cases, randomized controlled trials (RCTs). RCTs have been criticized, however, for not adequately addressing value in applied settings. In addition, limiting pharmacoeconomic data to those derived from RCTs would reduce the body of timely information. Alternative methods, such as observational studies, utilized by health economists and epidemiologists for assessing questions of value are often compromised by methodological inadequacies, particularly selection bias. Developing a better understanding of when different methods are appropriate will improve the usefulness and validity of pharmacoeconomic information.

Issues

This paper addresses four primary issues:

1. Under what circumstances should RCTs be the primary approach for assessing questions of value?
2. How can RCTs be modified to improve their usefulness to inform economic decision-making?
3. Under what circumstances can observational studies be used to assess questions of value?
4. How can observational studies be modified to improve their usefulness to inform economic decision-making?

Under What Circumstances Should RCTs Be the Primary Approach for Assessing Questions of Value?

Pharmacoeconomic data can be derived from randomized trials, but randomized controlled trials are not always the first choice for a variety of reasons, both practical and scientific. Randomized controlled trials should be considered the primary approach, however, under the following circumstances:

- The significance of the results justifies the expense. Not every question merits the level of assurance that a result is replicable and free from most types of bias that an RCT can provide. The “cost-effectiveness” of the study needs to be evaluated. For example, the Federal Trade Commission Policy Statement Regarding Advertising Substantiation [5] specifically mentions that the cost of developing the evidence for substantiation of a cost-effectiveness conclusion is one of the key issues in deciding the adequacy of the scientific basis for support. The FTC’s extensive experience in regulating the promotion of nonprescription drug claims provides an important perspective when considering whether to invest in a clinical trial to evaluate the economic impact of a particular drug.
- There is a reasonable likelihood of observing patient outcomes within a relatively short time, and/or intermediate endpoints (e.g., success in controlling blood pressure) are acceptable. When RCTs are used to study long-term outcomes, they can be very costly. Care must be exercised, however, in extrapolating to long-term outcomes that have not been causally linked to intermediate endpoints. When intermediate outcome variables are used, the data supporting the level of certainty in the link between the intermediate variable and final outcomes need to be well established and disclosed.
- There is consensus on the appropriate comparator(s). If there is agreement (within a defined practice setting) that one intervention is the standard treatment approach, a single large study (RCT) with a limited number of treatment arms presents a practical approach.
- Effect sizes are predicted to be small, but still important from a clinical or public health perspective. Under these circumstances, observational studies are of limited usefulness because the magnitude of the observed effect of the drug may be similar to the amount of uncon-

trolled confounding. A randomized controlled trial will provide the strongest evidence that the observed differences between treatments are not due to confounding and that the association between treatment and outcome is one of cause and effect.

- Randomization is acceptable to (physician) investigators and to patients.

How Can RCTs Be Modified to Improve Their Usefulness to Inform Economic Decision-Making?

Most of the concern expressed about RCTs, particularly from decision-makers, is not attributable to methodological flaws, but to questions about the generalizability of the results to practice settings. To mitigate these concerns:

- Conduct RCTs in settings where decisions are made, for example, in managed care organizations.
- Adopt more naturalistic study designs.
- Liberalize inclusion criteria and minimize exclusion criteria:
 - Use active comparators where reasonable. Placebos are generally perceived by decision-makers as not pertinent. When expected measures of effectiveness are less robust (e.g., in studies of depression), a placebo comparator may be needed for credible assessments of effectiveness.
 - Use relevant comparators. Select comparators based on community accepted standard treatments whenever feasible. If the optimal dose of the comparator agent is unknown, it may be necessary to test multiple doses.
 - Mimic usual care as closely as possible. The effectiveness of new interventions is critical to providers, as well as to decision-makers. Protocol required assessments and follow-ups should be structured to align as closely as possible with acceptable community standards. Whenever possible, the frequency of those assessments should not be dictated by the protocol.
 - Reduce the burden of adverse event reporting, if possible, without compromising patient safety. For a marketed drug, or even a drug in late phase III testing, it is not necessary to collect data on all adverse drug events (ADEs). It would be possible, for example, to collect data only on ADEs that lead to changes in drug dose, medication

switching, an intervention, and/or a premature termination of a patient from a study.

- Use nested designs. For example, answer secondary questions in subpopulations if the power is adequate.
- Report variances around economic measures. Variance around economic measures has generally been less adequately described than that around clinical outcomes, limiting the usefulness of this data for decision-makers.

Under What Circumstances Can Observational Studies Be Used to Assess Questions of Value?

Discussion here is limited to observational studies that use primary data sources. The use of administrative databases in economic evaluations is addressed by Panel 3. Primary data in this context refers to data that are derived from prospectively defined study questions that specify the particular endpoints of interest. The data may then come from a number of sources, including the medical record or a relevant database.

Observational studies should be considered as a primary approach when the following criteria are met:

- The effect size is expected to be large.
- The cost-effectiveness of an RCT is questionable. (Many of the following circumstances contribute to the unacceptable cost of an RCT.)
- The primary outcome is a rare event. Observational studies may be a more appropriate approach for rare outcomes requiring very large sample sizes. For example, if the true rate of an event is 0.001 and you want to detect a two-fold increase in risk, it would be necessary to follow 40,000 individuals [6]. Randomized controlled trials of this size are prohibitively expensive for all but the most important questions.
- There are multiple appropriate comparators. Variation in clinical practice is a well-described phenomenon and there are often multiple drugs that can be used to treat a particular condition. Because relevant comparators typically differ across practice settings, it is unlikely that all possible drugs will be evaluated in a randomized controlled trial.
- Observation of the primary outcome requires extended follow-up. This is particularly true for studies in chronic diseases. Long-term complications of diabetes and rheumatoid arthritis, for example, often manifest 10–20 years after

onset of disease. It is neither easy nor practical to conduct randomized clinical trials for such long durations. In addition, if there is a considerable loss-to-follow-up over time (as is likely), then results from an RCT will suffer from validity as well as prohibitive cost. The only feasible solution under such circumstances is longitudinal observational studies.

- The intervention to be studied has been available for a long time and data are available from multiple sources including government, third-party payers, and managed care organizations and vendors.
- People are likely to refuse randomization (for example, comparing invasive procedures to noninvasive treatments).
- The primary goal of the study is to assess practice patterns, as is typically the case in cost of illness studies.

How Can Observational Studies Be Modified to Improve Their Usefulness to Inform Economic Decision-Making?

Although the use of good clinical research practices obviously applies to the conduct of both RCTs and observational studies, observational studies are more frequently criticized for methodological inadequacies. To address this concern, the International Society for Pharmacoepidemiology (ISPE) has developed guidelines for good epidemiology practices for drug, device, and vaccine research in the United States [7]. While acknowledging the limitations of nonexperimental data, the ISPE authors conclude that properly conducted epidemiologic studies provide valuable information about the relationship between drugs and human health. The good epidemiology practices (GEPs) provide guidance on those issues that are under the control of the investigator, specifically data quality, study design, and study conduct. The GEPs do not dictate specific research methods, but rather propose minimum practices and procedures to help ensure the quality of the data used in observational studies and to provide adequate documentation of the study design and methods.

We recommend a similar approach be adopted for the use of observational data in the conduct of economic evaluations. Emphasis should be placed on the importance of specifying a priori:

- the research question(s) (which has been hypothesis driven);
- the study design;
- the data collection methods;

- the data analysis plan;
- dissemination plans.

Additional recommendations include replication of studies in more than one environment, and adjustment for factors believed to be associated with assignment of treatment. Analytic rigor for observational studies involves adjusting for differences in baseline characteristics across treatment groups. The measures of association between intervention and outcome should take into account those factors that are likely to influence treatment selection or outcome.

Recommendations and Next Steps

The following areas would benefit from further methodological development:

- Develop new methods to account for protocol-related costs (particularly in studies where these costs cannot be equated across study groups).
- Develop consensus on the definition of usual care. With the recommendation that trials should adopt more naturalistic designs, it is anticipated that the use of “usual care” as a comparator will continue to increase.
- Explore methods to supplement intent-to-treat analyses in usual care trials. Not to minimize the importance of the intent-to-treat analysis, but there may be research questions important to decision-makers that may require an alternate or additional analysis. For example, because economic trials are typically open-label, there can be a tendency for differential rates of crossover between treatment arms. Often patients randomized to usual care have higher rates of crossover to the new drug treatment arm, resulting in patients receiving the same treatment (the new drug) in both arms. In this situation useful information can still be derived from subgroup analyses of completers.
- Address problems of pooling economic data from multiple study sites. This is particularly problematic in studies when significant differences in practice styles exist across study sites (e.g., urban and rural sites, Veterans’ Administration and managed care settings). This problem is exacerbated in international studies where data are collected across countries.
- Improve statistical methods for adjusting for selection bias, which is a major drawback of observational studies. There are several methods to reduce this bias, from commonly used techniques such as adjustment by analysis of

covariance or other multivariate techniques, to more sophisticated techniques such as instrumental variables, two-stage regressions, and propensity scores. More research is needed to determine when a particular technique is appropriate for pharmacoeconomic analysis.

- Use better methods for estimating variance around resource utilization and cost. Traditional approaches that assume independence of observations and a normal Gaussian distribution are unsuitable for calculating variance around economic parameters. Techniques such as bootstrapping and jack-knifing can be used in these scenarios. These resampling methods are also useful in estimating variances around cost-effectiveness ratios.
- Conduct systematic comparisons of RCTs and observational studies of the same interventions. It would be useful whenever an RCT is used to answer a pharmacoeconomic question to also carry out observational studies of the same issues to look at correlation, to explain differences, etc.
- Explore approaches from other disciplines (e.g., psychology, sociology, marketing research) to enhance current methods, particularly in the areas of data collection, instrumentation, and analytic techniques.
- Measure resource utilization in large simple trials. Resources currently dedicated to marketing studies could be reallocated for the identification and collection of relevant economic data.
- Create better methods of measuring direct medical costs that are not routinely captured (e.g., nursing time, telephone care).
- Create better methods of measuring relevant indirect costs (e.g., caregiver, lost productivity). Pharmacoeconomic analyses are often conducted from a payer's viewpoint, ignoring indirect costs. Yet time lost from work, lost productivity, caregiver expenses, etc., are important to the patient and to society. There needs to be a greater appreciation of these cost drivers, and simple, easy methods to measure and document them.
- Encourage inclusion of standardized outcome measures in the evolving electronic medical record.

Summary

Useful pharmacoeconomic data can be derived from RCTs and observational studies. There is, however, no single pharmacoeconomic study that will establish the value of an intervention. Different audiences bring a variety of perspectives, as well as definitions of value, to the generation, framing, and interpretation of research questions. Establishing value can result from the synthesis of data from a variety of sources, including RCTs, observational studies, and modeling, with appropriate attention to the weaknesses and limitations of each kind of data.

The most important factors for a researcher to consider in designing a clinical trial to answer questions of value are knowing the precise question(s) that the study is required to answer and knowing the informational needs of the target audience. Over the long term, it is recommended that effort be directed towards better consensus and a greater degree of standardization of definitions and methodologies for assessing health economic questions through clinical studies.

References

- 1 Drummond M, Davies L. Economic analysis alongside clinical trials: revisiting the methodological issues. *Int J Tech Ass Health Care* 1991;7:561-73.
- 2 Drummond M. The future of pharmacoeconomics: bridging science and practice. *Clin Ther* 1996;18:969-78.
- 3 Haycox A, Drummond M, Walley T. Pharmacoeconomics: integrating economic evaluation into clinical trials. *Br J Clin Pharmacol* 1997;43:559-62.
- 4 O'Brien B. Economic evaluation of pharmaceuticals: Frankenstein's monster or vampire of trials? *Med Care* 1996;34:D599-DS108.
- 5 FTC Bureau of Economics and Consumer Protection. In the Matter of Pharmaceutical Marketing and Information Exchange in the Managed Care Environments: Public Hearings. Docket no. 95-N0228. Washington, DC: Federal Trade Commission, 1996.
- 6 Faich GA, Guess HA, Kuritsky JN. Postmarketing surveillance for drug safety. In: Cato AE (ed.), *Clinical Drug Trials and Tribulations*. New York: Marcel Dekker, Inc., 1988.
- 7 International Society for Pharmacoepidemiology. Good research practice guidelines. *Pharmacoepidemiol Drug Safety* 1996;5:333-8.