BIO
BIOTECHNOLOGY
INDUSTRY ORGANIZATION

The Complexities of Comparative Effectiveness

Ted Buckley, Ph.D.
BIO Director, Economic Policy
October 25, 2007

1201 Maryland Avenue, SW • Suite 900 • Washington, DC 20024-2149 • 202.962.9200 • Fax 202.488.6301 • www.bio.org

Executive Summary

The purpose of this white paper is educational. It is designed to impart a better understanding of what is meant by the term comparative effectiveness and to communicate the complexities that surround it. The primary purpose of comparative effectiveness research is to inform healthcare related decisions. While comparative effectiveness can be a valuable tool to inform the decision making of the clinician and the patient, it can not provide simple, "one size fits all" answers. The paper demonstrates that the seemingly simple idea of comparative effectiveness is actually quite complex.

The issues surrounding the implementation of comparative effectiveness studies include, but are not limited to:

- There is no standard definition for comparative effectiveness.
- Given that healthcare providers and patients do not currently utilize all relevant information, it is unclear how findings from new comparative effectiveness studies will impact clinical practice.
- The types of studies that may be used to produce comparative effectiveness results face significant challenges and have many limitations.
- Population-based study results may not be applicable to the needs of the individual patient.
- The absence of finding statistically significant differences among treatments does not mean that the outcomes of the compared treatments are the same.
- Patient treatment self-selection may affect comparative effectiveness results.
- Different perspectives (e.g., societal, patient, employer) may yield different results.

For many of these issues, there is no single correct way to conduct the study. Rather, the strategy that one should employ is dependent on who is asking the question, the patient population under evaluation and for what purpose the answer will be utilized.

The paper is not an exhaustive review of comparative effectiveness. In particular, it does not aim to analyze the existing international implementations of comparative effectiveness or the practical institutional difficulties and potential pitfalls of implementing comparative effectiveness. Rather, the paper is focused on examining some of the conceptual complexities inherent in comparative effectiveness analysis and should serve as a basis to begin understanding the topic of comparative effectiveness and the related complexities.

I. Introduction

The purpose of this white paper is educational. It is designed to impart a better understanding of what is meant by the term comparative effectiveness and to communicate the related complexities in execution and application. The primary purpose of comparative effectiveness research is to inform healthcare related decisions. As is discussed later in the paper, depending on the definition that is used, comparative effectiveness may or may not include a cost component in the analysis.[1] Many have expressed the hope that a comparative effectiveness study will come up with the answer as to what a clinician should or should not do in a particular case. The thought is that an analysis will compare two or more treatments and come up with a number or a simple graphic, in much the same way as Consumer Reports' rates cars or Zagat's rates restaurants. That is, it is hoped that the analysis will provide the answer as to which treatment is best for all patients. Unfortunately, such hope is misplaced.

While comparative effectiveness can be a valuable tool to inform the decision making of the clinician and the patient, it can not provide simple, "one size fits all" answers. As this paper will demonstrate, the seemingly simple idea of comparative effectiveness is actually quite complex. Further, there are many issues regarding the implementation of comparative effectiveness studies – ranging from what perspective to take to the type of data to use to conduct the study – that will affect the answer produced by the analysis. For many of these issues, there is no single correct way to conduct the study. Rather, the strategy that one should employ is dependent on who is asking the question, the patient population under evaluation and for what purpose the answer will be utilized.

As was stated previously, the purpose of this paper is educational. In particular, it does not aim to analyze the existing international implementations of comparative effectiveness or the practical institutional difficulties and potential pitfalls of implementing comparative effectiveness. Rather, the paper is focused on examining some of the conceptual complexities inherent in comparative effectiveness analysis and should serve as a basis to begin understanding the topic of comparative effectiveness and the related complexities.

The paper is structured as follows. The first section explores proposed definitions for comparative effectiveness. The second section examines some of the issues that should be deliberated when undertaking a study. The third section discusses the different

[1] The inclusion of a cost component to the analysis leads to further complexities on top of the complexities associated with comparing different treatment options. Some of these complexities are discussed in greater detail in the paper.

perspectives that may be utilized when conducting a comparative effectiveness study. A brief conclusion follows.

II. Definitions of Comparative Effectiveness[2]

There is currently no standard definition of comparative effectiveness. Neither the ISPOR Book of Terms nor AcademyHealth's Glossary of Terms Commonly Used in Health Care lists the term comparative effectiveness. Wilensky discusses comparative effectiveness at length in her 2006 Health Affairs' article, "Developing a Center for Comparative Effectiveness Information." However, no explicit definition for comparative effectiveness was given.

Thus, when approaching a discussion or outlining a position concerning comparative effectiveness, care must be used to make sure that there is a common understanding of what is meant by the term comparative effectiveness.

While no standard definition of comparative effectiveness has emerged, the following definitions of comparative effectiveness have been used:

> **Sean Tunis' (Center for Medical Technology Policy) definition:** "…a set of analytic tools that allow for the comparison of one treatment – drug, device, or procedure - to another treatment on the basis of risks, benefits, and potentially, cost."[3] The tools include: systematic reviews of evidence; modeling; retrospective analyses of databases (either electronic health records (EHRs) or administrative data used to process claims); and prospective, but non-randomized controlled trials research (adaptive trials, practical trials, etc). The research setting is "real world" health care interactions, rather than randomized and controlled trials.[4,5]

---

[2] Other related definitions can be found in Appendix One.

[3] http://www.avalerehealth.net/conferences/EBM_Conference_Summary_5.22.07.pdf - last accessed on August 16, 2007, page 6.

[4] Note: Randomized controlled trials are designed to achieve the most robust scientific data about the safety and efficacy of an intervention or product and are required for regulatory approval.

[5] http://www.allhealth.org/briefingmaterials/Tunis4-27-07-699.pdf

Congressional Budget Office (CBO).[6] **2007.** ***Research on the Comparative Effectiveness of Medical Treatments: Options for an Expanded Federal Role:*** a comparison of the impact of different options that are available for treating a given medical condition for a particular set of patients. Such studies may compare similar treatments, such as competing drugs, or they may analyze very different approaches, such as surgery and drug therapy. The analysis may focus only on the relative medical benefits and risks of each option, or it may go on to weigh both the costs and the benefits of those options.

One company's outcomes research group: Comparative clinical effectiveness studies assess the benefits and risks of one treatment (e.g. medication, procedure, etc) for a clinical condition compared to one or more other treatments for the same condition.

Institute of Medicine.[7] **2007.** ***Learning What Works Best: The Nation's Need for Evidence on Comparative Effectiveness in Health Care:*** Within the overall umbrella of clinical effectiveness research, the most practical need is for studies of comparative effectiveness, the comparison of one diagnostic or treatment option to one or more others. In this respect, primary comparative effectiveness research involves the direct generation of clinical information on the relative merits or outcomes of one intervention in comparison to one or more others. Secondary comparative effectiveness research involves the synthesis of primary studies (usually multiple) to allow conclusions to be drawn. Secondary comparisons of the relative merits of different diagnostic or treatment interventions can be done through collective analysis of the results of multiple head-to-head studies, or indirectly, in which the treatment options have not been directly compared to each other in a clinical evaluation but reside in larger databases. Conclusions utilize inferential adjustments based on the relative effect of each intervention to a specific comparison, often a placebo.

The definitions have the following in common:

1. Comparison of one treatment to one or more other treatments
2. Comparison of treatments is not limited to medications
3. Inclusion of both risk and benefits in the assessment – explicitly mentioned in the Tunis, CBO and company definition – alluded to in the IOM definition.

---

[6] http://www.cbo.gov/ftpdocs/82xx/doc8209/Comparative_Testimony.pdf - last accessed on August 16, 2007.

[7] http://www.iom.edu/Object.File/Master/43/390/Comparative%20Effectiveness%20White%20Paper%20(F).pdf – last accessed on September 5, 2007.

The definitions differ in the following ways:

1. Type of studies proposed for the assessment.[8]
    a. Both Sean Tunis and IOM list the types of studies to be utilized in comparative effectiveness research.
        i. Randomized controlled trials (RCT): Sean Tunis specifically rejects the use of randomized controlled trials, while the IOM definition does not.
        ii. Secondary data: Tunis specifically mentions retrospective analysis of databases, whereas the IOM definition implies its use among the types of research studies.
    b. The company's definition, which is less detailed, does not list the types of studies. The CBO also does not list the types of studies; however, it does outline different types of analyses that might be undertaken.
2. Whether or not cost is included in the analysis.
    a. Both Sean Tunis' definition and the CBO definition say that cost may be included.
    b. The IOM definition does not explicitly mention cost in the definition; however, cost is mentioned in other parts of the document. Therefore, it is implied that cost may be included in the analysis.
    c. The company definition does not mention cost.

III. Issues to Consider

When considering the generation of comparative effectiveness evidence, there are a number of issues that need to be considered. Nine issues are discussed below. None of the discussions are meant to be an exhaustive treatise on the issue discussed. Rather, the discussions serve to raise awareness that the issue exists.

Changing practice patterns

The purpose of comparative effectiveness studies is to provide evidence concerning the appropriateness of a particular health treatment for a particular disease. However, even when knowledge is widespread, the knowledge may not influence practice patterns. As several reports have shown, changing practice patterns can be difficult. For instance,

---

[8] Note: In a hierarchy of clinical evidence, as described in the article, "Evidence-based Medicine: A New Approach to Teaching Medicine" (JAMA ,1992), evidence based medicine clearly defines as appropriate the research methods which have the most scientific rigor and quality versus those methods which introduce biases, system errors, and confounding factors or background issues that appear to impact conclusions but may not be related to the interventions under evaluation.

- According to the Agency for Healthcare Research and Quality (AHRQ) 2006 National Healthcare Quality Report:[9]
    - In 2003 only 66.1% of smokers with routine physician visits reported that their healthcare provider advised them to quit.
    - In 2004, influenza vaccination status assessment/vaccine provision occurred in 43.1% of Medicare patients who received care for pneumonia while pneumococcal vaccination status assessment/vaccine provision was only 43.5%.

These scores have shown improvement since 2000. However, a substantial number of patients are still not getting standard care. Given this, thought must be given as to how findings from comparative effectiveness studies will be used by providers and patients when currently there is information that is not being incorporated into clinical practice.

Impact of a fragmented US healthcare delivery system

The U.S. healthcare delivery system is fragmented. The transition of care from one setting to the next or from one clinician to the next is far from optimal. The incentives of the healthcare delivery system are often not aligned to ensure the best quality of care for the patients. Indeed, many who discuss the issue of comparative effectiveness believe that one should examine different healthcare delivery models to assess the risks and benefits of different delivery models on patient outcomes in much the same way as one would examine the risks and benefits of different treatment options on patient outcomes. Further, others wonder how one can compare and evaluate interventions without considering the environment and behavior of users within the context of the patient experience. For instance, a novel product may offer tremendous value to the long term outcome of a patient, but the benefit may not be realized when administered inappropriately or without monitoring the patient if this is a recommended behavior.

Type of studies[10]

Randomized controlled trials (RCTs) are considered to be the "gold standard" of robust scientific evidence. Indeed, the Food and Drug Administration requires RCTs to prove safety and efficacy. The design is well described with controls to exclude potential sources of systematic bias or confounding factors – examples include the use of inclusion or exclusion criteria and use of blinded populations so investigator or analytical bias does not influence results.

---

[9] http://www.ahrq.gov/qual/nhqr06/nhqr06report.pdf - last accessed on August 16, 2007.
[10] See also: Garrison, Louis P. et al. "Using Real-World Data for Coverage and Payment Decisions: The ISPOR Real-World Data Task Force Report." Value in Health forthcoming.

While RCTs are critical to determine the efficacy of a drug, they may be less informative in determining the effectiveness of a drug. That is, how does a treatment work in a real world setting as opposed to the clinical research setting. For instance, imagine that one wants to determine which of two treatments works better. For simplicity, let us assume neither have side effects. One treatment needs to be taken 5 times per day for 4 days; while the other treatment needs to be taken 1 time per day for 7 days. Under a controlled clinical trial environment, the 5 time per day treatment may prove to be better in treating the condition; however, in the real world, given the challenges associated with patient compliance, the 1 time per day treatment may prove to be better for treating the condition. Further, RCTs utilize a select population chosen to demonstrate reproducible and confirmatory evidence of the safety and efficacy of the product. This allows practitioners to achieve confirmatory validation of the intervention or product's performance when used in a group of fairly homogeneous patients. In the real world the combination of multiple medical conditions or multiple treatments may require provider decisions to be made for not only similar patients, but also for patients who are unlikely to be representative of the patients evaluated when generating the scientific evidence.

Given the difference between efficacy and effectiveness, some suggest using retrospective studies to determine the benefits of different treatments. There are several advantages of retrospective studies. The first is cost. Because the data has already been generated, there is no need for costly clinical trials or additional patient exposure. Secondly, the outcomes represent the effect of a treatment in the real world. However, retrospective studies come with their own set of limitations. First, the data that are used for retrospective studies are generally administrative data that were not designed for an effectiveness study. Thus, data elements that may be needed to adjust for differences in multiple medical conditions and overall health risks may be missing or incorrectly coded. Second, there may be bias for certain patients to be selected for treatment for which one can not adjust with the previously collected data. That is, patients may have received a certain treatment based on some observed characteristics not reflected in the administrative data. This selection bias could result in an analysis with spurious results.

Databases derived from health care claims from interactions in the health care system are not designed as research tools. They are used to document the patients' experience, diagnosis, treatment and outcome in the health care system for the purposes of paying respective providers. There are important variables that may capture the patients' experience which may not be reflected in payment claims. Databases used to describe the medical encounter (e.g. electronic medical records) may have more relevant detail but are not currently standardized or widespread enough to allow researchers to compare similar patients.

Population-based study results may not be applicable to the needs of the individual patient

The purpose of comparing two treatment options is to inform the decision making processes of those involved with patient care – both the patients and the healthcare providers. The hope is that increased information will lead to better decision making and thus better outcomes. While this is true in the aggregate, it must be recognized that the results generated from both RCT and research studies represent the "average" result. Thus, the general or "average" results of a study may not inform the decisions regarding an individual patient who is not "average". For instance, if two treatments A and B have been compared, and treatment A has been found to be more effective than treatment B, this means that treatment A will work on average better or be safer than treatment B for a random patient. However, that does not imply that treatment A will always work better for all patients. Rather, there may be a subset of patients for whom treatment B will be more effective. It is imperative for decision-makers to realize the limitations of population-based study results.

Lack of findings of statistically significant differences among treatments

A study that finds no difference in outcomes between two treatment options does not mean necessarily that there is not a difference in the outcomes. What the study has found is that the statistical analysis of the data in that particular study can not detect a difference. This is a subtle, yet important point. This lack of a statistically significant finding can be caused for a number of reasons including lack of a large enough patient sample, flaws in the study design or that there actually is no difference in the outcomes.

Conversion to a common outcome metric

While the conversion of different treatment outcomes into a common metric, such as the quality-adjusted life year (QALY), allows for the comparison of different treatments for the same disease state – and even comparisons of treatments of different disease states, there are issues involved in the conversion of outcomes to a common metric. The methodologies used to compare outcomes can be controversial and are imprecise at best. A fuller discussion of these issues can be found in Appendix 2.

The six issues discussed above apply whether or not one considers an economic component in the evaluation. When adding an economic dimension to the analysis, it is important to recognize that the economic evaluations are supplementary to evidence-based medicine evaluations. The economic evaluations are based on the clinical evidence and depend on the robustness of the systematic evaluation of the research. Given the absence of clear standards in the approach to evidence based medicine, the absence of

economic evaluation standards is equally as problematic for conclusions derived from the underlying science. Indeed, the three issues below (which generally apply only when costs are modeled into the analysis) demonstrate how the addition of cost considerations in the analysis leads to even more complexity in the analysis and less certainty concerning the conclusions of the study.

Impact on Innovation

This issue pertains to those therapies which require large amounts of research and development before being brought to market (e.g., medical devices and pharmaceuticals.) Comparative effectiveness generally focuses on a static world – that is, evaluation of a current therapy in the current state of the world in the short term. Evaluations by foreign agencies have led to decisions where an innovative therapy is not covered by a country's health service. However, these evaluations, which are effectively a cost hurdle, do not take into account the effect the decisions may have on research and development of future products. That is, they do not take into account future patients and whether the decisions implemented now will have deleterious effects on the availability of future therapies.

It is interesting that agencies such as the UK's National Institute for Health and Clinical Excellence (NICE) sometimes appear to consider the static case and do not necessarily utilize a dynamic perspective in their evaluations. It is interesting because these countries issue patents, which are designed to encourage and to reward innovation. Patents balance the need to reward innovators and broad societal access to an innovation. The patent bestows the inventor with a property right, which means that the invention cannot be copied or used without the permission of the inventor. One result of this exclusive ownership is that the price of the invention that can be charged is higher than it would be in a competitive market. The inventor, therefore, makes a higher profit for the invention that has been patented. The ability to charge the higher price for innovative products provides innovators with an incentive to develop innovative products.

The cost hurdle that countries have erected is a type of price control on an innovative product. The price control benefits present day patients; however, there is a cost borne by future patients.[11] That cost is fewer innovative therapies due to a lower level of research and development undertaken.[12]

[11] Please note: If the effect of the agency is to encourage the introduction of an innovative therapy at a lower cost, the present day consumers benefit. If, however, the agency refuses to reimburse the innovative treatment, then the present day patients are worse off because they do not have access to the innovative therapy. In both of these cases the agency hurts future innovation, and thus, it hurts future patients.

A dynamic perspective, or long-term perspective, is one that takes into account the future and the effects that a decision will have on future innovation. The static perspective focuses only on the present. It is easy to understand that taking a dynamic perspective can lead to a different conclusion than may be reached when taking a static perspective.

Product life cycle approach to comparative effectiveness[13]

When an economic evaluation is included in a comparative effectiveness analysis, the analysis of the treatment typically examines a single, specific indication. However, there are numerous innovative therapies that receive regulatory approval for one indication and then later receive supplemental approval for new indications. Cancer therapies are a good example of this phenomenon. A cancer therapy may be approved for treatment use against a cancer that has metastasized and then later obtain a supplemental approval for use as an adjuvant therapy. In the case where the cancer has metastasized, the goal of the use of the cancer therapy is to prolong life but a cure of the cancer is often not possible. However, when used in the adjuvant setting, a cure may be more likely. Thus, while a product life cycle approach is seldom undertaken when conducting a comparative effectiveness study, failing to perform this more inclusive type of analysis may result in a significant underestimate of the benefits that a therapy can provide.

The effect of patients' self-selection (patients' choice of treatment) on comparative effectiveness[14,15]

When one approaches comparative effectiveness research, it is important not to regard the patient population as homogeneous. Thinking of a patient population in monolithic terms ignores the fact that every patient is different. Each patient has his/her own preferences that need to be taken into account by the clinician when prescribing a course of treatment. Patients place different values on different levels of health status and also

---

[12] U.S. Department of Commerce, "Pharmaceutical Price Controls in OECD Countries: Implications for U.S. consumers, Pricing, Research and Development, and Innovation" Washington DC. 2004

13 Garrison Louis P. and David L. Veenstra. "The Economic Value of Innovative Treatments Over the Product Life Cycle: The Case of Trastuzumab" Poster Presentation ISPOR 9th Annual European Congress.

[14] Huang, Elbert S. et al. "The Impact of Patient Preferences on the Cost-Effectiveness of Intensive Glucose Control in Older Patients with New-Onset Diabetes." Diabetes Care 29.2 (2006): 259-264.

[15] Meltzer, David O. et al. "Effects of Patient Self-Selection on Cost Effectiveness of Diabetes Therapy by Age." Society for Medical Decision Making. Presentation October 17, 2006. See the following URL for an overview: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=996268

show different tolerance levels for side effects. Therefore, when evaluating different treatment options, it is important to recognize the diversity within the patient community. Indeed, researchers have found that taking patients' self-selection for treatment options (that is, patients' choice of treatment options) into account when evaluating a treatment results in very different analytical outcomes than when employing population averages because of the different valuations of health status and tolerance for different treatment courses by individual patients. This once again underscores the importance of the patient and clinician dialoguing about treatment options and preferences before deciding on a course of action.

The following illustrative example may help to underscore this issue. Imagine 3 women who 2 years ago completed treatment for breast cancer. Recently, they have discovered that the cancer has recurred in the breast and spine. One woman is 35 years old, married and has two children under the age of 10. Another woman is 65, has just retired and has 2 grown children and several young grandchildren. She is looking forward to spending time with her husband who has also just retired. The third woman is 85 years old, a widow and has two children approaching retirement age and several older grandchildren. One can imagine that each of these women with different menopausal status and potentially different bone mineral density will have different preferences and different treatment goals. Even if the three women were all the same age and had the same life situations, their preferences and utilities regarding treatment choices could be very different.[16] These different preferences and utilities could affect the patients' preference and assessment of the value of a particular treatment. What might be appropriate for one patient may not be appropriate for another because of the individual patients' preferences. If a clinician or agency relied upon the results of a comparative effectiveness study which was based upon population averages, a clinician or agency could decide on a treatment that is inappropriate for one or all of the above individuals.

IV. Perspectives

When discussing the benefits and/or costs of an intervention there are a variety of perspectives that can be undertaken. The perspectives range from a societal perspective to an individual perspective. This section describes the different perspectives and what should be included or excluded when examining one or more treatments. Because some include a cost component in their definition of comparative effectiveness, costs will be included in the discussion. However, it is possible to utilize the different perspectives without including costs.

[16] I am indebted to Margaret Kirk of the Y-Me National Breast Cancer Organization for this example.

We examine four perspectives: Societal, patient, employer and payer. While the perspectives are illustrative, they are not an exhaustive list of perspectives that can be examined.

Societal

A societal perspective includes all the benefits and costs associated with a treatment. It does not represent only one particular stakeholder in the system, but rather attempts to capture the net effects on society as a whole. Examples of effects that would be included in the analysis follow:
Treatment benefits: The benefit to the individual would be taken into account. In addition, if the treatment had externalities associated with it these would also be included in the analysis (e.g., an individual being vaccinated for influenza lowers the risk that others will catch influenza).
Productivity: If the treatment increased an individual's productivity (e.g., allowed one to return to work or have fewer missed days) the increase in both the individuals' and business' productivity would be included in the analysis. Double counting of the increase of productivity is avoided. Further, if there are caregivers who are no longer required, this productivity gain would also be included.
Costs: Because cost of treatment can represent just a transfer of money, only the marginal cost of treatment should be included in the cost assessment. The time cost of the treatment should also be included in the cost. This would account for productivity lost during the treatment.
Time horizon: The benefits and costs that accrue over the long run as a result of the treatment that was undertaken would be included in the analysis. Thus, if daily medication is associated with prevention of an illness in the future (e.g., blood pressure medication and stroke) the benefit of not having a stroke in the future would be included in the analysis.

While the societal perspective takes into account all of the costs and benefits, the other perspectives are more limited in which benefits and costs are included.

Patient

A patient perspective includes only the benefits and costs of a treatment that impact the patient. It represents only the patient; effects on other parts of society are not included in the analysis.
Treatment benefits: As with the societal perspective, the benefit to the individual would be taken into account. However, no externalities associated with the treatment would be included.
Productivity: The benefit that accrues to a business from an increase in the individual productivity would not be counted. The benefit to the individual from an increase in productivity would be counted. However, if the individual is paid whether or not he or she is at work (e.g., sick time

or disability), then only the net increase to the individual would be counted. Further, the productivity gain to the caregivers would not be counted.
Costs: Concerning the cost of treatment, only the part of the cost for which the individual pays out-of-pocket would be included in the cost assessment. Other costs that the individual incurs would also be included in the assessment such as time costs of treatment (e.g., travel to and from the doctor's office.)
Time horizon: As in the societal perspective, the time horizon that is utilized in this analysis is the long run. Thus, if daily exercise is associated with prevention of an illness in the future, the benefit of not having the illness would be included in the analysis as would the daily cost (e.g., time) of the exercise regime.

Employer

An employer perspective includes only the benefits and costs of a treatment that impacts the employer. It represents only the employer; effects on other parts of society are not included in the analysis.
Treatment benefits: The benefit to the individual and externalities are not included in this analysis.
Productivity: The benefit that accrues to the business because of increased productivity is included in the analysis. In addition, if other employees are more productive because of the treatment, these benefits are also included (e.g., externalities associated with vaccination of one employee results in a reduced risk of other employees catching the flu and the company incurring the cost of lost employee productivity).
Costs: Concerning the cost of treatment, only the part of the cost for which the employer pays would be included in the cost assessment. Also, any other costs that the employer incurs would be included in the assessment such as lost productivity due to the time away from work due to the treatment.
Time horizon: The time horizon used to calculate the benefits and costs is the expected length of employment.[17] That is, if a benefit from a treatment is expected to take place 10 years in the future (e.g., reduced risk of disability and thus, reduced loss of productivity), but the employee is only expected to be employed with the company for 5 years, the employer perspective would not take this benefit into account.

[17] If there is a defined pension benefit, then this may also be included. If the intervention will increase life expectancy of the individual and thus, s/he will draw a pension from the company for a longer period of time, this would be included in the company perspective.

Payer A payer perspective includes only the benefits and costs of a treatment that impact a payer. It represents only the payer; effects on other parts of society are not included in the analysis.
Treatment benefits: The benefit to the individual and externalities are not included in this analysis.
Costs: Concerning the cost of treatment, only the part of the cost for which the payer pays would be included in the cost assessment.
Productivity: The benefit that accrues to the individual or business because of increased productivity is not included in the analysis.
Time horizon: The time horizon used to calculate the benefits and costs is the expected length that an individual is expected to be enrolled with the insurance plan. That is, if a benefit from a treatment is expected to take place 10 years in the future (e.g., reduced risk of disability and thus, reduced loss of productivity), but the individual receiving the treatment is expected to switch to a different payer within 5 years, the payer perspective would not take this benefit into account.

Below are some of the ways in which the perspectives differ:

1. Treatment benefits to individual
    a. Both the societal and individual perspectives take the benefit of the treatment to the individual into account.
    b. The employer perspective takes the benefit of the treatment to the individual into account only as it impacts the employer.
2. Treatment effects on others
    a. The societal perspective accounts for the effects of the treatment on other individuals including care givers and in the case of a contagious condition, reduced risk of other individuals contracting the disease.
    b. The employer perspective only includes the treatment effects on others if it impacts the business.
    c. The individual perspective does not include the treatment effects on others.
3. Time horizon
    a. Both the societal and individual perspectives take into account the benefits and costs of the treatment over the long run.
    b. The employer perspective examines the impact of the treatment during the time period for which the individual will be employed by the company and if there is a defined pension benefit, the treatment effect on the duration which the patient will draw a pension may also be included
    c. The time horizon used in the payer perspective to calculate the benefits and costs is the expected length of time that an individual is expected to be enrolled with the insurance plan.

As is apparent from the discussion above, the societal perspective attempts to account for all of the benefits and costs associated with a treatment option over the long run. While the most difficult to undertake, it attempts to provide the true overall benefit and cost of all those affected.

V. Conclusion

Many believe that comparative effectiveness provides clinicians and patients with a simple answer as to what the best treatment option is. Unfortunately, as has been discussed, there are many issues regarding comparative effectiveness that render it much more complicated than it first appears. Indeed, the complication even extends to the fact that there is currently no widespread agreement on the definition of comparative effectiveness. If one includes an economic component in the analysis, the study becomes even more complex and casts more doubt on the conclusions that are reached. Further, there are a number of issues that need to be examined in more depth. These include, but are not limited to, impact on innovation, the role of patient self-selection, the type of study to undertake and the issues regarding how to encourage the use of clinical information that may be generated. Finally, a variety of perspectives could be considered when designing a study which will influence what should and should not be included in a particular study.

Once again, it must be stressed that this paper should not be viewed as an exhaustive review of comparative effectiveness and the associated issues. Rather, it should be seen as an introduction to the nuances surrounding this topic.

Appendix One: Related Definitions[18]

Main Economic Terms

Whereas comparative effectiveness may or may not consider cost when comparing treatments, there are 5 main economic analyses of health care interventions that seek to compare different alternatives in terms of both costs and outcomes (cost-effectiveness, cost-benefit, cost-utility, cost-consequence and cost-minimization.) They differ primarily in how outcomes are measured and assessed. Cost-effectiveness utilizes one single, natural parameter such as symptom free days to compare alternatives. Cost-utility and cost-benefit both convert the health consequences into equivalent units. In the case of cost-utility that may be quality adjusted life years (QALYs); whereas, cost-benefit translates the health benefit into a monetary unit using a technique such as willingness to pay. Alternatively, cost-consequence simply lists all the consequences for the decision maker. Cost-minimization is used in comparing alternatives where the outcomes are identical.

While the conversion of different treatment outcomes into a common metric allows for the comparison of different treatments for the same disease state – and even comparisons of treatments of different disease states, there are issues surrounding the conversion of outcomes to a common metric. These are discussed in more detail in the definitions of the "QALY" and "Willingness to Pay." All 5 of these analyses can be performed with different perspectives in mind ranging from societal to individual. A fuller discussion of the different perspectives can be found in section IV of the paper.

Cost-Effectiveness Analysis — A tool in which the effects and costs of a health intervention are compared to a placebo or to another treatment, or even to compare two treatments for unrelated conditions. The effects are health outcomes and must be expressed in the same units (e.g., life years gained); however, they are not expressed in monetary units as in cost-benefit

18 The following terms have generally agreed upon definitions. The definitions were compiled from the following sources:

Academy Health. Glossary of Terms Commonly Used in Health Care 2004 edition.
Berger, Marc L. Health Care Cost, Quality and Outcomes: ISPOR Book of Terms International Society for Pharmacoeconomics and Outcomes Research. 2003.
Drummond, Michael F. et al. Methods for Economic Evaluation of Health Care Programmes 2nd edition. Oxford University Press. 1997.
Gold, Marthe R. et al. editors. Cost-Effectiveness in Health and Medicine Oxford University Press. 1996.

analysis. If there are just two alternatives being compared, a comparison is made by dividing the difference in cost of the two treatments by the difference in outcomes. As with the cost-benefit analysis, the conversion of health improvements (outcomes) into a common metric can be quite challenging (e.g., one treatment increases the quality of life but not survival time, but another treatment increases survival time.)

Cost-Benefit Analysis

Derived from economic theory, this analytical technique lists and compares the net costs and net benefits of a health care intervention. Both the net costs and net benefits must be expressed in monetary units. In order to value the net benefits, the improved health outcomes must be expressed in monetary units. A variety of techniques can be used to value improved health outcomes, one of which is willingness to pay. The bottom line of the analysis is net benefit. It may be used to compare a treatment to a placebo or to another treatment, or even to compare two treatments for unrelated conditions. A key problem with the methodology is the conversion of health improvements into monetary units. Many of these analyses require significant assumptions in the models that are employed.

Cost-Utility Analysis

A comparison of different health interventions where the health outcomes are translated into units of utility (e.g., QALYs.) The analysis is expressed in terms of a ratio of the incremental cost of the two alternatives to the incremental health effects of the two alternatives. The result of the analysis is the "cost/QALY" of the intervention. It is used to determine the relative value of alternative health interventions. As with the cost-benefit and cost-effectiveness analyses, the conversion of health improvements into a common metric can be quite challenging. Note: some consider the cost-utility analysis to be a special type of cost-effectiveness.

The three terms above all have the same challenge of finding a common metric that describes the change in health status associated with the treatment options that are being evaluated. Terms to describe health outcomes range from very specific measures (e.g., symptom free days) to more abstract terms such as life years gained, healthy life years gained, QALYs etc. Differences in the description of outcomes may further impair the ability to condense various studies into one common health outcome result.

Cost-Consequence Analysis	A comparison of alternative health interventions in which the outcomes and costs are listed without aggregating the results. Whereas, the cost-benefit ratio results in a single aggregated result, the cost-consequence analysis does not. There is no prescribed weighting system to indicate the relative importance of different benefits or costs. Thus, while this analysis avoids the problems of the conversion of different outcomes into a common metric, it leaves the comparison of outcomes that may be difficult to compare to the decision maker.
Cost-Minimization Analysis	When the outcomes are equivalent, this tool is used to compare the net costs between different health interventions. Because the outcomes must be equivalent, the value and use of this technique is very limited for assessing new medical interventions.

Related Terms

The following terms are closely related to those described above.

Contingent Valuation	Determining an individual's maximum willingness to pay for a good or service that is not available in the marketplace. The determination is often made through hypothetical survey questions. Drawbacks include need for a large sample size and starting point bias.
Cost-Comparison Analysis	Compare only the costs associated with two or more health care treatments. There is no inclusion of health benefits in the analysis. As with all

	accounting of costs, care must be taken to include all costs. Uncertainty exists in determining the proper allocation of certain costs (e.g., one time, shared and fixed costs.)
Cost-Identification Analysis	Identification of all relative costs and their importance.
Cost-of-Illness Study	A study to determine the total cost, including treatment costs, of a disease or health condition on society.
Effectiveness	Represents outcomes achieved from a treatment or health intervention in real, practical settings (e.g., the real world)
Efficacy	Represents outcomes achieved from a treatment or health intervention under ideal circumstances (e.g., clinical trials)
Evidence-Based Medicine	Based on systematic review of all available data usually in published domain or available from organizations, the identification of best evidence to inform decision making about the care of individual patients. The information may also be utilized with same or different conclusions to describe best evidence to inform decision making about the care of populations. Evidence-based medicine (EBM) requires that physicians have access to critical, unbiased reviews of all currently available information. The goal is to enable physicians to bring unbiased sources of information into the patient encounter and use them in the clinical decision-making process. EBM will take into account specific patient characteristics and preferences. Note: many consider comparative effectiveness to be a type of EBM.
Evidence Synthesis	A meta analysis – also referred to as secondary clinical effectiveness research – which is a

structured assessment of evidence from multiple primary studies to develop a conclusion.

Health Technology Assessment Also known as HTA. It is an evaluation that examines the effects and impacts of health care technology or treatment. The effects and impacts are broad and can include, but are not limited to, safety, efficacy, effectiveness, economic, political and ethical. It is intended to educate decision makers as to the direct and indirect consequences of a given technology of treatment.

Primary Clinical Effectiveness Research Design and implementation of structured research protocols to produce data on the results of one or more diagnostic or therapeutic interventions of interest. The evaluation can measure either effectiveness or efficacy.

Quality-Adjusted Life Year Combines gains or losses in both quantity of life (mortality) and quality of life (morbidity) into a single measure that is years of life saved by a health intervention, adjusted according to the quality of those years. The adjustment is made according to some evaluative measure. Typically, the range is from 0 to 1 where 0 is death and 1 is optimal health. Thus a year of life in non-optimal health would be rated somewhere between 0 and 1. The use of Quality-Adjusted Life Years (QALYs) enables comparison across diseases and programs. The precise score given to different levels of non-health are based on tools such as utility measures (e.g., standard gamble), Health Utilities Index, EuroQol EQ-5D. QALYs can be used in studies with or without an economic component. There are a variety of issues associated with the conversion of health status into QALYs. The first is that conversion of certain health states must be based on preferences. However, whose preferences should they be based upon - individual patient preferences, preferences of an informed public, or some other group? For instance the informed public might

	value the quality of life of a year in a wheelchair very differently than a person who is in a wheelchair. Another issue is that different tools may give systematically different results. The QALY does not differentiate whether a gain (loss) in QALYs comes from a small gain (loss) to a large number of people or a large gain (loss) to a small number of people.
Randomized Control Trial	A clinical trial in which subjects are randomly assigned to different treatment groups. The research study is well described with controls to exclude potential sources of systematic bias or confounding factors – examples include the use of inclusion or exclusion criteria, use of blinded populations so investigator or analytical bias does not influence results. The randomized control trial is an effort to identify the most robust knowledge about the treatment groups without systematic bias or confounding factors.
Retrospective Analysis	Analysis based on data that is currently available (i.e., that has been already generated). Generally, the data comes from sources such as insurance claims data or hospital discharge data. These data are usually derived from health care claims used to secure payment in health care systems. For that reason, their primary goal is to insure payment and not research on patients as an outcome. Because the data is usually not designed with research in mind, the data quality may be lacking important clinical data.
Willingness to Pay	An estimation of the maximum dollar amount an individual would pay to obtain a good, service or reduction in risk. Willingness to Pay (WTP) may be used in cost-benefit analysis to determine how much one is willing to pay for a certain health outcome. The estimation is made in order to measure the value of the good, service or reduction in risk to the individual. There are several issues

associated with using WTP. An individual's ability to pay constrains his or her WTP. Therefore, if there are health care treatments for diseases that affect primarily the wealthy, those outcomes and treatments will receive higher WTP than health care treatments for diseases that affect primarily the poor. If open ended questions are used to ascertain individuals WTP, the results will vary widely and include many non-responses. Alternatively, closed-ended questions may cause starting point biases.

Appendix Two: Conversion to a Common Metric

Quality-Adjusted Life Years (QALYs) are commonly used and enable comparison across diseases and programs. The precise score given to different levels of non-health are based on tools such as utility measures (e.g., standard gamble), Health Utilities Index, EuroQol EQ-5D. QALYs can be used in studies with or without an economic component. However, there are a variety of issues associated with the conversion of health status into QALYs. The first is that conversion of certain health states must be based on preferences. However, whose preferences should they be based upon - individual patient preferences, preferences of an informed public, or some other group? For instance the informed public might value the quality of life of a year in a wheelchair very differently than a person who is in a wheelchair. Another issue is that different tools may give systematically different results. The QALY does not differentiate whether a gain (loss) in QALYs comes from a small gain (loss) to a large number of people or a large gain (loss) to a small number of people.

If an economic component is part of the analysis, willingness to pay (WTP) may be used, in conjunction with or in addition to the QALY measure, to determine how much one is willing to pay for a certain health outcome, that is, to determine the value of the health outcome. The estimation is made in order to measure the value of the good, service or reduction in risk to the individual. There are several issues associated with using WTP. An individual's ability to pay constrains his or her WTP. Therefore, if there are health care treatments for diseases that affect primarily the wealthy, those outcomes and treatments will receive higher WTP scores than health care treatments for diseases that affect primarily the poor. If open ended questions are used to ascertain individuals' WTP, the results will vary widely and include many non-responses. Alternatively, closed-ended questions may cause starting point biases. As the description of WTP shows, the addition of economic considerations in the analysis leads to more complexity in the analysis and less certainty concerning the conclusions of the study.

Bibliography:

Academy Health. Glossary of Terms Commonly Used in Health Care 2004 edition.

Agency for Healthcare Research and Quality. "2006 National Healthcare Quality Report" AHRP Publication No. 07-0013. December 2006.

Berger, Marc L. Health Care Cost, Quality and Outcomes: ISPOR Book of Terms International Society for Pharmacoeconomics and Outcomes Research. 2003.

Drummond, Michael F. et al. Methods for Economic Evaluation of Health Care Programmes 2nd edition. Oxford University Press. 1997.

Garrison, Louis P. et al. "Using Real-World Data for Coverage and Payment Decisions: The ISPOR Real-World Data Task Force Report." Value in Health *forthcoming*.

Garrison Louis P. and David L. Veenstra. "The Economic Value of Innovative Treatments Over the Product Life Cycle: The Case of Trastuzumab" Poster Presentation ISPOR 9th Annual European Congress.

Gold, Marthe R. et al. editors. Cost-Effectiveness in Health and Medicine Oxford University Press. 1996.

Evidence Based Medicine Working Group. "Evidence Based Medicine: A New Approach to Teaching the Practice of Medicine." JAMA 268.17 (1992): 2420-2425.

Huang, Elbert S. et al. "The Impact of Patient Preferences on the Cost-Effectiveness of Intensive Glucose Control in Older Patients with New-Onset Diabetes." Diabetes Care 29.2 (2006): 259-264.

Institute of Medicine. "Learning What Works Best: The Nations' Need for Evidence on Comparative Effectiveness in Health Care"

Jena, Anupam B. and Tomas Philipson. "Cost-Effectiveness As a Price Control." Health Affairs 26.3 (2007): 696-703.

Meltzer, David O. et al. "Effects of Patient Self-Selection on Cost Effectiveness of Diabetes Therapy by Age." Society for Medical Decision Making. Presentation October 17, 2006.

Phelps, Charles E. Health Economics 2nd edition. Addison-Wesley Education Publishers Inc. 1997.

U.S. Department of Commerce, "Pharmaceutical Price Controls in OECD Countries: Implications for U.S. consumers, Pricing, Research and Development, and Innovation" Washington DC. 2004.

Wilensky, Gail. "Developing a Center for Comparative Effectiveness Information." Health Affairs 25.6 (2006): w572-w585