Acceptance and rejection: Cost-effectiveness and the working nephrologist

PHILIP A. MCFARLANE and AHMED M. BAYOUMI

Division of Nephrology, St. Michael’s Hospital, University of Toronto, Toronto, Ontario, Canada; and Department of Medicine, Inner City Health Research Unit, St. Michael’s Hospital, University of Toronto, Toronto, Ontario, Canada

Acceptance and rejection: Cost-effectiveness and the working nephrologist.

While many nephrologists have developed a sophisticated approach to appraising clinical trials, an equal comfort in critiquing cost-effectiveness literature is often lagging. Readers can wonder how new results compare to those from other cost-effectiveness trials, and whether they should accept a new intervention as cost-effective or reject it as too costly for the benefit it produces.

Critical readers should first judge whether the authors have made the correct trade-off between complexity and generalizability when selecting a study perspective, and should examine the method of linkage between costs and effectiveness. The most popular method is the incremental cost-effectiveness ratio (ICER), which has limitations that have led some authors to prefer the net monetary benefit (NMB), where confidence intervals are more easily determined and which can more readily be used in regression analyses. Interpretation of the ICER and NMB require the choice of a cost-effectiveness ceiling, representing the maximum that society would be willing to pay for an incremental health benefit, and the development of a decision rule based on this maximum. Comparing cost-effectiveness studies from different disciplines requires the use of “universal” effectiveness measures, such as the quality-adjusted life-year (QALY).

An understanding of study perspective, the relative strengths of different cost-effectiveness measures, the methods for measuring uncertainty in these estimates, and how to select and use cost-effectiveness ceiling ratios will help the critical reader to determine if a new intervention should be accepted or rejected.

Restoring and preserving health is a pricey enterprise. The costs and effectiveness of an intervention are factors that must be balanced by physicians when prescribing an intervention, and by health care planners when considering how to allocate resources between programs. Health economics can guide medical decision making through analysis of the relative costs and outcomes of the relevant options. While nephrologists are aware that dialysis is expensive (and often quoted as a benchmark when considering whether therapies are cost-effective), many are uncomfortable appraising the results of a health economics study. Readers are often left with concerns such as “What is the meaning of the estimate of cost-effectiveness? How confident are the authors in their results? How do I compare these results to those from other cost-effectiveness studies? Should this new intervention be accepted as cost-effective or rejected as too costly for the incremental health benefit it provides?” We will address these questions by focusing on concepts important to the critical reading of studies that link costs and outcomes, and interpreting these studies through the use of decision rules. These issues are of particular importance in nephrology, as dialysis is expensive, repetitive, ongoing, and statutory.

We will build on previous works introducing the concepts of cost-effectiveness analysis by familiarizing readers with the recent methods to quantify uncertainty in cost-effectiveness estimates, including the net monetary benefit framework, by highlighting cost-effectiveness issues from studies in nephrology, and by examining the development of decision rules as a critical step in the analysis process. While many cost-effectiveness analyses are performed using modeling techniques, we will focus on studies that where costs and health benefits have been directly measured. Readers desiring a greater understanding of costing methodologies, measuring effectiveness or standards in health economics are directed to existing works [1–9].

GETTING SOME PERSPECTIVE

An essential decision for an economic analysis is which perspective to use when measuring costs and outcomes. Options include the viewpoint of a patient, a hospital, a health care provider organization [such as a health maintenance organization (HMO) or a provincial Ministry of Health], or society as a whole. While analyses based on
narrow perspectives are often easier to perform (as they focus on fewer components), they may miss important costs. For example, discharging a patient early from hospital may reduce costs from the unit manager’s perspective but not for the health insurer, who must consider increased costs related to home care. A “societal” perspective would be the broadest, including potentially important costs such as an informal caregiver’s lost wages. Similarly, limited perspectives may exclude important health outcomes. For example, a program of intensive dialysis may reduce the requirement for antihypertensives, but also improve sexual function (important from the patient’s perspective) and restore social functioning (important from the societal perspective).

It is important that the perspective taken be the same for both costs and health outcomes. A study should not capture health outcomes from a societal perspective, while restricting costs to those borne by the dialysis unit. As we will see later, the chosen perspective will also affect how we judge which interventions are cost-effective and which are not.

Guidelines for conducting economic analyses generally favor using broad perspectives, particularly the societal perspective [2, 10], although this can increase study complexity, and require measurement of some costs and outcomes that are difficult to quantify. The critical reader should judge whether the authors have made the correct trade-off between complexity and generalizability when selecting a perspective.

RELATING COSTS AND OUTCOMES

It is rare for costing studies to be published without a comparison of effectiveness, as cost-minimization is rarely the primary aim of medical interventions [2]. Cost-effectiveness analyses describe the impact of an intervention on both costs and patient outcomes. The term “cost-effectiveness study” is, somewhat confusingly, used in two contexts: generically, to refer to any study that links costs and outcomes (Fig. 1) and specifically, to refer to studies that measure effectiveness with a “natural unit,” an outcome that intuitively captures the effect of an intervention. Examples of natural unit analyses include studies that related the cost of OKT3 to the number of years of graft survival ($8335/year of additional graft survival) [11], the cost of vaccination in predialysis clinics to the number of cases of hepatitis B prevented ($856/case) [12], and the cost of erythropoietin alpha to improvement in hematocrit ($370/hematocrit percent) [13].

When the difference in costs is divided by the difference in effectiveness, the result is one of the most commonly used metrics in health economics, the incremental cost-effectiveness ratio (ICER) (Fig. 2) [14], which is often plotted in one of the quadrants of a cost-effectiveness plane (Fig. 3) [15]. Consider an analysis that calculates the ICER for a new therapy relative to the current standard of care. A new therapy both more effective and less costly than conventional care (a “dominant” and therefore usually preferred therapy) is plotted in the southeast quadrant, while one more costly and less effective (and therefore usually rejected) is plotted in the northwest quadrant. Many new interventions land in the northeast quadrant (more effective but also more expensive). As we will see, these require the development of a “decision rule” before being accepted or rejected. Interventions in the southwest quadrant cost are both less costly and less effective than the standard of care. Acceptance or rejection of such an intervention will be a difficult decision, as patients must sacrifice health in order to realize the savings [16].

\[
\text{Net monetary benefit by regression} = NMB_i = \alpha + \sum_{j=1}^{p} \beta_j x_j + \delta t + \epsilon
\]

where \(\alpha\) is the regression intercept, \(p\) is the number of covariates \(x_i\), \(t\) is a dummy variable for treatment group (0 = standard treatment, 1 = investigational treatment), \(\delta\) is the regression coefficient for the treatment variable and \(\epsilon\) is the estimate of the incremental net benefit, and

\[
\text{Net monetary benefit by regression} = NMB = \lambda \times \Delta E - \Delta C
\]

where \(\lambda\) = maximum willingness to pay for a benefit.

**Fig. 1.** Categories of cost-effectiveness analyses by type of effectiveness measure. QALY is quality-adjusted life-year.

**Fig. 2.** Equations used to estimate cost-effectiveness by calculating an incremental cost-effectiveness ratio or net monetary benefit. Adapted from Hoch JS, et al: Health Econ 11:415–430, 2002.
pressing benefit in monetary terms may be intuitive, by reducing erythropoietin use [17]. While the transferrin saturation saved $5 per dollar spent on screening dialysis patients for iron deficiency with a cost-effective approach to a medical problem, or to determine allocative efficiency, which considers how best to use scarce resources. Allocative efficiency is particularly important for health policy decision makers considering decisions within the health care sector (for example, to build a new dialysis unit or fund a new medication for human immunodeficiency virus (HIV)), or between health and other sectors (for example, to build a new dialysis unit or a new high school). Analysts using cost-effectiveness studies to assess allocative efficiency must be able to compare results across studies; that is, there must be standard measures of effectiveness. How would a health care provider choose between OKT3 to prevent rejection or hepatitis B vaccination in the predialysis setting, assuming that budget constraints forced a choice between them? While both studies examined the technical efficiency of the interventions, the dissimilarity in their outcome measures prevents an easy comparison. Cost-benefit and cost-utility studies are two types of economic analyses that attempt to address this issue by incorporating “universal” effectiveness measures.

In cost-benefit analyses, the benefit of an intervention is expressed in dollars. For example, a study concluded that screening dialysis patients for iron deficiency with a transferrin saturation saved $5 per dollar spent on screening by reducing erythropoietin use [17]. While the expression of benefit in monetary terms may be intuitive, cost-benefit trials are uncommon in health care due to methodologic issues (it is difficult to assign a dollar value to some health states), moral dilemmas (some respondents say that “you can’t put a price on health”), and ethical concerns (the use of monetary valuations can bias results in favor of those with more money) [18]. For these reasons, cost-utility studies are a more popular alternative. Cost-utility studies measure effectiveness through the use of utility scores, which reflect a person’s preference for health states [19]. In health economics, utilities are usually expressed on a scale ranging from 0 (a health state equivalent to death) to 1 (equivalent to the best imaginable health). Discussion of how utility scores are measured can be found elsewhere [1, 2]. Cost-utility studies often express results using the popular outcome of quality-adjusted life-years (QALYs), which are the product of the utility score and the number of years spent in the health state, and therefore simultaneously capture both quantity and quality of life in a single measure. Studies measuring cost-per-QALY permit the direct comparison of studies from disparate medical fields.

**THE GREAT EQUALIZERS**

Cost-effectiveness can be used as a measure of technical efficiency, which deals with the selection of the most cost-effective approach to a medical problem, or to determine allocative efficiency, which considers how best to use scarce resources. Allocative efficiency is particularly important for health policy decision makers considering decisions within the health care sector (for example, to build a new dialysis unit or fund a new medication for human immunodeficiency virus (HIV)), or between health and other sectors (for example, to build a new dialysis unit or a new high school). Analysts using cost-effectiveness studies to assess allocative efficiency must be able to compare results across studies; that is, there must be standard measures of effectiveness. How would a health care provider choose between OKT3 to prevent rejection or hepatitis B vaccination in the predialysis setting, assuming that budget constraints forced a choice between them? While both studies examined the technical efficiency of the interventions, the dissimilarity in their outcome measures prevents an easy comparison. Cost-benefit and cost-utility studies are two types of economic analyses that attempt to address this issue by incorporating “universal” effectiveness measures.

In cost-benefit analyses, the benefit of an intervention is expressed in dollars. For example, a study concluded that screening dialysis patients for iron deficiency with a transferrin saturation saved $5 per dollar spent on screening by reducing erythropoietin use [17]. While the expression of benefit in monetary terms may be intuitive, cost-benefit trials are uncommon in health care due to methodologic issues (it is difficult to assign a dollar value to some health states), moral dilemmas (some respondents say that “you can’t put a price on health”), and ethical concerns (the use of monetary valuations can bias results in favor of those with more money) [18]. For these reasons, cost-utility studies are a more popular alternative. Cost-utility studies measure effectiveness through the use of utility scores, which reflect a person’s preference for health states [19]. In health economics, utilities are usually expressed on a scale ranging from 0 (a health state equivalent to death) to 1 (equivalent to the best imaginable health). Discussion of how utility scores are measured can be found elsewhere [1, 2]. Cost-utility studies often express results using the popular outcome of quality-adjusted life-years (QALYs), which are the product of the utility score and the number of years spent in the health state, and therefore simultaneously capture both quantity and quality of life in a single measure. Studies measuring cost-per-QALY permit the direct comparison of studies from disparate medical fields.

**TO ACCEPT OR REJECT: USING THE COST-EFFECTIVENESS CEILING RATIO**

If both costs and effectiveness are higher with a new intervention, the question becomes whether the increase in effectiveness is worth the extra expense, or colloquially, do we get enough bang for the buck? To answer this question we need a “decision rule,” a guiding principle (preferably prespecified) that defines the circumstances where we will accept a new intervention as cost-effective. For example, when considering antirejection therapy following renal transplantation, we might create a decision rule stating that any new drug costing less than $10,000 per year of additional graft survival is cost-effective. Decision rules can be used to serve other purposes, such as determining whether a new treatment should be funded, however, we will restrict our discussion to the use of decision rules in determining which of the available options is the most cost-effective. Developing decision rules requires us to know the maximum amount that the payer would be willing to spend for an additional benefit. If the intervention’s ICER is less than this maximum, the new intervention is deemed cost-effective; otherwise it is rejected as too costly for the benefit achieved. This maximum willingness to pay for a benefit is known as the cost-effectiveness ceiling ratio, which can be plotted as a line through the origin on the cost-effectiveness plane (Fig. 3). Interventions falling on or to the right of this line are generally considered cost-effective, while those falling to the left are not. While theoretic work justifies the existence of a maximum ceiling ratio [20], in practice it is difficult to agree on what this value should be [21]. In our previous...
example, the use of OKT3 following a kidney transplant costs an additional $8335 for each additional year of graft survival [11]. A decision rule based on the ICER would dictate that if society were willing to pay $10,000 or more for a year of graft survival, then OKT3 would be cost-effective; however, if society were willing to pay only $5000 for a year of graft survival, OKT3 is economically unattractive.

When creating a decision rule, how are we to know the maximum amount a payer would be willing to pay for a year of graft survival? What is preventing an author from choosing a ceiling ratio that leads to favorable interpretation of their conclusions [21]? Authors can reduce such concerns by referencing previous work that has suggested standards for ceiling ratios or by presenting their results using a variety of ceiling ratios.

When using the first approach, the author can take advantage of a broader study perspective. Consider erythropoietin therapy for anemia associated with renal disease. Erythropoietin costs $370 per percent increase in hematocrit [13]. However, as a society we have not defined how much (if anything) we would be willing to pay to increase hematocrit. Our determination of whether erythropoietin therapy should be accepted or rejected requires comparison of the study results to the cost-effectiveness ceiling ratio, which in this case is undefined, making interpreting of this study difficult. But what if the perspective were broadened to incorporate other outcome measures? Erythropoietin has also been shown to improve quality of life in dialysis patients, a valuable outcome for which ceiling ratios ranging from $60,000 to $160,000/QALY have been suggested [22–25]. Knowing this, we could suggest a decision rule that anemia therapy must cost less than $100,000 per QALY gained in order to be cost-effective. In a recent study, when erythropoietin was used to raise the hemoglobin from the 9.5 to 10.5 range to the 11.0 to 12.0 range, the ICER was $55,295/QALY. Raising the hemoglobin to higher values increased the ICER to more than $600,000/QALY [26]. Armed with our decision rule, we can conclude that erythropoietin is likely cost-effective when raising hemoglobin to the 11.0 to 12.0 range, but not cost-effective when a higher hemoglobin is targeted. When broad perspective outcomes such as QALYs are used, we can begin to determine which clinical scenarios are cost-effective.

Alternatively, authors can calculate the probability of cost-effectiveness across a range of ceiling ratios, allowing the astute reader to select a cost-effectiveness ceiling that they feel is appropriate. When authors present the probability of cost-effectiveness across a range of ceiling ratios, readers using different decision rules will be equally well served. The major limitation of this approach is the lack of a clear conclusion for inexperienced readers, who may not be able to suggest a cost-effectiveness ceiling for an intervention.

MEASURING UNCERTAINTY IN A COST-EFFECTIVENESS STUDY

One of the most significant drawbacks of the ICER relates to the mathematical difficulty in creating confidence intervals for a ratio [27, 28]. Accordingly, ICERs are often published with no accompanying assessment of uncertainty, an event that would be considered extraordinary in any other type of study! Authors often sidestep this concern by presenting confidence intervals for the individual differences in costs and outcome, but not for the ICER itself. Recently, several methods have been developed to estimate uncertainty for the ICER, including techniques such as “bootstrapping,” where statistical precision is estimated by repetitively generating hypothetical substudies through resampling (randomly selecting data from the original data set) with replacement (allowing each original patient to be sampled more than once). The uncertainty regarding the ICER can be demonstrated graphically by plotting each of these substudies on the cost-effectiveness plane. If a significant proportion (for example, 95%) fall on or to the right of the line representing the selected cost-effectiveness ceiling ratio, then one can be suitably confident that the new intervention is cost-effective. Another graphic approach is the construction of a cost-effectiveness acceptability curve, in which the probability that the new intervention is cost-effective is presented as a function of the ceiling ratio [29]. The curve is drawn by calculating the proportion of ICER points (generated by bootstrap or decision analysis techniques) that fall on or to the right of a series of cost-effectiveness ceiling lines. As an example of these two graphic illustrations of uncertainty in ICER estimates, we can consider recent work by our research group examining the costs and effectiveness of home nocturnal hemodialysis versus conventional in-center hemodialysis [30]. We determined that home nocturnal hemodialysis was dominant, with an ICER of $45,932 saved per QALY gained, but for the reasons mentioned above we did not provide confidence intervals for this result. To illustrate uncertainty in our results we created 2500 hypothetical study groups using a bootstrap technique, and plotted these on the cost-effectiveness plane, along with an ellipse that encompassed 95% of data points, the boundaries of which fell entirely to the right of the right of the $50,000 cost-effectiveness ceiling line (Fig. 4). We then determined the proportion of ICER points that were cost-effective at a variety of ceiling ratios, and generated an acceptability curve (Fig. 5).

While graphical techniques help illustrate the uncertainty in ICER estimates, numeric estimates of confidence intervals are also desirable. As we will see, a newer measure of cost-effectiveness known as the net monetary benefit (NMB) can provide these estimates, and correct for other shortcomings of the ICER.
would be entirely different. Finally, while uncertainty in the ICER estimate can be illustrated graphically, calculating it mathematically is difficult. Many of these deficiencies can be avoided by using an approach based on the concept of an NMB. The NMB is calculated by assigning a monetary value to the incremental benefit achieved (equal to the product of the cost-effectiveness ceiling ratio for one unit of benefit and the number of units of benefit achieved), and subtracting from this the incremental cost of achieving the benefit (Fig. 2) [32]. A positive NMB always implies that the additional value of a new therapy is more than the extra cost, and therefore it should be considered cost-effective. A negative NMB implies that an intervention should be rejected, as its value is less than the additional cost of the benefit. The NMB is well suited for use in regression analyses, and confidence intervals and tests of statistical significance can be determined using standard statistical tests. The major limitation of the NMB approach is that it requires the analyst to explicitly place a monetary value on health outcomes through the selection of a cost-effectiveness ceiling ratio. If there is controversy regarding which ceiling ratio to use, an analyst can choose to present a series of NMB values calculated using a range of cost-effectiveness ceiling ratios.

Returning to our comparison of home nocturnal and in-center hemodialysis, we calculated the NMB over a range of cost-effectiveness ceilings using a multivariate regression analysis that accounted for the independent effects of age and the presence of cardiovascular disease. In all cases, the NMB of home nocturnal hemodialysis was positive, implying that the value of the incremental benefit exceeded the incremental costs. We concluded that home nocturnal hemodialysis was cost-effective over a range of cost-effectiveness ceilings from $0 (if society were only willing to consider interventions that were cost-saving) to $100,000 per QALY gained. The ability to calculate the NMB using regression techniques helped us to address criticism that the differences in cost-effectiveness between the two study groups were related to demographic differences rather than the type of dialysis. The ability to determine the probability of a positive NMB over a wide range of cost-effectiveness ceiling ratios allowed us to satisfy readers who believe the societal cost-effective ceiling ratio is different from that which we had presented in our primary analysis.

**CONCLUSION**

When reviewing a cost-effectiveness paper, the reader must ultimately decide whether to accept a new intervention as cost-effective, or reject it as too expensive for the benefit gained. This determination is aided by consideration of study design and perspective, the method of relating costs and effectiveness, and most importantly, the comparison of study results to a reasonable ceiling ratio (Table 1). Those interested in a more detailed approach...
Table 1. Suggested steps in critiquing a cost-effectiveness analysis

Step 1: Study design and perspective
How methodologically rigorous was the study design? (in order of preference)
Meta-analysis of randomized controlled trials, randomized controlled trial, prospective cohort, retrospective cohort, case-series
What population was studied (consider age, gender, comorbid conditions, dialysis vintage, socioeconomic factors, country of study)
Was the study perspective explicitly described and sufficiently broad? (in order of preference)
Societal, health care payor [example, Ministry of Health, health maintenance organization (HMO)], regional/multicenter program, single hospital, single program (example, dialysis unit or transplant program)
Did the list of costs and benefits studied match the selected perspective?
Were the list of costs and benefits sufficiently broad to capture important differences in costs and benefits of the studied interventions?

Step 2: Relating costs and benefits
Were confidence intervals provided for the individual estimates of differences in costs and effectiveness?
Which measure was used to relate costs and benefits?
For incremental cost-effectiveness ratios (ICER)
Was an estimate of the uncertainty in the ICER provided?
For Net Monetary Benefit (NMB)
Was an estimate of the uncertainty in the NMB provided?
Was a linear regression model testing the effect of baseline and demographic variables on NMB presented?
Were sensitivity analyses provided that test the effect of critical variables on the robustness of primary results?

Step 3: Interpreting the results
Was a predefined decision rule explicitly stated?
Is this cost-effectiveness measure compared against a cost-effectiveness ceiling ratio?
Do the authors provide convincing justification for their choice of ceiling ratios?
Does the reader agree with the chosen cost-effectiveness ceiling ratio?
Was an estimate of probability of cost-effectiveness at the chosen threshold provided?
Were estimates of probability of cost-effectiveness at different ceiling ratios provided (example, a cost-effectiveness acceptability curve)?

to reviewing cost-effectiveness studies are directed to the “User’s Guide” series in the Journal of the American Medical Association [6–8, 33, 34].

The push toward evidence-based practice has resulted in many health professionals being able to critique clinical trials with confidence and sophistication. These skills form the first step in critiquing a health economics paper. A familiarity with the methodology and terminology used in cost-effectiveness analyses completes the skill set needed to review these papers and to determine whether new interventions should be accepted or rejected. As cost-effectiveness studies appear more often in mainstream medical journals, readers should develop the expertise required to appraise these works with confidence.

Reprint requests to Philip A. McFarlane, Director of Home Dialysis, St. Michael’s Hospital, 61 Queen St. E., 9th Floor, Toronto, Ontario, Canada, M5C 2T2.
E-mail: phil.mcfarlane@utoronto.ca

REFERENCES

3. Guyatt GH, Sackett DL, Cook DJ: Users’ guides to the medical literature. II. How to use an article about therapy or prevention. B. What were the results and will they help me in caring for my patients? Evidence-Based Medicine Working Group. JAMA 271:39–63, 1994


