Standardizing cost-utility analysis in neurosurgery

JARED D. AMENT, M.D., M.P.H., AND KEE D. KIM, M.D.

Department of Neurosurgery, University of California, Davis, Sacramento, California

This review seeks to introduce the concept of cost-utility analysis in neurosurgery and to highlight its essential components. It also includes a suggested approach to standardization, which would help bring more credence to this research and potentially affect management choices, reimbursement, and policy. (http://thejns.org/doi/abs/10.3171/2012.4.FOCUS1288)

Key Words • comparative effectiveness • cost-effectiveness • cost-utility • neurosurgery

Neurosurgery, like many subspecialized fields in medicine, has become increasingly technological. And, expectedly, this is associated with escalating costs. In an era of cost containment and accountability, neurosurgery resultantly finds itself in a quandary. The awestruck days of “well, it’s neurosurgery” are fleeting, as is the reliance on mild improvements in functional domains or clinical outcomes. We contend that addressing and attempting to maximize health-related QOL in neurosurgery is an important, relatively underutilized, investigative approach that may deviate the field from its current course during the next decade.

Health-related QOL may be defined as the extent to which one’s usual or expected physical, emotional, and social well-being are affected by a medical condition or treatment. Quality of life measurement tools can be dichotomously classified as either “health status” or “preference-based” instruments. Health status instruments break QOL down into several domains based on a conceptual model and provide a score in each of the domains. The instruments are typically multiple choice questionnaires asking about current symptoms and functioning, and responses are used to calculate scores. The most widely used generic health status instrument include the SF-36, SF-12 (12-Item Short-Form Health Survey), and more recently, the economic utility derivative, SF-6D (6-dimensional Short-Form Health Survey).

In contrast, preference-based QOL instruments elicit patients’ valuations for their current health state. The instruments generate a single value expressed on a 0-to-1 scale, where 0 represents the value of death and 1 represents the value of perfect health. This valuation of a health state is also known as “utility,” a concept developed by economists to indicate the strength of an individual’s preference. It is important to note that these instruments are, at best, surrogates for determining utility values and health states, and they are used primarily because of their ease of administration. The standard gamble approach has long been considered the “gold standard” for assessing utilities; however, it has also been criticized for overestimating risk aversion and being too cognitively demanding. As a result, the “time trade-off” method was designed as an alternative. Both methodologies generate utilities that can be combined with economic data to determine cost-effectiveness.

Utilities

The standard gamble method generates utilities that are fully consistent with expected utility theory and the axioms of rational individual behavior. It requires an individual to choose between 2 alternatives: a health state that is certain, such as low-back pain from degenerative disc disease, and a gamble with an outcome that is either better, often perfect health, or worse, often death. Respondents must then choose a probability that makes them inclined to take or not take the gamble. For example, if participants would take the gamble with a 90% chance...
of perfect health (and hence a 10% chance of death), but would not with a 0.89:0.11 breakdown, then 0.89 represents the utility of low-back pain due to degenerative disc disease. The concern with the standard gamble approach is that it correlates with risk behavior. In general, individuals are risk averse and therefore tend toward an outcome that is certain rather than taking the gamble. This artificially inflates utilities (compared with risk-seeking individuals) for the health states being examined. In contrast, with the time trade-off approach, an individual is presented with life-expectancy estimates (based on available data and approximations), followed by how many years he or she would be willing to trade in return for perfect health (assuming less than perfect health due to a particular health state). Thus, if expected to live another 20 years and be willing to trade 5 of them for perfect health (that is, resolution of low-back pain and return to work), the utility value for her current health state would be calculated as 1.0 – (5/20) = 0.75.

**Terminology**

The term “cost-effective” requires defining, as it can include several disparate analyses that are often confused. The first is “cost-minimization analysis.” Typically, such an analysis compares 2 interventions of comparable effectiveness to determine the less costly. In contrast, a “cost-benefit analysis” compares the costs of an intervention against the money saved as a result and is widely used in health care policy decision making. In both of these examples, however, QOL issues are ignored. The “cost-effectiveness analysis” measures the output (benefit) on a different scale, rather than in monetary terms. Output units are usually described in terms of life years gained or saved, for example, QALYs gained. Average cost-effectiveness evaluates a single intervention against its baseline option. Incremental cost-effectiveness represents the additional cost of 1 unit of outcome gained by an intervention, when compared with the next best alternative. Any cost-effective analysis that accounts for the perceived value of an intervention or health state is termed a “cost-utility analysis.”

**Current Practice**

Despite the increasing prominence of evidence-based medicine practices, a standardized rubric for what is considered cost-effective has yet to be clearly defined. A commonly cited guideline considers interventions costing less than $20,000 US dollars per QALY strongly cost-effective and interventions costing more than $100,000 per QALY as not cost-effective. The UK National Institute for Health and Clinical Excellence (NICE) publicly announced their threshold definition for cost-effective treatments to range between $40,000–$60,000 US dollars per QALY. Indeed, regulatory and insurance bodies have repeatedly recognized the inherent limitations in using dollars/QALY as a proxy for value in medicine, and therapeutics costing beyond these guidelines are frequently funded.

Preference-based QOL instruments, such as the SF-36, SF-12, EQ-5D (EuroQol-5D health survey) often used in Europe, and the HUI (Health Utilities Index) are increasingly common in the medical literature and are arguably superior to health status or functional measures, since preference-based measures incorporate individual attitudes toward functional status, pain, or disability and integrate these attitudes proportionate to their importance to each patient. Despite this advantage, a standardized approach for using preference-based QOL instruments and for conducting CUAs has received little attention in neurosurgical studies.

**Methodology**

**Model Assumptions**

To begin making cost-utility calculations, several model assumptions and considerations must be made from the onset. The notion of creating a model around sensible and realistic assumptions is fundamental to all CUAs. Unlike typical outcomes data, cost-utility data are appreciably more dynamic. Pecuniary data are subject to inflation, discounting, interest, and exchange rates. Similarly, one of the principal criticisms of preference-based QOL measurements is the substantial variance over time. Consequently, rules must be established prior to data collection and mathematical manipulation of the data. If the limitations of the model are realistic, then the results are more likely to be accurate and generalizable.

The basic construct of economic models and CUAs should include a standardized set of model assumptions and/or considerations. These include the following:

1) The investigative sample is representative of the population.
2) Uniformity of intervention. Patients included in the study must all receive the same intervention/care (without variation in procedural technique, surgical instruments, or follow-up).
3) Impact of intervention. The impact should be a known value or approximation, available in the literature, that provides the basis for an internal quality control and for the CUA (for example, it would be inappropriate to conduct a CUA on an intervention that is known to have a poorer outcome than an alternative).
4) Large effect size: the population or societal impact of the pathology or injury being addressed.
5) Average versus incremental cost-effectiveness. Determine which will be evaluated by the CUA based on the definitions outlined in Terminology.
6) Account for disutilities, risks, and complications associated with the intervention (numerous methods exist, such as a weighted cost-evaluation decision analysis tree illustrated in Fig. 1).
7) Retention rate or survival benefit. In surgical CUAs, this refers to the longevity of a device/implant or the duration for which an intervention is predicted to sustain a positive outcome.
8) Timeframe for study.
9) Timeframe for incremental utility calculation, which is similar to or different from the study timeframe, depending on the model’s power to support mathematical extrapolations.
10) Available data or need for extrapolation/interpolation. 

11) Survival data for population. The life expectancy of the population involved in the CUA establishes the reliability of the portended utility benefit (for example, if the patients in the study are only going to live 5 more years on average, then the CUA should not be modeled on QOL gains of 10 years).

12) Direct (standard gamble or time trade-off) or indirect (for example, SF-36 or EQ-5D) ascertainment of utilities. As described previously, each has its benefits and limitations.

13) Ascertainment of costs, direct and indirect—please refer to Cost Assessment.

14) Utility stability during study period. This is important to address if utilities can change throughout the study period (for example, do preferences change with time?). This may suggest that the length of the study period, especially if extrapolated, is unrealistic.

15) Cost and utility perspectives (societal, patient, provider, and payers)—please refer to Utility Assessment and Cost Assessment.

16) Univariate versus multivariate sensitivity analysis—please refer to Sensitivity Analysis.

17) Modeling (Markov and Monte Carlo). This refers to economic tools that are often used to represent stochastic processes (random processes that evolve over time) and are particularly useful when modeling chronic disease processes. 

Utility Assessment

As described, the time trade-off approach or alternative preference-based QOL metrics ought to be used pre-operatively and at predetermined postoperative encounters to measure incremental utility changes. Incremental utility change is calculated by determining for each patient the difference between preoperative and postoperative utility calculated and the mean of these differences obtained. The total gain in QALY (comparative effectiveness) can also be determined by multiplying the years of utility gain by years of benefit duration and comparing it to the preoperative utility (QOL) state.

Cost Assessment

All pertinent costs for the intervention, including direct medical costs, costs associated with complications, insurance costs, third-party payer costs, hospital costs, surgical costs, societal costs, and indirect costs (often expressed in terms of opportunity costs or disability-adjusted life years), such as follow-up care, time from work/school, time from caregiver responsibilities, and perceived value saved must be accounted for. Indirect costs are often underreported in the literature, especially in surgical fields. Medicare reimbursement rates and CPT (Current Procedure Terminology) codes have inherent limitations but are often used to represent societal impact. Although by no means the only available tool, UB-04 (Uniform Billing) forms report admission-specific charges to Medicare and are a valuable resource. 

An expected-value decision tree that accounts for probabilities of events is an essential component of this process (Fig. 1). The expected-value decision tree includes follow-up visits and possible complications requiring reoperation, all of which are weighted based on normal clinical practice and complication rates observed in the study sample. Total costs ought to be calculated by the net present value method, accounting for initial costs while...

Fig. 1. Example of an expected-value decision tree from the literature. Discounted costs are italicized. *If conducted, are single initial costs at time of KPro and are therefore not discounted. **BCL fitting and replacement is $300 per year after initial year. BCL = bandage contact lens; ECCE = extracapsular cataract extraction; IOL = intraocular lens; KPro = keratoprosthesis; PK = penetrating keratoplasty; PPV = pars plana vitrectomy; YAG = yttrium aluminum garnet. Used with permission from Dr. Ament.
appropriately discounting future costs. Discounted future costs include inevitable costs and the average of weighted possible future costs. This requires sufficient sample size to ensure generalizability.

**Discounting**

There is a consensus that utilities and costs should be discounted in health care economics analyses. The 3% conventional discount rate is appropriate initially, but this value should be modified in a sensitivity analysis.\(^5\)

**Calculation of QALY**

Quality-adjusted life years are to be calculated using the following formula:\(^5,20\)

\[ \sum_{x=1}^{n} t^x u \frac{1}{(1 + d_y)^x} \]

where \(n\) is the study period, \(t\) is the yearly retention rate of system or intervention being studied, \(u\) is the average incremental utility, and \(d_y\) is the discounting rate for QALYs (3%).

**Calculation of Cost**

As described, calculating costs is most commonly achieved by the net present value method. An example of the equation used to determine the total discounted cost associated with the system or intervention being studied is as follows:

\[ \sum_{x=1}^{n} \frac{Sa}{(1 + d_y)^x} + \sum_{x=1}^{n} \frac{Sb}{(1 + d_y)^x} + \sum_{x=1}^{n} \frac{Sc}{(1 + d_y)^x} + \sum_{x=1}^{n} \frac{Sd}{(1 + d_y)^x} + \frac{Se}{(1 + d_y)^x} \]

An initial cost, \(\theta\), is normally incurred at or immediately prior to or after the time of surgery and is not discounted. Costs paid over the initial year alone will be discounted accordingly, as are costs paid throughout the entire study period. In the equation, \(x\) represents the year of follow-up and \(d_y\) is the discounted rate. Future value analysis is a sophisticated economic method that starts with the net present value and uses integration and calculus to account for inflationary pressures and continuous ongoing expenses. This type of analysis, however, is beyond the scope of this review.

**Sensitivity Analysis**

The model ought to undergo a univariate, and when feasible, a multivariate sensitivity analysis. For any CUA, this entails manipulating the appropriate parameters, such as utility value, retention rate, discounting rate for QALYs, and discounting rate for costs. Markov models, Monte Carlo simulations, bootstrapping, and other computer-dependent methods exist that enable simultaneous multivariate manipulation to address all probable scenarios (exponential permutations).

**Discussion**

**Cost-Utility Analysis in Practice**

The cost-utility value of various medical interventions is listed in Table 1. The wide range in this sample illustrates the numerous variables that may influence the final cost-effectiveness of an intervention. For example, although the mean cost of an MRI study for equivocal neurological symptoms is less than $2000 per scan, the lack of utility gained by the test in returned health years, that is, QALYs, results in it being markedly cost-ineffective at $142,206/QALY. Similarly, although cephalexin may be an inexpensive drug, the probability of an infection after dental work among patients with prosthetic joints is so low that the cost of providing the antibiotic on a population level is very high.\(^{14}\) These examples illustrate how we ought to question our health care practices, especially with respect to conferred value and QALYs. Despite lacking standardization, neurosurgery has started to demonstrate productivity in the areas of health-related QOL and CUA research.

Spine surgery has been a principal focus for CUA among neurosurgeons in part due to the existence of standardized and validated outcome assessment tools, the expense associated with spine surgery, the morbidity and disability associated with spine disease/injury, and the availability of nonsurgical management options. Tosteson et al.\(^{23}\) published a well-structured analysis using the SPORT (Spine Patient Outcomes Research Trial) data. These authors showed that surgery for lumbar disc herniation was more costly than nonoperative treatment ($14,137) but that the treatment, when evaluated over 2 years, was moderately cost-effective ($69,403/QALY using general adult surgery costs and $34,355/QALY using Medicare population surgery costs).\(^{23}\) This study illustrates how the economic value of an intervention can vary considerably according to the method used for assigning costs. It also highlights how robust modeling and thorough methodological planning can result in meaningful conclusions from CUA.\(^{23}\) But, to expand on this foundation, mitigate the variability between CUA, and ensure a greater understanding and appreciation for CUA in neurosurgery, standardization is essential.

**TABLE 1: Cost-utility of medical interventions in the US, adjusted to 2011 US dollars***

<table>
<thead>
<tr>
<th>Authors &amp; Year</th>
<th>Intervention</th>
<th>Cost in $/QALY</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ament et al., 2010</td>
<td>keratoprosthesis surgery for severe corneal injury</td>
<td>17,034</td>
</tr>
<tr>
<td>Mushlin et al., 1997</td>
<td>chemoprophylaxis after occupational exposure to HIV</td>
<td>51,752</td>
</tr>
<tr>
<td>Busbee et al., 2002</td>
<td>primary pediatric heart transplant</td>
<td>55,321</td>
</tr>
<tr>
<td>Revicki &amp; Kaplan, 1993</td>
<td>lumbar discectomy for intervertebral disc herniation</td>
<td>69,403</td>
</tr>
<tr>
<td>McCabe et al., 2008</td>
<td>MRI for equivocal neurological symptoms</td>
<td>142,206</td>
</tr>
<tr>
<td>Griffin et al., 2007</td>
<td>1 day of chemoprophylaxis prior to receiving dental work for patients with prosthetic joints</td>
<td>735,284</td>
</tr>
</tbody>
</table>

Standardizing cost-utility analysis in neurosurgery

**Standardization**

A paradigm shift is underway in health care, but despite the increasing prominence of CUAs, a standardized rubric for collecting, analyzing, and presenting health-related QOL data in neurosurgery has not been established. When this review is deconstructed, it attempts to put forward a pragmatic approach to standardizing CUAs in neurosurgery. The salient steps and considerations are as follows:

1. Choose a question that warrants a CUA.
   a) Consider patient, provider, and societal perspectives.
   b) Know the literature and first demonstrate positive outcomes.

2. Adhere to the methodological recommendations (see Methodology).
   a) Ensure that model assumptions are stable.
   b) Choose between average and incremental cost-effectiveness.

3. Use the time trade-off approach whenever feasible.
   a) Preference-based QOL instruments are acceptable, but ensure that they are being used appropriately (with correct population weights).
   b) Handle the data equally.
   a) Costs are as equally important as utilities.
   b) Both require discounting.
   c) Indirect costs should be incorporated in every CUA.

4. Present all the appropriate perspectives.
   a) The source of the data will limit what perspectives can be presented; this will also limit the usability and generalizability of the results.

5. Conduct a robust sensitivity analysis.
   a) This tests the stability of your model.
   b) Uni- and multivariate analyses.

6. Discuss the results in context.
   a) Results should always be discussed in the context of impact and policy implications.
   b) Compare (in terms of $/QALY) with other similar neurosurgical interventions and modalities.

**Conclusions**

Standardizing CUAs and health-related QOL research in neurosurgery will result in reliable and generalizable information. This should help reduce uncertainty about which treatments are not only best but also cost-effective. What is learned can then be readily disseminated and inevitably direct care decisions. This may also significantly improve patient outcomes by imposing a close examination of standard clinical practices for efficacy of treatment. With sensitivity about resource allocation and cost containment at its pinnacle, expensive, highly technological, and highly specialized fields like neurosurgery ought to spearhead efforts toward cost-effectiveness and health care improvement. A standardized approach to the research that will allow us to make informed and consistent decisions is the initial step.

**Disclosure**

The authors report no conflict of interest concerning the materials or methods used in this study or the findings specified in this paper.

**Author contributions to the study and manuscript preparation**

Include the following: Conception and design: Ament. Acquisition of data: Ament. Analysis and interpretation of data: Kim. Drafting the article: both authors. Critically revising the article: both authors. Reviewed submitted version of manuscript: both authors. Approved the final version of the manuscript on behalf of all authors: Ament. Administrative/technical/material support: Kim. Study supervision: Kim.

**References**