Health-related Quality of Life in Emphysema

Robert M. Kaplan1 and Andrew L. Ries2

1Department of Health Services, University of California, Los Angeles, Los Angeles, California; and 2Departments of Medicine and Family and Preventive Medicine University of California, San Diego, San Diego, California

Patients with emphysema may experience reduced health-related quality of life (HRQOL). HRQOL measures have evolved from two different measurement traditions: psychometric theory and decision theory. Psychometric methods typically create a profile of outcomes, whereas decision theory methods offer a summary score on a continuum ranging from 0.0 (for death or worst possible health) to 1.0 (for best possible health). Decision theory methods are better suited for cost-effectiveness studies. Generic HRQOL measures can be applied to any disease population, whereas disease-targeted measures are tailored to a specific clinical condition. Disease-targeted measures are typically more sensitive to clinical change, but cannot offer a comparison basis for different clinical conditions. This article reviews the measurement of HRQOL in patients with emphysema. The National Emphysema Treatment Trial (NETT) offers an example of the application of both generic and disease-targeted, as well as profile and decision theory, methods. The NETT illustrates how HRQOL measures can be used to assess outcomes and estimate cost-effectiveness in a major clinical trial.

Keywords: health-related quality of life; emphysema; outcomes assessment; quality-of-life measurement

Patients with emphysema experience significant limitations in daily activities and reduced quality of life. Evidence supports the value of medical management and pulmonary rehabilitation for emphysema. Despite important advances in medical management, the promise of medical therapy is limited. Lung volume reduction surgery (LVRS) offers the potential to enhance quality of life for selected patients who have completed pulmonary rehabilitation.

The goals of surgical and medical treatments for emphysema are to extend life expectancy and to improve quality of life. Medicines have side effects and surgery is associated with operative and perioperative risks. Recovery from surgery necessarily reduces quality of life and functioning in the short term. These effects are important, but often are difficult to quantify. Health-related quality of life (HRQOL) measures attempt to characterize the negative and positive effects of a disease as well as its effect on quality of life.

HRQOL measures have been developed over the past 30 years, and have commonly been used to study emphysema. A PubMed search cross-referencing chronic obstructive pulmonary disease (COPD) with quality of life (in June 2007) produced 1,704 references and the intersection of emphysema and quality of life yielded 313 references. The American Thoracic Society maintains a website that summarizes quality-of-life measures that can be used in outcomes research for lung disease (see www.atsqol.org). The site lists measures by disease and offers references on their use.

The purpose of this article is to review HRQOL measurement in emphysema studies with emphasis on measures chosen for the National Emphysema Treatment Trial (NETT). More comprehensive reviews of HRQOL in chronic lung disease are available elsewhere (1).

DISTINCTIONS BETWEEN MEASURES OF HRQOL

Measures of HRQOL are classified as either disease targeted or generic. Disease-targeted measures are used for patients with a particular diagnosis, whereas generic measures can be used in any population. There are two major approaches to assessment of HRQOL: psychometric and decision theory.

The psychometric approach provides a profile of measures, each summarizing a different dimension of HRQOL. The best-known example of the psychometric tradition is the Medical Outcomes Study 36-item Short Form (SF-36) (2), a generic measure of HRQOL.

The decision theory approach weights the different dimensions of HRQOL to provide a single expression of quality of life. Supporters of this approach argue that psychometric methods fail to consider that different health problems cause different levels of concern. A minor itch and coughing blood are both symptoms, but they are not equal in importance. Furthermore, a treatment may improve some aspects of quality of life while causing others to deteriorate. For example, a medication might reduce coughing, but increase skin problems or reduce energy. When components of outcome change in different directions, an overall subjective evaluation is required to integrate the components and offer a summary of whether the patient is better or worse off. The decision theory approach attempts to provide an overall measure of HRQOL that integrates subjective function states, preferences for these states, morbidity, and mortality.

Measures can also be distinguished by their possible uses. Most measures can be used to characterize populations and to study clinical changes. However, only generic decision theory–based measures can be used to evaluate cost-effectiveness.

DISEASE-TARGETED METHODS OF HRQOL

Commonly used measures of HRQOL in patients with lung diseases are described briefly below and summarized in Table 1.

The Chronic Respiratory Questionnaire

The Chronic Respiratory Questionnaire (CRO) is a 20-item interviewer-administered instrument that evaluates four domains: dyspnea, fatigue, emotional function, and mastery. For the dyspnea measure, patients rate breathlessness on the five most important activities of daily living (ADLs) chosen from a standard list. Therefore, the measure of dyspnea is unique to each individual and cannot be compared among subjects. Changes on
TABLE 1. DISEASE-TARGETED QUALITY-OF-LIFE MEASURES COMMONLY USED TO EVALUATE OUTCOMES IN ADULTS WITH CHRONIC LUNG DISEASES

<table>
<thead>
<tr>
<th>Measure</th>
<th>Type</th>
<th>Uses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chronic Respiratory Questionnaire (CRQ)</td>
<td>Profile</td>
<td>Descriptive studies, assess clinical change</td>
</tr>
<tr>
<td>St. George’s Respiratory Questionnaire (SGRQ)</td>
<td>Profile</td>
<td>Descriptive studies, assess clinical change</td>
</tr>
<tr>
<td>University of California, San Diego, Shortness of Breath Questionnaire (SOBQ)</td>
<td>Profile/dyspnea-specific</td>
<td>Descriptive studies, assess clinical change</td>
</tr>
</tbody>
</table>

Data from Reference 1.

Note: More detailed descriptions of the measures and validity studies are summarized in Reference 11.

the questionnaire have been shown to correlate with changes in spirometric values, exercise performance, and subjective ratings of improvement by both the patients and physicians (3). The questionnaire is sensitive to a variety of interventions, including bronchodilator treatment (4–6). The original CRQ required a trained interviewer. More recently, a validated self-administered form has become available (7).

St. George’s Respiratory Questionnaire

The St. George’s Respiratory Questionnaire (SGRQ) is a self-administered 50-item instrument with three separate scales (symptoms, activity, and impact on daily life). A total score can also be calculated. The individual SGRQ questions are weighted differentially to produce an overall score. The questionnaire has been evaluated for reliability and validity in several chronic lung diseases of varying severity, particularly COPD and asthma (8). We offer more on the sensitivity of the SGRQ in a later section.

DYSPNEA

Despite the advantages of disease-specific and generic quality-of-life measures, pulmonologists have been concerned that these approaches may miss some of the most important patient outcomes. For example, several studies have observed low correlations between measures of generic quality of life and lung function. Dyspnea is the one common self-reported patient outcome that may have a profound effect on HRQOL (9). Several studies show that the correlation between generic quality of life and dyspnea is substantial. Schrier and colleagues found no correlation between lung function tests and scores on a generic quality-of-life measure known as the Sickness Impact Profile (SIP) (10). However, they did observe substantial correlations between symptoms of wheezing and dyspnea and SIP scores. Yet, the general measures miss many of the subtle characteristics or subtle aspects of the clinically important symptoms. These findings suggest that measures of shortness of breath or dyspnea may be of central importance for evaluating outcomes in emphysema.

We believe that HRQOL assessment in studies of patients with emphysema should be complemented by measures of shortness of breath. However, measurement of dyspnea presents several methodologic and technical challenges. A number of instruments are available to assess dyspnea, including structured interviews and self-report questionnaires that evaluate a patient’s historical recall of breathlessness. These measures compare shortness of breath associated with daily activities or exercise to symptoms occurring at that specific time when the assessment is made (Table 2). A review of dyspnea measures by Eakin and colleagues offers more details about the measures and their properties (11).

The University of California, San Diego, Shortness of Breath Questionnaire (UCSD-SOBQ) exemplifies a validated dyspnea measure. The SOBQ is self-administered and asks subjects to rate their breathlessness for 21 various daily activities (plus 3 overall items) on a 6-point scale from none at all (0) to severe (4), to maximal or unable to do because of breathlessness (5) (12). For activities that they do not typically perform, respondents estimate their breathlessness for that activity. The 21 ADLs are grouped into four categories of ADLs: rest and light ADLs (factor 1), 8 questions; moderate ADLs (factor 2), 5 questions; walking (factor 3), 4 questions; and strenuous ADLs (factor 4), 4

TABLE 2. COMMONLY USED DYSPNEA MEASURES

<table>
<thead>
<tr>
<th>Measure</th>
<th>Description</th>
<th>Validity Evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>British Medical Research Council (MRC)</td>
<td>Five yes/no questions about shortness of breath. Requires recall. Higher score indicates more shortness of breath.</td>
<td>$r = -0.53$ with other dyspnea measures, $r = -0.42$ with FEV$_{1.0}$</td>
</tr>
<tr>
<td>American Thoracic Society Respiratory Questionnaire</td>
<td>Five yes/no items. Similar to MRC, but asks about magnitude of effort and impairment. Requires recall. Higher score indicates more impairment.</td>
<td>$r = 0.50$ with MRC questionnaire</td>
</tr>
<tr>
<td>Indexes (BDI-TDI) (57)</td>
<td>Single ratings in three categories. Ratings of transition require recall of baseline dyspnea. Higher score indicates more dyspnea.</td>
<td>$r = 0.53$ with other dyspnea indexes, $r = 0.60$ with six-minute-walk distance</td>
</tr>
<tr>
<td>Modified Dyspnea Index (58)</td>
<td>Ratings for impairment at work and impairment at home. Similar to BDI. Requires recall. Higher score indicates more impairment.</td>
<td>$r = 0.71$ with FEV$_{1.0}$</td>
</tr>
<tr>
<td>Oxygen Cost Diagram (OCD) (59)</td>
<td>Ratings of breathlessness with different activities on 10-cm line associated with activities of daily living. Requires recall. Higher score indicates more breathlessness.</td>
<td>$r = 0.68$ with six-minute-walk distance, $r = -0.053$ to $-0.79$ with other dyspnea measures</td>
</tr>
<tr>
<td>Visual Analogue Scale (VAS) (60)</td>
<td>10-cm line to rate perceived symptom associated with exercise or physical stimulus. This measure does not require recall—ratings are taken during activity. Higher score indicates greater severity of symptom.</td>
<td>$r = -0.85$ with peak expiratory flow rate</td>
</tr>
<tr>
<td>Borg scale (61)</td>
<td>Rating scale of 0 to 10 for rating intensity of breathlessness on 1–10 log scale. This measure does not require recall—ratings are taken during activity. Higher score indicates greater breathlessness.</td>
<td>$r$ with FEV$_{1.0} = 0.88$</td>
</tr>
<tr>
<td>University of California, San Diego, Shortness of Breath Questionnaire (UCSD-SOBQ) (12)</td>
<td>Self-ratings of breathlessness for 21 various daily activities (plus 3 overall items) on a 6-point scale from none at all (0) to severe (4) to maximal or unable to do because of breathlessness (5).</td>
<td>$r = -0.40$ with QWB, $-0.68$ with six-minute-walk distance, $-0.67$ with DLCO % predicted</td>
</tr>
</tbody>
</table>

Data from Reference 1.
questions. The score on each of the 24 items is summed to produce a total score (range of 0 to 120). To represent a range of disease-specific outcomes, the NETT selected the SGRQ and the UCSD-SOBQ.

**GENERIC METHODS OF HRQOL**

Generic HRQOL measures can evaluate outcomes for any population (as compared with disease-targeted measures that measure outcomes for people with a particular health condition). Some of the common generic approaches are summarized in Table 3. Readers interested in more detail should consult Shumaker and Berzon (13) or Walker and Rosser (14) and McDowell and Newell (15).

**SF-36**

Perhaps the most commonly used generic HRQOL measure in the world today is the Medical Outcomes Study SF-36. The SF-36 uses the psychometric or profile approach and grew out of work by the Rand Corporation and the Medical Outcomes Study (2). The SF-36 includes eight health concepts: physical functioning, role-physical, bodily pain, general health perceptions, vitality, social functioning, role-emotional, and mental health (16). The SF-36 is either administered by a trained interviewer or self-administered. The SF-36 can be machine scored and has been evaluated in large-population studies. The reliability, validity, and sensitivity of the SF-36 are well documented (17–19). Despite its many advantages, the SF-36 also presents some disadvantages. For example, it does not have age-specific questions, and one cannot clearly determine whether it is equally appropriate at each level of the age continuum. The items for older retired individuals are the same as those for children (17).

Some approaches to the measurement of HRQOL combine measures of morbidity and mortality to express health outcomes in units analogous to years of life. The years-of-life measure can be adjusted for diminished quality of life associated with diseases or disability (20). We consider three utility-based approaches.

**Quality of Well-Being Scale**

The Quality of Well-Being scale (QWB) combines preference-weighted values for symptoms and functioning. The preference weights were obtained by ratings of 856 people from the general population. These judges rated the desirability of health concepts (as compared with disease-targeted measures that measure outcomes for people with a particular health condition). Some of the common generic approaches are summarized in Table 3. Readers interested in more detail should consult Shumaker and Berzon (13) or Walker and Rosser (14) and McDowell and Newell (15).

**EuroQol Group 5 Dimension**

The approach most commonly used in the European community is the EuroQol Group 5 Dimension (EQ-5D). This method has been advanced by a collaborative group from Western Europe known as the EuroQol group (44). The original version of the EuroQol had 14 health states in six different domains. In addition, respondents placed their health on a continuum ranging from death (0.0) to perfect health (1.0). The method was validated in postal surveys in England, Sweden, and the Netherlands. More recent versions of the EuroQol, known as the EQ-5D, are now in use in a substantial number of clinical and population studies (40, 41). Although the EQ-5D is easy to use and comprehensive, there have been some concerns about ceiling effects. Substantial numbers of people obtain the highest possible score.

**Health Utilities Index**

The Health Utilities Index (HUI) is a family of health status and preference-based HRQOL measures suitable for use in clinical and population studies (42, 43). Each member of the family includes a health status classification system, a preference-based multiattribute utility function, data collection questionnaires, and algorithms for deriving HUI variables from questionnaire responses. The most commonly used versions are the HUI Mark 2 (HUI2) and the HUI Mark 3 (HUI3). HUI3 consists of eight attributes: vision, hearing, speech, ambulation, dexterity, emotion, cognition, and pain (43). There are five to six levels per attribute. HUI3 focuses on capacity rather than performance. Multiplicative multiattribute utility functions based on community preferences have been estimated for HUI2 (45) and HUI3 (42). Evidence on responsiveness and construct validity is provided by Torrance and coworkers (45). Evidence of

**TABLE 3. COMMONLY USED GENERIC MEASURES OF QUALITY OF LIFE**

<table>
<thead>
<tr>
<th>Method</th>
<th>Description</th>
<th>Advantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Medical Outcomes Short Form 36 (SF-36)</td>
<td>Includes eight health concepts: physical functioning, role-physical, bodily pain, general health perceptions, vitality, social functioning, role-emotional, and mental health (16).</td>
<td>The reliability and validity of the SF-36 are well documented (17–19).</td>
</tr>
<tr>
<td>EuroQol Group 5 Dimension (EQ-5D)</td>
<td>Includes five dimensions, each with three levels; it has been validated in postal surveys in England, Sweden, and the Netherlands</td>
<td>EQ-5D is now in use in a substantial number of clinical and population studies (40, 41).</td>
</tr>
<tr>
<td>Health Utilities Index (HUI2)</td>
<td>Consists of seven dimensions of health status: sensation, mobility, emotion, cognition, self-care, pain, and fertility (43). There are five to six levels per dimension.</td>
<td>Widely used in Canada and Europe. Excellent record of reliability and validity. Well-grounded in economic theory.</td>
</tr>
<tr>
<td>Self-administered Quality of Well-Being scale (QWB-SA)</td>
<td>Combines preference-weighted values for symptoms and functioning.</td>
<td>Has been used in numerous clinical trials and population studies (21–35). Can be used to estimate utility data for cost-effectiveness analysis using existing SF-36 data.</td>
</tr>
<tr>
<td>Fryback Index (48)</td>
<td>Estimates QWB scores based on SF-36 responses.</td>
<td></td>
</tr>
</tbody>
</table>
construct validity in the 1990 Ontario Health Survey is provided in Grootendorst (46). HUI3 described the burden of morbidity for both stroke and arthritis. Problems were found in the attributes that had been expected to be affected.

QUALITY-ADJUSTED YEARS OF LIFE

The QWB, EQ-5D, and the HUI are important because they can be used to obtain quality-adjusted life-years (QALYs). QALYs combine measures of quality of life and years of life to adjust the years-of-life measure for the diminished quality of life associated with diseases or disabilities (20). Modern measures of health outcome consider future as well as current health status. Lung cancer, for example, may have very little impact on current functioning but may have a substantial impact on life expectancy and functioning in the future. Today, a person with a malignant tumor in a lung may be functioning very much like a person with a chest cold. However, the patients with cancer is more likely to remain dysfunctional or to die in the near future. Comprehensive expressions of health status need to incorporate estimates of future outcomes as well as measure current status (47). For example, surgical intervention for emphysema may have short-term consequences that will reduce HRQOL. However, both HRQOL and life expectancy gains may result from the surgery. QALY methods can express this comprehensive effect of the treatment.

Most of the methods for obtaining QALYs are similar (35). The most commonly used methods are the EQ-5D, the HUI, and the QWB. Methods for translating SF-36 scores into the QWB scores (so that the SF-36 scores can be used to calculate QALYs) have emerged (48). After considerable discussion, the NETT selected the QWB.

INTEGRATING COST WITH OUTCOME DATA

Although treatment programs provide health benefits, they also have costs. Not all health care interventions equally return benefit for the expended dollar. Resources are limited, and good policy requires that they be used wisely. Cost-effectiveness studies have gained in popularity because health care costs have grown so rapidly in recent years. Objective cost studies might guide policymakers toward an optimal and equitable distribution of scarce resources. Methodologies for estimating costs have now become standardized (49). From an administrative perspective, cost estimates include all treatment costs and any costs associated with treating side effects. Typically, economic discounting is applied to adjust for using current assets to achieve a future benefit. From a social perspective, costs are broader and may include indirect costs of family members missing work to provide care. Comparing treatment programs for populations with a given medical condition, cost-effectiveness is measured as the change in costs of care for the new intervention compared with the existing therapy or program, relative to the change in health measured in a standardized unit such as the QALY. The difference in costs over the difference in effectiveness is the incremental cost-effectiveness and is usually expressed as the cost/QALY. Because the objective of all programs is to produce QALYs, the cost–QALY ratio can be used to show the relative health benefits from investing in different programs (35). Several different governments have proposed allocating resources based on systematic data (37). For example, the Australian government now requires evidence of effectiveness, as do a variety of European governments. In Ontario, Canada, QALYs have been considered as a basis for formulary decisions (50). Similar proposals have been considered in the United Kingdom (51).

HRQOL ASSESSMENT IN THE NETT

The NETT was one of the first randomized clinical trials of emphysema treatment to prospectively integrate quality-of-life assessment. The NETT compared LVRS plus maximal medical therapy with medical therapy alone. Four quality-of-life measures were administered at baseline and again at 6, 12, 24, 36, 48, and 60 months after treatment assignment. The trial used two generic HRQOL measures (the QWB-SA and the SF-36) and two disease-targeted measures (the SGRQ and the UCSD-SOBQ).

HRQOL at Baseline

Pairwise correlations between HRQOL measures obtained at baseline (SOBQ total score, QWB-SA total score, SGRQ total score, SF-36 physical component score [PCS] and mental component score [MCS]) were all statistically significant and ranged from 0.31 to 0.70 except for the PCS with the MCS; the PCS and MCS scores were originally derived from factor analysis and are expected to be uncorrelated (52). On average, the NETT participants had lower scores on the QWB-SA compared with a comparable normative population. With the exception of the SF-36 scores for bodily pain and MCS, all other SF-36 measures show NETT participants to be well below the population norms. For the QWB-SA and the SF-36 subscales (other than the PCS and MCS), the normative sample was older adults selected from the general population in Beaver Dam, Wisconsin (mean age, 64.1 yr) (53). Because the Beaver Dam sample did not include PCS and MCS scores, SF-36 general population norms for the U.S. population in the age category of 55–64 years (54) were used to assess the PCS and MCS scores of the NETT participants. These summary measures suggest that the NETT participants were quite ill. All of the HRQOL measures were significantly associated with the number of hospital days in the 3 months before NETT enrollment (all P values < 0.001) (52).

HRQOL Changes after Rehabilitation

The NETT evaluated HRQOL before and after comprehensive pulmonary rehabilitation. There were significant improvements on all HRQOL measures after rehabilitation. The correlations between the changes in HRQOL measures and the change in the six-minute-walk distance between baseline and the completion of rehabilitation were all modest, but statistically significant. This suggests that those who improved in exercise also improved on HRQOL measures (52). The pre- and postrehabilitation means show NETT participants to be well below the population norms. For the QWB-SA and the SF-36 subscales (other than the PCS and MCS), the normative sample was older adults selected from the general population in Beaver Dam, Wisconsin (mean age, 64.1 yr) (53). Because the Beaver Dam sample did not include PCS and MCS scores, SF-36 general population norms for the U.S. population in the age category of 55–64 years (54) were used to assess the PCS and MCS scores of the NETT participants. These summary measures suggest that the NETT participants were quite ill. All of the HRQOL measures were significantly associated with the number of hospital days in the 3 months before NETT enrollment (all P values < 0.001) (52).

<table>
<thead>
<tr>
<th>Variable</th>
<th>Pre</th>
<th>Post</th>
<th>P Value</th>
<th>Effect Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>SF-36 Physical Health (PCS)</td>
<td>28.3</td>
<td>29.7</td>
<td>&lt;0.001</td>
<td>0.19</td>
</tr>
<tr>
<td>SF-36 Mental Health (MCS)</td>
<td>53.2</td>
<td>55.3</td>
<td>&lt;0.001</td>
<td>0.21</td>
</tr>
<tr>
<td>SGRQ</td>
<td>56.5</td>
<td>53.1</td>
<td>&lt;0.001</td>
<td>0.35</td>
</tr>
<tr>
<td>UCSD-SOBQ</td>
<td>66.7</td>
<td>62.5</td>
<td>&lt;0.001</td>
<td>0.24</td>
</tr>
<tr>
<td>QWB-SA</td>
<td>0.537</td>
<td>0.571</td>
<td>&lt;0.001</td>
<td>0.29</td>
</tr>
</tbody>
</table>

**Definition of abbreviations:** MCS = mental component score; PCS = physical component score; QWB-SA = self-administered Quality of Well-Being scale; SF-36 = Medical Outcomes Study Short Form 36; SGRQ = St. George’s Respiratory Questionnaire; UCSD-SOBQ = University of California, San Diego, Shortness of Breath Questionnaire.

Data from Reference 52.

* n = 1,209–1,216 pairs.
are a variety of indexes of effect size, but most consider the difference between treatment and control means divided by the standard deviation of the pooled groups. The effect sizes were in a comparable range for generic and disease-targeted measures. The NETT results are consistent with a variety of studies, summarized in a recent Cochrane review, that have documented HRQOL benefits resulting from pulmonary rehabilitation (55).

HRQOL after Surgery
In May 2001, recruitment into the NETT trial was halted for patients with an FEV₁ < 20% predicted and either homogeneous emphysema on computed tomography scan or DLCO < 20% predicted. A total of 1,078 patients who did not have high-risk characteristics completed the trial (56). Patients randomly assigned to LVRS obtained lower (more favorable) scores on the UCSD-SOBQ and the total score on the SGRQ. For these measures, groups began to separate by the 6-month follow-up and the differences remained significant through 60 months. Similar trends were observed for the generic QWB-SA and the physical component of the SF-36.

Overall, the NETT trial suggests that LVRS leads to improvements in HRQOL for non–high-risk patients with emphysema. The benefits may be stronger for certain subgroups of patients. Patients randomly assigned to LVRS achieved better outcomes on the SGRQ, the UCSD-SOBQ, and the QWB-SA. These findings suggest that LVRS may result in enhancements in everyday functioning, particularly in activities associated with shortness of breath.

Cost-effectiveness Analysis
The outcomes on the QWB-SA were combined with mortality outcomes to express the benefits of LVRS in terms of QALYs. These measures are the central metric for cost-effectiveness analysis (see Ramsey and coworkers, pages 406–411, in this symposium [62]). The NETT suggested that, in comparison to maximal medical therapy, LVRS produces a QALY for about $190,000, considering a 3-year time horizon. If the time horizon is expanded to 10 years, the cost–QALY was estimated to be $53,000. The cost–QALY was even more attractive for some subgroups.

CONCLUSIONS
Quality-of-life measures evolve from two different measurement traditions: psychometric theory and decision theory. Psychometric methods, such as the SF-36, typically create a profile of outcomes, whereas decision theory methods attempt to portray an integrative summary judgment of health. Decision theory methods are better suited for cost-effectiveness studies.

A review of current outcomes research for emphysema shows that quality-of-life measures are now commonly used. There is substantial evidence for the validity of these measures for patients with chronic lung diseases. The NETT showed that, in comparison to maximal medical therapy alone, non–high-risk patients undergoing maximal medical therapy plus LVRS experienced improved HRQOL. Data from the prerandomization phase of the NETT indicate that patients with emphysema who complete rehabilitation experience significant improvements on several indexes of life quality. The NETT also illustrated how quality-of-life data can be used to assess outcomes in major clinical trials and how they can be applied in cost-effectiveness analysis (see Ramsey and coworkers, pages 406–411, in this symposium [62]). Analyses from NETT suggest that the cost-effectiveness of LVRS is justifiable in comparison to other medical and surgical treatments and can be as low as $21,000 per QALY for patients with upper-lobe emphysema and low exercise capacity if a 10-year time horizon is used. We urge that quality-of-life measurement become a standard part of outcomes assessment in emphysema clinical research.

Conflict of Interest Statement: R.M.K. does not have a financial relationship with a commercial entity that has an interest in the subject of this manuscript. A.L.R. received $50,000 from Boehringer Ingelheim as research grants for participating in a multicenter clinical trial.

References
51. Williams A. Cochrane Lecture. All cost effective treatments should be free, or, how Archie Cochrane changed my life! J Epidemiol Community Health 1997;51:116–120.