The assessment of health status among patients with COPD

J.R. Curtis*, D.L. Patrick[#]

The assessment of health status among patients with COPD. J.R. Curtis, D.L. Patrick. ©ERS Journals Ltd 2003.
ABSTRACT: Survival and physiological measures alone do not represent the full experiences of patients with chronic obstructive pulmonary disease. Reducing the personal and social burden of disease by improving patients' symptoms, functional status and quality of life are also important goals.

There has been a substantial increase in the use of newly developed tools that measure health status and it is important for clinicians and researchers to understand these instruments' strengths and weaknesses in providing insight into a patient's condition and experience.

Relying only on mortality and physiological outcomes could blind a clinician to significant benefits patients may receive from a treatment. A growing body of research utilises end-points assessed directly by patients whose self-reported health status includes health-related quality of life and functional status.

This article reviews major concepts and methods in health-status assessments for patients with chronic obstructive pulmonary disease, which will have an important role in assessing the efficacy and effectiveness of new treatments.
Eur Respir J 2003; 21: Suppl. 41, 36s–45s.

*Division of Pulmonary and Critical Care Medicine, Dept of Medicine, and [#]the Social and Behavioral Sciences Dept of Health Services, University of Washington, Seattle, WA, USA.

Correspondence: J. Randall Curtis, Division of Pulmonary and Critical Care Medicine, Harborview Medical Center, Box 359762, 325 Ninth Avenue, Seattle, WA 98104-2499, USA.
Fax: 1 2067318584
E-mail: jrc@u.washington.edu

In the past three decades, a number of important advances have been made in the treatment of patients with chronic obstructive pulmonary disease (COPD) [1]. For example, supplemental oxygen therapy and smoking cessation have resulted in improved traditional outcome measures, such as mortality [2] and rate of forced expiratory volume in one second (FEV1) decline [3]. Although these end-points are important to clinicians and patients alike, survival and physiological measures do not fully represent the experiences of patients with COPD. Relying exclusively on mortality and physiological outcomes for evaluating treatment effectiveness could result in the decision that some treatments offer no benefits, when in fact they provide significant and important benefits to patients and their families. A treatment that has no effect on mortality or FEV1, for example, may significantly improve the patient's vitality or other functional areas of daily living. Needless to say, reducing the personal and social burden of disease by improving patients' symptoms, functional status and quality of life are important goals. Self-perceived health outcomes are often most relevant and important to patients and their loved ones because they capture the patients' experiences and perspectives.

Increasingly, it has been recognised that health status, especially health-related quality of life, is an important outcome of medical care [4]. The development of reliable and valid tools that measure health status has contributed to a substantial increase in the use of these instruments. Since more of these tools are being used and reported, it is important for clinicians and researchers to understand their strengths and limitations [5].

This article reviews major concepts and methods in health-status assessment for patients with COPD. Its focus is for the clinician and the clinical researcher looking to have a better understanding of how health status and, in particular, health-related, quality-of-life measures are used in clinical trials. In addition, this article will address what work needs to be done so that clinicians can interpret and apply the results of studies using health status as an outcome measure with the ultimate goal of improving their patients' lives.

Terms and definitions

Relationship between health outcomes, health status and quality of life

Health outcomes represent a broad group of end-points used in clinical trials and other clinical research to assess the efficacy or effectiveness of interventions and to assess disease outcomes [6]. Traditional health outcomes include mortality, number of hospital admissions and FEV1. More recently, there has been a growing body of research concerning end-points that are assessed directly by patients and can be termed "patient-reported health outcomes". These can be divided into four categories: health status, health utilities, adherence to treatment and patient satisfaction with healthcare. One perspective on the definition of these terms will be described, as well as the distinctions and overlap in these terms [7, 8], although other conceptualisations also exist [9]. Health status can be defined as the impact of health on a person's ability to perform and derive fulfilment from the activities of daily life. A patient's self-reported health status thus includes health-related quality of life and functional status.

The expression "quality of life" seems easy to understand and yet it can be difficult to define [7]. One definition of quality of life is the "holistic, self-determined evaluation of satisfaction with issues important to the individual" [10, 11]. A person's quality of life can be influenced by a number of factors. The degree to which a patient's health status affects their self-determined evaluation of satisfaction, or quality of

life, has been defined as health-related quality of life [10]. In theory, individuals with no health problems should enjoy a good health-related quality of life. This does not preclude them from other experiences that may affect their overall quality of life, such as poverty or family strife. Quality of life is considerably more comprehensive than health status and includes aspects of the environment that may or may not be affected by health or treatment. The term health-related quality of life is used to indicate that the outcome measure is focused on the health concept or aspects of human life and activities that are generally affected by health conditions or health services [8].

Functional status refers to a person's ability to perform a variety of physical, emotional and social activities. An individual's functional capacity will be influenced and potentially limited by their overall health. Although functional status and quality of life overlap, they are conceptually distinct from one another and sometimes do not correlate highly [10, 12].

Many health-status instruments include items that measure both functional status and health-related quality of life, making it difficult to separate the effect of health on these two concepts. Furthermore, for some patients and caregivers, health-related quality of life may affect diverse realms of their overall quality of life not usually considered "health related". Whatever concepts are applied to a health status or quality of life in various applications of the measuring instrument, they should be matched as closely as possible to the measure's purpose for each specific application and with reference to the theoretical model constructed for the given application.

Assessment of health-status instruments

Before researchers and clinicians use health-status instruments there should be established evidence (preferably published) of the instrument's reliability, validity and responsiveness. An instrument by itself is not inherently valid, but rather it is valid for the specific uses for which it has been evaluated. Without performing the research to document these basic measurement properties of an instrument, it is not possible to determine if the instrument is capable of detecting a true signal as opposed to noise [4].

Reliability

Reliability refers to the amount of error found in any form of measurement and can be considered as the degree to which an instrument will give the same result when measuring the same phenomenon under varying circumstances. Practically, however, reliability translates into how reproducible the results of an instrument are when applied under various conditions. Reliability can be assessed by a number of methods that include internal consistency, intra- and inter-rater reliability and stability. An instrument is internally consistent if its different components, administered at the same time, yield similar results; the commonly used statistical method to assess internal consistency is Cronbach's alpha. Reliability includes the agreement between different observers (inter-rater reliability) and between the same observer on different occasions (intra-rater reliability). Stability refers to the reproducibility over time and can be described by assessing the same subject on different occasions under circumstances when the concept being measured, such as health status, has not changed (test/retest reliability). Once an instrument has been demonstrated to be reliable, it does not necessarily need to be retested for reliability with every use, unless it is to be applied in a new, untested population.

Validity

Whereas reliability refers to reproducibility, validity commonly refers to an instrument's ability to measure what it purports to measure. Classically, validity has been described by the "3 Cs": content, construct and criterion. Content validity refers to the instrument's ability to reflect the domains of the concepts it purports to measure. Many experts agree that the content validity of any patient's self-reported measure can be judged best by the persons or populations being assessed (although there is a long tradition of professionals and "experts" defining the domains and the items contained in the measures) [7]. For example, the Chronic Respiratory Questionnaire (CRQ) was developed to assess the health-related quality of life for patients who suffer from COPD and the items were generated from interviewing COPD patients [13]. It would be expected that a group of patients with COPD would feel that the questionnaire reflected their health-related quality of life, but it would likely have poor content validity if it were administered to a group of diabetics to measure their health-related quality of life. The more that the content of an instrument deviates from the content of the concept under study, the greater the error that is introduced and the lesser the accuracy of the inferences. Ideally, a health-related, quality-of-life instrument should contain questions that cover all domains of health-related quality of life for the population under study.

Criterion validity refers to the ability of an instrument to test a subject in comparison with an accepted "gold" standard. This can occur in the present (concurrent validity) or can be used to predict the future (predictive validity). Concurrent validity is most commonly tested when comparing a new test with an existing standard, with the intent of replacing that existing standard. Predictive validity can be assessed only by applying the instrument and finding out how well it predicts the outcome under study at a later time. Few opportunities exist for applying the logic of concurrent or gold-standard validity to self-reported, health-status measures, particularly for the concepts and domains that are inherently subjective and unobservable, since no gold standard exists.

How is it shown that an instrument is valid if a gold standard does not exist? This is the situation often faced when measuring relatively new concepts, such as health-related quality of life. When there is no gold standard, theoretical "constructs" must be relied upon to infer validity. A number of methods can infer construct validity, but because no measure can unequivocally prove it, construct validity takes the form of ongoing hypothesis testing and rarely can be considered "finished". The most common methods of assessing construct validity are through convergent and divergent validity. How well does the construct under question correlate with other measures that assess the same or related constructs? The instrument should correlate with both related clinical and health-status instruments (convergent) and should not correlate with unrelated or dissimilar ones (divergent). For example, the physical function domains of the Seattle Obstructive Lung Disease Questionnaire (SOLDQ), a disease-specific, health-status instrument, has been demonstrated to correlate with physiological parameters (FEV1, 6-min walk), as well as previously validated and reliable generic and disease-specific, quality-of-life measures (Short Form (SF)-36 and CRQ, respectively) [14]. The consistency across these comparisons supports its validity.

As a specific example of assessing construct validity, its theory would suggest that the SOLDQ physical function domain score should have a positive and significant correlation with FEV1 [14]. In fact, the physical function correlation coefficient (r) is 0.26, suggesting that as FEV1 increases so does health

status. That the correlation is relatively low, however, suggests that the physical function domain may be capturing information in addition to the physical limitation explained by a subject's FEV1.

Determining the responsiveness of an instrument is imperative in trials that test an intervention. The responsiveness of a measure refers to the ability of a test to detect change over time, and it can be viewed as the ability of a particular measure in a particular application to detect minimally important changes, which is important when testing an intervention. Unlike FEV1, where clinicians have a conceptual understanding of how much clinical change occurs with an improvement of 200 mL, health-status values often do not have the same shared meaning.

It is important for investigators to report not only the degree of change, but whenever possible, to determine and explicitly state what defines a clinically relevant change. This is true whether it is either a positive change (such as would define the success of a treatment) or a negative change (reflecting significant deterioration in health). The best method to determine the "minimally important difference" (MID) is a matter of some contention [15–18]. Some authors have used patients' or clinicians' judgments of whether the patient improved since the prior measurement period (often after treatment), and divided patients into groups with significant change and those without significant change. The MID was estimated as the amount of change seen in the group identified as slightly or mildly improved. In a different approach of defining the MID, patients were asked to judge themselves in relation to others currently, rather than in relation to the way they were at a prior assessment. Both methods seemed to define the same MID on the CRQ. Although both of these methods have been criticised, they remain the most commonly used and best accepted methods. The MID will be discussed further in the section on improving instrument use.

Additional measurement issues

Another important issue in the use of health-status instruments is the concept of the ceiling and floor effect, which occurs when a large proportion of scores cluster at the highest or lowest possible value, respectively. For example, among patients who had moderate-to-severe COPD, approximately two-thirds of subjects had the worst possible score on the role-physical domain of the SF-36 version of the Medical Outcomes Study (MOS) questionnaire [19]. This problem limits the ability of an investigator to make any distinction among subjects or to demonstrate worsening physical function in any subject that scores at the "floor" or the worst possible score. Understanding the intended population for which the instrument was designed and the anticipated distribution of scores may help avoid this problem.

Patients are not the only source of reports. The use of proxy reports is an important issue because participants in studies of chronic diseases, such as COPD, include individuals who cannot speak for themselves at some or all points in the course of the condition or the evaluation of the treatment. Proxy reports may help to limit missing data in longitudinal evaluation. This issue is of major significance for measuring health and quality of life in patients with end-stage respiratory diseases. Confining an evaluation to a patient's report only excludes an important part of this population. Observable domains, such as behaviours, are more accessible and are reported more accurately by family members or other caregivers than domains such as feelings or mood [7, 20]. Self-reports from persons who are cognitively impaired or persons with communication problems cannot be neglected, just as the reliability and validity of proxy reports must be examined.

Relevance and use of health-status measurements

Health-status measurements are by definition "subjective" and thus the data generated from questionnaires are sometimes viewed as "soft", but it is axiomatic that health-status measurements are assessed from a patient's perspective. The true underlying question is "Are patients' perspectives reliable?" An essential feature of a "hard measure" is its reliability [21] and thus health-status measures can be considered "hard" if they are reliable. In fact, many instruments have been demonstrated to have adequate reliability and some are regarded as excellent in this respect [14, 19, 22–28].

Health-status instruments do not have a strong correlation with physiological measures, such as FEV1 [14, 23, 26, 28–30]. This is not a limitation, but rather this reflects that different individuals with the same physiological limitations will experience different effects on their health status. The purpose of health-status instruments is not to replace physiological measurements but to add to the understanding of what variables lead to decrements in health status [29]. If the FEV1 by itself explained all the decrements in the coping, physical and emotional functions of patients with COPD, then, when compared with reliable health-status measures, the correlation coefficient would be one. Consequently, there would be no need to measure health status because FEV1, which is easier to measure, would capture all the information.

Measuring a person's health status, however, may not always be appropriate in every study. Health-status measures should be used only if they are relevant to the hypothesis being tested. Although health status can provide useful information, the major end-point should sometimes be a physiological or survival outcome. For example, a trial that examined the effect of pneumococcal vaccination on the incidence of pneumococcal pneumonia would not likely benefit from administering a health-status instrument. It is unlikely that the vaccination would produce an effect sufficient to make measuring health status feasible or reasonable. In contrast, health status can be of critical importance when an intervention's major goal is to improve a patient's functional status, such as with lung volume reduction surgery (LVRS). When measuring health-related quality of life before and after the operation, estimates of the effect of LVRS can be made on patients' quality of life [31–33].

Cost-effectiveness analyses are a common approach for evaluating the economic impact of medical care technologies [7, 34]. A cost-effectiveness analysis produces a ratio, such as the cost per year of life gained, where the denominator reflects the gain in health from a specific intervention and the numerator reflects the cost in dollars of obtaining that gain [34]. Cost utility is a type of cost-effectiveness where effects are expressed as utilities, such as quality-adjusted survival, facilitating comparisons across different diseases and interventions (*e.g.* quality-adjusted life years (QALYs)). The core purpose of these analyses is to determine the value or trade-off of a therapy or programme. In other words, for a therapy known to be effective, cost-effectiveness analysts ask "What is the cost to achieve that effect (gain in survival)?" This is expressed as the ratio of the incremental, or additional, costs divided by the incremental effects. Measuring the denominator, or QALYs, requires information about quality of life that permits the investigator to adjust survival for the quality of that survival. The QALY is a measure of health outcome that assigns to each period a weight ranging from zero to one, corresponding to the quality of life during that time, where a

weight of one corresponds to perfect health and a weight of zero corresponds to a health state judged equivalent to death. The number of QALYs represents the number of "healthy years of life" that are valued equivalently to the actual outcome [34]. This numeric value for a given health state is a utility. Utilities can be measured either by comparing one health-outcome state with another, using techniques such as the standard gamble [34] or time trade-off [7, 35] or by administering a health-status instrument for which standard utility weights have been determined, such as the Quality of Well-Being (QWB) scale [36, 37], the European Quality of Life Questionnaire EQ-5D [38] or the Health Utilities Index [39]. These different methods have not been compared directly in respiratory disease; the QWB has been used most extensively in COPD [40].

Generic *versus* disease-specific instruments

Health-status instruments can be designed to measure either general health status or the effect of a specific disease on health status. The decision about whether to use a disease-specific *versus* a generic questionnaire should depend on the question being asked. In many instances, both generic and disease-specific questionnaires should be considered [41–43]. Generic measures are broader in scope and applicability. The strengths of generic measures are that they are capable of detecting the effects of the diverse aspects one disease beyond those captured by a disease-specific measure and they are capable of comparing health status across multiple diseases [44]. In addition, these measures may be more likely to detect unexpected effects of an intervention that does not relate to respiratory health. No definitive evidence in head-to-head comparisons supports the use of one generic health-status measurement over another in COPD. Domains that reflect physical limitations generally share similar correlations to a 6-min walk and FEV1 [16, 23, 45–48]. In a direct comparison between the Nottingham Health Profile (NHP) and the SF-36 for patients with mild-to-severe COPD, both instruments demonstrated similar area under the receiver-operating characteristics curves, however, the NHP had greater ceiling and floor effects than the SF-36 [23]. The instruments are designed to be broad in scope and applicability and thus generic measures are often limited by questions that do not fully cover a disease-specific condition and may not be as responsive to change as disease-specific measures [46].

The disease- or condition-specific instruments focus on one condition and attempt to define its effects on a patient's health status. For example, as with many respiratory disease-specific questionnaires, the St George's Respiratory Questionnaire (SGRQ) and the CRQ include questions designed to assess the effect of dyspnoea on everyday activities. These instruments are more likely to be responsive to change in clinical status with treatments that target respiratory symptoms [46, 49]. Among patients with severe COPD, the SGRQ demonstrated a floor effect in 26% of patients, however, both the CRQ and the SGRQ were discriminative in distinguishing patients with different levels of disease severity [19]. In head-to-head comparisons, both the CRQ and the SGRQ physical domains demonstrated similar correlations with physiological measurements and exercise tolerance [47].

Generic health-status instruments used in chronic obstructive pulmonary disease

A number of generic health-related, quality-of-life instruments have been used to characterise COPD. The following is a discussion of the instruments that are most widely used.

The Sickness Impact Profile (SIP) was developed in 1972 as a result of collecting statements describing behavioural dysfunction attributable to illnesses from sick and healthy patients, as well as professional and nonprofessional caregivers. This 136-item questionnaire covers a broad number of domains and dimensions that include physical activity (ambulation, mobility, body care and movement), psychosocial functioning (societal interactions, alertness behaviour, emotional behaviour and communication), as well as five independent domains: sleep and rest, eating, work, home management, and recreation and past-times [27]. The SIP has been well validated and demonstrated to be reliable and responsive among patients with COPD [43, 50–53]. It has been used to describe the effect of COPD on patients' health status, although it may not be discriminative of mild COPD. It has also been used as an outcome variable in multicentre studies such as the Nocturnal Oxygen Therapy Trial and the Intermittent Positive-Pressure Breathing study [2, 54] and in interventional studies to assess the effects of LVRS [55] and antidepressant therapy [56] on health status. This instrument was designed to be either interviewer- or self-administered. Its major disadvantage is the relatively long time it takes to complete, which is ~20–30 min. Another potential disadvantage is that no studies clearly demonstrate how to define the MID in scores. Unpublished data suggest the MID of the SIP is five points [7].

The MOS questionnaire has undergone a number of revisions. The most commonly used form is the previously mentioned SF-36. This instrument has been demonstrated to be reliable [57] and responsive in COPD [24, 28, 58, 59]. The SF-36 is divided into eight domains: physical-functioning, role-physical, bodily pain, general health, vitality, social-functioning, role-emotional and general health. The SF-36 also has two summary scores, a physical component scale and a mental component scale that were developed with factor analysis. These component scores have been demonstrated to explain 70–80% of the variance in the individual domain scores, however, the component scores may not have as good a sensitivity for detecting change as the individual domain scores. The SF-36 component scores have been standardised against national population-based samples to have a mean of 50 and a standard deviation of 10. The physical component scale has been demonstrated to predict hospitalisations and mortality among patients who have self-reported obstructive lung disease [60]. The domains that have the greatest linear trend with dyspnoea scores and FEV1 are the role-physical, physical function and general health scales [28, 61]. The SF-36 offers significant advantages in that it is self-administered, easily completed in ~5 min and has been translated and validated in several languages. Although not originally designed to be a utility measure, attempts have been made to create a single index to be used in the calculation of QALYs [62]. The MID of the SF-36 is reported to be five units, although this has not been replicated in patients with COPD [63].

The NHP uses statements that measure departures from "normal" functioning by affirming particular statements or items that describe health status. The first part contains 38 items that fall into six domains: physical mobility, energy, sleep, pain, social isolation and emotional reactions. The second section contains seven items that address areas of daily living affected by a patient's health status [64]. This instrument has been demonstrated to be valid and reliable among patients with COPD [28, 65] and has been administered among patients with COPD in trials of LVRS [31], inhaled corticosteroids [66], bronchodilator [67] and pulmonary rehabilitation [68]. Very little data exist regarding how much difference represents an MID. This instrument is self-administered and can be completed in ~10–15 min.

The QWB is a health-status instrument that was developed

as part of the Health Status Index [69]. It contains three scales: mobility, physical activity and social activity. Questions determine the functional level of each scale. Weighted values for the different combinations of functional level and symptomatology on each scale were assigned from a randomly selected population-based sample. Interviewees responding to questions were assigned values derived from the population-based sample. This method of scoring allows for the QWB to be transformed to a scale from zero to one and has the unique advantage of being able to be used to calculate QALYs in cost-effectiveness analysis, meaning that the QWB is the only health-status measure described here that allows for the calculation of QALYs. A single cross-sectional study has been performed, demonstrating validity among patients with COPD [36] and it is unclear what scoring difference counts as an MID. Direct comparisons of multiple health-status instruments suggest that this generic utility measure may be less responsive at detecting health-status changes in patients who have undergone pulmonary rehabilitation than some of the disease-specific measures [46, 70]. This instrument can be either interviewer or self-administered and takes ∼20 min to complete.

Disease-specific measures in chronic obstructive pulmonary disease

The CRQ was developed >10-yrs-ago through qualitative interviews of patients with chronic lung disease and it has been used extensively in examining patients′ health status among COPD patients [45]. The instrument has proven to be reliable, valid and responsive to change [46–48, 71]. Administered by an interviewer, it contains 20 questions that cover four domains: dyspnoea, fatigue, emotional function and mastery. The CRQ individualises the dyspnoea domain by asking the patient to identify activities that make them dyspnoeic and to rate the degree of dyspnoea. A difference in score of 0.5 per question has been determined to be an MID [18, 72, 73]. A unique feature of this questionnaire is the ability to assess limitations in patient-specific activities. As the dyspnoea scale is individualised in this way, comparison between patients is difficult on this scale, but the scale may be put to better use when comparisons are made in the same individual [22, 74]. This instrument has been demonstrated to be more sensitive to change than generic health-status instruments, such as the NHP and the SIP [29, 46]. The CRQ has been used to assess the effects of bronchodilators [71, 75, 76] and to make comparisons between bronchodilators or aerosol delivery mechanisms [45], modes of ventilation [77], long-term oxygen therapy [78] and pulmonary rehabilitation [73, 79–85]. It has been translated into many languages including Dutch [86], Spanish [25, 87] and German [88]. Although the CRQ has been adapted for self-administration, reliability and validity data are limited. Therefore, its primary limitation for large population-based use is the need for in-person interviews. Skilled interviewers can generally complete the first administration of the CRQ in 20–30 min with subsequent administrations taking ∼10 min.

The SGRQ has been used extensively for patients with COPD and several other chronic pulmonary diseases. This instrument has been demonstrated to be valid, reliable and responsive among patients with COPD [16, 30, 49, 52]. It contains 50 items with 76 weighted responses that cover three domains: symptoms, activity and impact. In addition to the domain scores, a total score is calculated. Each item has been weighted, based on empirically derived values from 160 patients with asthma [16], and the weighting was validated among patients with COPD [30]. The SGRQ is scaled from zero to 100 (with zero representing the best health-related quality of

life). The MID in a score that signifies a clinically significant change is four points.

The SGRQ is able to detect decrements in health-related quality of life among patients with mild disease [89] and has been demonstrated to discriminate between those patients who had mild-to-severe COPD, determined by the American Thoracic Society (ATS)-staging classification. Furthermore, the SGRQ had better sensitivity to detect impairments in health-related quality of life than the generic measure, the SIP [90]. It has also been used to describe the magnitude of the effect of COPD exacerbations on health status [91]. Among elderly patients, completion rates may be lower than other instruments with structured interviews [19]. The SGRQ has been used to assess the effects of medications, such as bronchodilator therapy [49], long-term oxygen therapy [92], inhaled corticosteroids [93] and antibiotic therapy [94]. In addition, it has been used to assess the effects of patient self-management [95], patient education strategies [96], nurse specialist care [97] and psychotherapy among patients with COPD [98]. Finally, noninvasive ventilation among patients with severe COPD has been assessed using the SGRQ [99]. The SGRQ was designed to be self-administered and usually can be completed within 20 min. It was developed in English, but it has been translated into over 35 languages and dialects. An American version of this UK-developed questionnaire has been developed and validated with some of the language modified to make it consistent with American-English [100].

The SOLDQ, a relatively new instrument, was designed to examine patients with asthma or COPD [14]. The instrument contains 29 items that cover three health domains and a satisfaction with care domain. These domains are the physical function, emotional function, coping skills and treatment satisfaction. Each domain score is transformed (but not normalised) on a scale of one to 100. The MID in the score of the physical function domain, adapted from the SF-36 physical function domain, has been estimated to be six points. The SOLDQ instrument has been demonstrated to be valid and responsive to change among American veterans, but has yet to be used extensively outside this population. A recent study demonstrated that the three health domains of the SOLDQ were predictive of all-cause mortality and hospitalisations for COPD and related illness [60]. The physical function component and domain scores of SF-36 and the SOLDQ had more predictive validity than the mental component or the other domain scores of the SOLDQ. Overall, the SOLDQ physical function score had greater predictive validity than the SF-36 physical component score. The SOLDQ is self-administered and can be completed in ∼10–15 min.

The Quality of Life for Respiratory Illness Questionnaire is a disease-specific, quality-of-life measure designed for patients with reversible and fixed airway obstruction. Patients are asked how much of a problem each item has been during the past year. This instrument contains seven domains and 55 items and each item is rated on a seven-point Likert scale. The domains are breathing problems, physical problems, emotions, situations triggering or enhancing breathing problems, daily and domestic activities, social activities, relationships and sexuality, and general activities. Although test/retest reliability and responsiveness have not been reported, the instrument shows internal consistency and construct validation [101, 102].

Finally, there is an additional health-status instrument that was developed for use among a subset of patients with COPD, those with chronic respiratory failure. The Maugeri Foundation Respiratory Failure Questionnaire-28 has 28 items and was developed in Italian but has been translated into many languages. Although this instrument is only validated in this subset of patients with severe COPD, in this group, the

instrument has been shown to be reliable, valid and responsive to change with treatment with noninvasive ventilation [103–105].

Choosing the "best" health-status instrument for an individual study

The choice of a particular instrument for an individual study must depend on the research questions, the conceptual model and definition of the concept to be measured and the ability of the candidate instruments to meet these specific needs. In addition, investigators must examine the measurement properties of each available instrument including reliability, validity, responsiveness and interpretability as well as administration mode, time and respondent burden. Table 1 describes the generic and disease-specific instruments that have been most commonly used among patients with COPD and lists some information about many of these characteristics. Table 2 lists these characteristics and their relative importance in three different kinds of studies. This list is based on the criteria defined by the Scientific Advisory Committee of the Medical Outcomes Trust [106]. Furthermore, although most investigators must limit the number of questionnaires used in any study to minimise the burden on study participants, expert opinion and some recent data suggest that both a generic and

a disease-specific instrument should be used in most studies concerning the health status of patients with COPD [41–43].

Improving instrument use

Despite the increased use of health-status instruments in observational and experimental studies, there remain many areas of improvement that could advance the use and interpretability of these instruments. A consensus group convened by the ATS, identified three specific recommendations [10]. First, to establish standards for evaluating, using and interpreting health-status instruments in clinical research concerning chronic lung disease. Second, to facilitate funding for research on health-status measurement in chronic lung disease. Third, to promote education and training concerning the measurement of health status. These recommendations are as applicable today as they were 6-yrs-ago when this group was convened.

In addition to these recommendations, three specific areas of research are needed to improve the use and interpretability of these instruments. These areas include head-to-head comparisons of existing instruments, increased research concerning the best methods to assess responsiveness and to identify the MIDs, and methods to increase the interpretability of these instruments for clinicians and policy-makers.

Table 1. – Selected commonly used health-related quality-of-life instruments

Name	Items n, estimated time	Administration	Domains	Scale (worse–best); MID	Validity and reliability in patients with COPD
Generic instruments					
SIP	136, 20–30 min	Designed for either self or interviewer, telephone	12 Domains, two component scores, physical and psychosocial scores, five independent categories, one total score	0–100; MID unknown	Well documented, may not be discriminative with mild COPD
MOS, SF-36	36, 5–10 min	Self, interviewer, telephone	Eight domains	PCS, MCS; 0–100 Normalised to population mean±SD, 50±10. Domains; 0–100; MID five points	Valid and reproducible, responsiveness not well documented in COPD
NHP	50	Interviewer	Three domains: mobility, physical activity, social activity	0–1; MID unknown	Valid, reproducible, responsiveness not demonstrated
QWB	38, 10–15 min	Self	Six domains	100–0; MID unknown	Valid, reproducible, responsiveness not demonstrated
Disease-specific					
CRDQ	20, 20–30 min initial; 10–20 min repetitive	Designed primarily for interviewer; may be self-administered	Dyspnoea, emotional function, fatigue mastery	1–7 Likert scale; MID 0.5 per question for all domains	Well documented in multiple trials
SGRQ	76, 20 min	Self	Symptoms, activities, impact, overall score "no well-being"	100–0; MID four for all domains	Well documented in multiple trials
SOLDQ	29, 10–15 min	Self	Physical function, coping skills, emotional function, satisfaction	0–100; MID six for physical function domain	Documented in original description

MID: minimally important difference; COPD: chronic obstructive pulmonary disease; SIP: Sickness Impact Profile; MOS: Medical Outcomes Study; SF: Short Form; PCS: Physical Component Score; MCS: Mental Component Score; NHP: Nottingham Health Profile; QWB: quality of well-being; CRDQ: Chronic Respiratory Disease Questionnaire; SGRQ: St George's Respiratory Questionnaire; SOLDQ: Seattle Obstructive Lung Disease Questionnaire.

Table 2. – Assessing health-status instruments for chronic obstructive pulmonary disease studies (based in part on the criteria defined by the Medical Outcomes Trust)

Quality to be assessed	Importance of each quality, for use in all or some studies using health-status measures		
	Outcome in a randomised, controlled trial	Discriminate between groups	Predict future outcomes
Measurement model			
Conceptual model specified	All	All	All
Model verified empirically	Some	Some	Some
Reliability			
Internal consistency	Some	All	All
Stability (test/retest reliability)	All	All	All
Inter-rater/intra-rater reliability	All	All	All
Validity			
Content validity	All	All	All
Construct validity	All	All	All
Predictive validity	Maybe	Maybe	All
Responsiveness			
Detection of change	All	Maybe	All
Identification of MID	All	Some	Maybe
Interpretability	All	All	All
Alternative modes of administration			
Testing for comparison	All	Maybe	Maybe
Cultural and language adaptations			
Conceptual and semantic equivalence	Maybe	Maybe	Maybe
Validity testing in different cultures	Maybe	Maybe	Maybe

MID: minimally important difference; All: important in all studies of this type; Some: important in most studies of this type; Maybe: may be important depending on the specific study. Adapted from data from [106].

Head-to-head comparisons of health-status instruments in chronic obstructive pulmonary disease

There are dozens of health-status instruments that assess COPD, but few head-to-head comparisons that evaluate these instruments in the same population, making it difficult to determine which instrument may perform best in a given setting. Head-to-head comparisons, although important, can be difficult to publish in the clinical literature unless the study has clinical implications that extend beyond measurement issues. Studies designed to compare health-status instruments, however, can be usually designed simultaneously to address other research questions of clinical importance.

Responsiveness and the minimally important difference

There is currently no standard or accepted method to assess what constitutes an MID or to compare the responsiveness of different instruments. As described above, two main methods have been used to assess the MID [16–18]. The important question being asked for both MID and responsiveness is "How does one know that the group who reported change really did change?" If a study is large enough, a statistically significant difference may not be clinically important. Therefore, it is important in COPD to identify external criteria to help interpret the magnitude of change seen in a given instrument [107]. Responsiveness involves hypothesis testing about the direction of the change (*i.e.* worse, same, better) and interpretation of the results as to the magnitude of the change (*i.e.* small, medium or large). Assessing the MID involves identifying the smallest change that is important to patients and, potentially, to their families and clinicians. Repeated use of these measures within COPD populations is perhaps the most assured way of identifying the MIDs for different instruments or across instruments over time.

A burgeoning literature focuses on responsiveness because all stakeholders in the evaluation of treatments want to be able to judge if the treatment is effective, how effective it is, under what conditions and in what populations. Statistical reporting of effect sizes has long been the focus of clinical trial methodologies, with consensus around the definition of effect size and standardisation of effect sizes [108, 109]. Alternative measures of responsiveness have been proposed, essentially changing what is in the denominator or the measure of variability or noise [110]. Which statistical measure of change should be calculated and reported across treatment effectiveness evaluations remains an important question for scientific inquiry and consensus development.

Interpretability

The field of health-status measurement is in need of the development of specific methods to translate instrument scores into clinically interpretable results. This process of increasing interpretability begins with selecting and justifying the external measure used to define the importance of the change, which is the same process needed to identify the MID. Interpretability also requires that the relationship between the self-reported, health-status measure and the most relevant external measures be examined. The external measures used could vary widely by including groups of patients with different severity of COPD defined by FEV1 or other physiological measures, or by using events, such as hospitalisation, for an exacerbation of COPD. In addition, these external measures can also include global evaluations of change made by patients, families, or clinicians, or any external measure that is logically related to the self-report measure, is moderately correlated with that measure and is more interpretable to patients or clinicians than the health-status measure under examination. This work needs to be done before greater acceptance of these instruments outside of the academic research community can be expected.

Novel uses of health-status instruments

Health-status instruments have been used primarily as outcome measures. Novel research uses of these instruments include the prediction of clinical outcomes, such as mortality or hospitalisations [60]. Such use may allow clinicians and healthcare systems to identify individuals at risk for poor outcomes who may benefit from specific interventions. This could lead to better allocation of healthcare resources and the ability to target patients who may be at higher risk of poor outcomes. There has also been interest in using these measures to improve clinical care by feeding results back to providers, but studies using health-status measures for this purpose have not convincingly demonstrated significant benefit [111].

Conclusions

Health-status instruments have provided valuable insights into the effects of disease and benefits of treatments for patients with chronic obstructive pulmonary disease. Many instruments have reported evidence on validity and reliability, although only a few have been demonstrated to be responsive to change. These instruments will have an important role in assessing the efficacy and effectiveness of new treatments for patients with chronic diseases, such as chronic obstructive pulmonary disease. To maximise the potential for these instruments, further work needs to be performed to compare different instruments, develop standardised methods for assessing the minimally important difference and increase the interpretability of the instruments. Novel uses for these instruments will undoubtedly improve their use and interpretability.

References

1. Pauwels RA, Buist AS, Calverley CR, Jenkins CR, Hurd SS, on behalf of Committee. Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease, NHLBI/WHO Global Initiative for Chronic Obstructive Lung Disease (GOLD) Workshop Summary. *Am J Respir Crit Care Med* 2001; 163: 1256–1276.
2. Nocturnal Oxygen Therapy Trial Group. Continuous or nocturnal oxygen therapy in hypoxemic chronic obstructive lung disease: a clinical trial. *Ann Intern Med* 1980; 93: 391–398.
3. Anthonisen NR, Connett JE, Kiley JP, *et al.* Effects of smoking intervention and the use of an inhaled anticholinergic bronchodilator on the rate of decline of FEV1. The Lung Health Study. *JAMA* 1994; 272: 1497–1505.
4. Guyatt GH, Feeny DH, Patrick DL. Measuring health-related quality of life. *Ann Intern Med* 1993; 118: 622–629.
5. Gill TM, Feinstein AR. A critical appraisal of the quality of quality-of-life measurements. *JAMA* 1994; 272: 619–626.
6. Clancy CM, Eisenberg JM. Outcomes research: measuring the end results of health care. *Science* 1998; 282: 245–246.
7. Patrick DL, Erickson P. Health Status and Health Policy: Quality of Life in Health Care Evaluation and Resource Allocation. New York City, Oxford University Press, 1993.
8. Patrick DL, Chiang YP. Measurement of health outcomes in treatment effectiveness evaluations: conceptual and methodological challenges. *Med Care* 2000; 9: II14–II25.
9. Carone M, Jones PW. Health status "quality of life". *In*: Donner CF, Decramer M, eds. Pulmonary Rehabilitation. *Eur Respir Mon.* 2000; 13: 22–35.
10. Curtis JR, Martin DP, Martin TM. Patient-assessed health outcomes in chronic lung disease: what are they, how do they help us, and where do we go from here? *Am J Respir Crit Care Med* 1997; 156: 1032–1039.
11. Patrick DL, Danis M, Southerland LI, Hong G. Quality of life following intensive care. *J Gen Intern Med* 1988; 3: 218–223.
12. Patrick DL, Kinne S, Engelberg RA, Pearlman RA. Functional status and perceived quality of life in adults with and without chronic conditions. *J Clin Epidemiol* 2000; 53: 779–785.
13. Guyatt GH, Townsend M, Berman LB, Pugsley SO. Quality of life in patients with chronic airflow limitation. *Br J Dis Chest* 1987; 81: 45–54.
14. Tu SP, McDonell MB, Spertus JA, Steele BG, Fihn SD. A new self-administered questionnaire to monitor health-related quality of life in patients with COPD. Ambulatory Care Quality Improvement Project (ACQUIP) Investigators. *Chest* 1997; 112: 614–622.
15. Jones PW. Interpreting thresholds for a clinically significant change in health status in asthma and COPD. *Eur Respir J* 2002; 19: 398–404.
16. Jones PW, Quick FH, Baveystock CM. The St George's respiratory questionnaire. *Respir Med* 1991; 85: Suppl. B, 25–31.
17. Juniper EF, Gordon GH, Willan A, Griffith LE. Determining a minimal important change in a disease-specific Quality of Life Questionnaire. *J Clin Epidemiol* 1994; 47: 81–87.
18. Redelmeier DA, Guyatt GH, Goldstein RS. Assessing the minimal important difference in symptoms: a comparison of two techniques. *J Clin Epidemiol* 1996; 49: 1215–1219.
19. Harper R, Brazier JE, Waterhouse JC, Walters SJ, Jones NM, Howard P. Comparison of outcome measures for patients with chronic obstructive pulmonary disease (COPD) in an outpatient setting. *Thorax* 1997; 52: 879–887.
20. Hinton J. How reliable are relatives' retrospective reports of terminal illness? Patients' and relatives' accounts compared. *Soc Sci Med* 1996; 43: 1229–1236.
21. Feinstein AR. Clinical biostatistics: hard science, soft data, and the challenges of choosing clinical variables in research. *Clin Pharmacol Ther* 1977; 22: 485–498.
22. Wijkstra PJ, TenVergert EM, van Altena R, *et al.* Reliability and validity of the chronic respiratory questionnaire (CRQ). *Thorax* 1994; 49: 465–467.
23. Prieto L, Alonso J, Ferrer M, Anto JM. Are results of the SF-36 health survey and the Nottingham Health Profile similar? A comparison in COPD patients. Quality of Life in COPD Study Group. *J Clin Epidemiol* 1997; 50: 463–473.
24. Keller SD, Bayliss MS, Ware JE Jr, Hsu MA, Aamiano AM, Goss TF. Comparison of responses to SF-36 Health Survey, questions with one-week and four-week recall periods. *Health Serv Res* 1997; 32: 367–384.
25. Guell R, Casan P, Sangenis M, *et al.* The Spanish translation and evaluation of a quality-of-life questionnaire in patients with chronic obstructive pulmonary disease. *Arch Bronconeumol* 1995; 31: 202–210.
26. Ferrer M, Alonso J, Prieto L, *et al.* Validity and reliability of the St George's Respiratory Questionnaire after adaptation to a different language and culture: the Spanish example. *Eur Respir J* 1996; 9: 1160–1166.
27. Bergner M, Bobbitt RA, Carter WB, Gilson BS. The Sickness Impact Profile: Development and final revision of a health status measure. *Med Care* 1981; 18: 787–805.
28. Alonso J, Prieto L, Ferrer M, *et al.* Testing the measurement properties of the Spanish version of the SF-36 Health Survey among male patients with chronic obstructive pulmonary disease. Quality of Life in COPD Study Group. *J Clin Epidemiol* 1998; 51: 1087–1094.
29. Tsukino M, Nishimura K, Ikeda A, Koyama H, Mishimia M, Izumi T. Physiologic factors that determine the health-related quality of life in patients with COPD. *Chest* 1996; 110: 896–903.
30. Jones PW, Quirk FH, Baveystock CM, Littlejohns P. A self-complete measure of health status for chronic airflow limitation. The St George's Respiratory Questionnaire. *Am Rev Respir Dis* 1992; 145: 1321–1327.
31. Cooper JD, Trulock EP, Triantafillou AN, *et al.* Bilateral pneumectomy (volume reduction) for chronic obstructive pulmonary disease. *J Thorac Cardiovasc Surg* 1995; 109: 106–119.

32. Ramsey SD, Patrick DL, Albert RK, Larson EB, Wood DE, Raghu G. The cost-effectiveness of lung transplantation. A pilot study. University of Washington Medical Center Lung Transplant Study Group. *Chest* 1995; 108: 1594–1601.

33. National Emphysema Treatment Trial Research Group. Patients at high risk of death after lung volume reduction surgery. *N Engl J Med* 2001; 345: 1075–1083.

34. Gold MR, Siegel JE, Russell LB, Weinstein MC. Cost-Effectiveness in Health and Medicine. New York, Oxford University Press, 1996.

35. Froberg DG, Kane RL. Methodology for measuring health-state preferences--III: Population and context effects. *J Clin Epidemiol* 1989; 42: 585–592.

36. Kaplan RM, Atkins CJ, Timms R. Validity of a quality of well-being scale as an outcome measure in chronic obstructive pulmonary disease. *J Chronic Dis* 1984; 37: 85–95.

37. Kaplan RM, Feeny DH, Revicki DA. Methods for assessing relative importance in preference based outcome measures. *Qual Life Res* 1993; 2: 467–475.

38. The EuroQol Group. EuroQol--a new facility for the measurement of health-related quality of life. *Health Policy* 1990; 16: 199–208.

39. Drummond MF, Stoddart GL, Torrance GW. Methods for the Economic Evaluation of Health Care Programmes. New York, Oxford University Press, 1987.

40. Ramsey SD, Sullivan SD, Kaplan RM, Hood DE, Chiang YP, Wagner JL. Economic analysis of lung volume reduction surgery as part of the National Emphysema Treatment Trial. *Ann Thorac Surg* 2001; 71: 995–1002.

41. Patrick DL, Deyo RA. Generic and disease-specific measures in assessing health status and quality of life. *Med Care* 1989; 27: 5217–5232.

42. Singh SJ, Sodergren SC, Hyland ME, Williams J, Morgan MD. A comparison of three disease-specific and two generic health-status measures to evaluate the outcome of pulmonary rehabilitation in COPD. *Respir Med* 2001; 95: 71–77.

43. Engstrom EP, Persson LO, Larsson S, Sullivan M. Health-related quality of life in COPD: why both disease-specific and generic measures should be used. *Eur Respir J* 2001; 18: 69–76.

44. Ren XS, Kazis L, Lee A, *et al.* Comparing generic and disease-specific measures of physical and role functioning: results from the Veterans Health Study. *Med Care* 1998; 36: 155–166.

45. Guyatt GH, Berman LB, Townsend M, Pugsley SO, Chambers LW. A measure of quality of life for clinical trials in chronic lung disease. *Thorax* 1987; 42: 773–778.

46. Guyatt GH, King DR, Feeny DH, Stubbing D, Goldstein RS. Generic and specific measurement of health-related quality of life in a clinical trial of respiratory rehabilitation. *J Clin Epidemiol* 1999; 52: 187–192.

47. Hajiro T, Nishimura K, Tsukino M, Ikeda A, Koyama H, Izumi T. Comparison of discriminative properties among disease-specific questionnaires for measuring health-related quality of life in patients with chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 1998; 157: 785–790.

48. Hajiro T, Nishimura K, Tsukino M, Ikeda A, Koyama H, Izumi T. Analysis of clinical methods used to evaluate dyspnoea in patients with chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 1998; 158: 1185–1189.

49. Jones PW, Bosh TK. Quality of life changes in COPD patients treated with salmeterol. *Am J Respir Crit Care Med* 1997; 155: 1283–1289.

50. McSweeny A, Grant I, Heaton R, Adams KM, Timms R. Life quality of patients with chronic obstructive pulmonary disease. *Arch Intern Med* 1982; 142: 473–478.

51. McSweeny A, Heaton R, Grant I, Cugell D, Solliday N, Timms T. Chronic obstructive pulmonary disease: socio-emotional adjustment and life quality. *Chest* 1980; 77s: 309s–311s.

52. Jones PW, Baveystock CM, Littlejohns P. Relationships between general health measured with the sickness impact profile and respiratory symptoms, physiological measures, and mood in patients with chronic airflow limitation. *Am Rev Respir Dis* 1989; 140: 1538–1543.

53. Prigatano GP, Wright EC, Levin D. Quality of life and its predictors in patients with mild hypoxemia and chronic obstructive pulmonary disease. *Arch Intern Med* 1984; 144: 1613–1619.

54. Intermittent Positive Pressure Breathing Trial Group. Intermittent positive pressure breathing therapy of chronic obstructive pulmonary disease. *Ann Intern Med* 1983; 99: 612–620.

55. Cordova F, O'Brien G, Furukawa S, Kuzma AM, Travaline J, Criner GJ. Stability of improvements in exercise performance and quality of life following bilateral lung volume reduction surgery in severe COPD. *Chest* 1997; 112: 907–915.

56. Borson S, McDonald GJ, Gayle T, Deffebvach M, Lakshminarayan S, van Tuinen C. Improvement in mood, physical symptoms, and function with nortriptyline for depression in patients with chronic obstructive pulmonary disease. *Psychosomatics* 1992; 33: 190–201.

57. Mahler DA, Mackowiak JI. Evaluation of the short-form 36-item questionnaire to measure health-related quality of life in patients with COPD. *Chest* 1995; 107: 1585–1589.

58. Limbos MM, Chan CK, Kesten S. Quality of life in female lung transplant candidates and recipients. *Chest* 1997; 112: 1165–1174.

59. Moy ML, Ingenito EP, Mentzer SJ, Evans RB, Reilly JJ Jr. Health-related quality of life improves following pulmonary rehabilitation and lung volume reduction surgery. *Chest* 1999; 115: 383–389.

60. Fan V, Curtis JR, MacDowell M, Fihn SD. Using quality of life to predict hospitalizations and death in chronic obstructive lung disease. *Chest* 2002; 122: 429–436.

61. Viramontes JL, O'Brien B. Relationship between symptoms and health-related quality of life in chronic lung disease. *J Gen Inter Med* 1994; 9: 46–48.

62. Brazier JE, Usherwood T, Harper R, Thomas K. Deriving a preference-based single index from the UK SF-36 Health Survey. *J Clin Epidemiol* 1998; 51: 1115–1128.

63. Ware JE Jr, Snow KK, Kosinski M, Gandek B. SF-36 Health Survey: Manual and Interpretation Guide. Boston, The Health Institute, 1993.

64. Hunt SM, McKenna SP, McEwen J, Backett EM, Williams J, Papp E. A quantitative approach to perceived health status: a validation study. *J Epidemiol Community Health* 1980; 34: 281–286.

65. Monso E, Fiz JM, Izquierdo J, *et al.* Quality of life in severe chronic obstructive pulmonary disease: correlation with lung and muscle function. *Respir Med* 1998; 92: 221–227.

66. van Schayck CP, Dompeling E, Rutten MP, Folgering H, van den Boom G, van Weel C. The influence of an inhaled steroid on quality of life in patients with asthma or COPD. *Chest* 1995; 107: 1199–1205.

67. van Schayck CP, Rutten-van Molken MP, van Doorslaer EK, Folgering H, van Weel C. Two-year bronchodilator treatment in patients with mild airflow obstruction. Contradictory effects on lung function and quality of life. *Chest* 1992; 102: 1384–1391.

68. Fuchs-Climent D, Le Gallais D, Varray A, Desplan J, Cadopi M, Prefaut C. Quality of life and exercise tolerance in chronic obstructive pulmonary disease: effects of a short and intensive inpatient rehabilitation program. *Am J Phys Med Rehabil* 1999; 78: 330–335.

69. Fanshel S, Bush J. A health-status index and its application to health-services outcomes. *Oper Res* 1970; 18: 1021–1066.

70. Ries AL, Kaplan RM, Limberg TM, Prewitt LM. Effects of pulmonary rehabilitation on physiologic and psychosocial outcomes in patients with chronic obstructive pulmonary disease. *Ann Intern Med* 1995; 122: 823.

71. Guyatt GH, Townsend M, Keller J, Singer J, Nogradi S. Measuring functional status in chronic lung disease: conclusions from a randomized control trial. *Respir Med* 1989; 83: 293–297.

72. Jaescheke R, Singer J, Guyatt GH. Measurement of health status. Ascertaining the minimal clinically important difference. *Control Clin Trials* 1989; 10: 407–415.

73. Lacasse Y, Wong E, Guyatt GH, King DR, Cook DJ, Goldstein RS. Meta-analysis of respiratory rehabilitation in chronic obstructive pulmonary disease. *Lancet* 1996; 348: 1115–1119.

74. Morgan MD. Experience of using the CRQ (Chronic Respiratory Questionnaire). *Respir Med* 1991; 85: Suppl. B, 23–37.

75. Mahon J, Laupacis A, Donner A, Wood T. Randomised study of n of 1 trials versus standard practice. *BMJ* 1996; 312: 1069–1074.

76. Mahon JL, Laupacis A, Hodder RV, *et al.* Theophylline for irreversible chronic airflow limitation: a randomized study comparing n of 1 trials to standard practice. *Chest* 1999; 115: 38–48.

77. Elliott MW, Simonds AK, Carroll MP, Wedzicha JA, Branthwaite MA. Domiciliary nocturnal nasal intermittent positive pressure ventilation in hypercapnic respiratory failure due to chronic obstructive lung disease: effects on sleep and quality of life. *Thorax* 1992; 47: 342–348.

78. McDonald CF, Blyth CM, Lazarus MD, Marschner I, Barter CE. Exertional oxygen of limited benefit in patients with chronic obstructive pulmonary disease and mild hypoxemia. *Am J Respir Crit Care Med* 1995; 152: 1616–1619.

79. Vale F, Reardon JZ, ZuWallack RL. The long-term benefits of outpatient pulmonary rehabilitation on exercise endurance and quality of life. *Chest* 1993; 103: 42–45.

80. Bendstrup KE, Ingemann Jensen J, Holm S, Bengtsson B. Out-patient rehabilitation improves activities of daily living, quality of life and exercise tolerance in chronic obstructive pulmonary disease. *Eur Respir J* 1997; 10: 2801–2806.

81. Singh SJ, Smith DL, Hyland ME, Morgan MD. A short outpatient pulmonary rehabilitation programme: immediate and longer-term effects on exercise performance and quality of life. *Respir Med* 1998; 92: 1146–1154.

82. White RJ, Rudkin ST, Ashley J, *et al.* Outpatient pulmonary rehabilitation in severe chronic obstructive pulmonary disease. *J R Coll Physicians Lond* 1997; 31: 541–545.

83. Wijkstra PJ, Ten Vergert EM, van Altena R, *et al.* Long term benefits of rehabilitation at home on quality of life and exercise tolerance in patients with chronic obstructive pulmonary disease. *Thorax* 1995; 50: 824–828.

84. Bernard S, Whittom F, Leblanc P, *et al.* Aerobic and strength training in patients with chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 1999; 159: 896–901.

85. Cambach W, Chadwick-Straver RV, Wagenaar RC, van Keimpema AR, Kemper HC. The effects of a community-based pulmonary rehabilitation programme on exercise tolerance and quality of life: a randomized controlled trial. *Eur Respir J* 1997; 10: 104–113.

86. van den Boom G, Rutten-van Molken MP, Tirimanna PR, van Schayck CP, Folgering H, van Weel C. Association between health-related quality of life and consultation for respiratory symptoms: results from the DIMCA programme. *Eur Respir J* 1998; 11: 67–72.

87. Sans-Torres J, Domingo C, Rue M, Duran-Tauleria E, Marin A. An assessment of the quality of life of patients with COPD and chronic hypoxemia by using the Spanish version of the Chronic Respiratory Disease Questionnaire. *Arch Bronchoneumol* 1999; 35: 428–434.

88. Kirsten DK, Wegner RE, Jorres RA, Magnussen H. Effects of theophylline withdrawal in severe chronic obstructive pulmonary disease. *Chest* 1993; 104: 1101–1107.

89. Ferrer M, Alonso J, Morera J, *et al.* Chronic obstructive pulmonary disease stage and health-related quality of life. The Quality of Life of Chronic Obstructive Pulmonary Disease Study Group. *Ann Intern Med* 1997; 127: 1072–1079.

90. Okubadejo AA, Jones PW, Wedzicha JA. Quality of life in patients with chronic obstructive pulmonary disease and severe hypoxaemia. *Thorax* 1996; 51: 44–47.

91. Seemungal TA, Donaldson GC, Paul EA, Bestall JC, Jeffries DJ, Wedzicha JA. Effect of exacerbation on quality of life in patients with chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 1998; 147: 1418–1422.

92. Okubadejo AA, O'Shea L, Jones PW, Wedzicha JA. Home assessment of activities of daily living in patients with severe chronic obstructive pulmonary disease on long-term oxygen therapy. *Eur Respir J* 1997; 10: 1572–1575.

93. Burges PS, Calverley PM, Jones PW, Spencer S, Anderson JA, Maslen TK. Randomised, double blind, placebo controlled study of fluticasone propionate in patients with moderate to severe chronic obstructive pulmonary disease: the ISOLDE trial. *BMJ* 2000; 320: 1297–1303.

94. Grossman R, Mukherjee J, Vaughan D, *et al.* A 1-year community-based health economic study of ciprofloxacin vs usual antibiotic treatment in acute exacerbations of chronic bronchitis: the Canadian Ciprofloxacin Health Economic Study Group. *Chest* 1998; 113: 131–141.

95. Watson PB, Town GI, Holbrook N, Dwan C, Toop LJ, Drennan CJ. Evaluation of a self-management plan for chronic obstructive pulmonary disease. *Eur Respir J* 1997; 10: 1267–1271.

96. Gallefoss F, Bakke PS, Rsgaard PK. Quality of life assessment after patient education in a randomized controlled study on asthma and chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 1999; 159: 812–817.

97. Ketelaars CA, Schloosser MA, Mostert R, Huyer Abu-Saad H, Halfens RJ, Wouters EF. Determinants of health-related quality of life in patients with chronic obstructive pulmonary disease. *Thorax* 1996; 51: 39–43.

98. Eiser N, West C, Evans S, Jeffers A, Quirk FH. Effects of psychotherapy in moderately severe COPD: a pilot study. *Eur Respir J* 1997; 10: 1581–1584.

99. Meecham Jones DJ, Paul EA, Jones PW, Wedzicha JA. Nasal pressure support ventilation plus oxygen compared with oxygen therapy alone in hypercapnic COPD. *Am J Respir Crit Care Med* 1995; 152: 538–544.

100. Barr JT, Schumacher GE, Freeman S, LeMoine M, Bakst AW, Jones PW. American translation, modification, and validation of the St George's Respiratory Questionnaire. *Clin Ther* 2000; 22: 1121–1145.

101. Maille AR, Koning CJM, Zwinderman AH, Willems LNA, Dijkman JH, Kaptein AA. The development of the "Quality-of-Life for Respiratory Illness Questionnaire (QOL-RIQ)": a disease-specific quality-of-life questionnaire for patients with mild to moderate chronic non-specific lung disease. *Respir Med* 1997; 91: 297–309.

102. Maille AR, Koning CJM, Kaptein AA. Developing a quality-of-life questionnaire for patients with respiratory illness. *Monaldi Arch Chest Dis* 1994; 49: 76–78.

103. Carone M, Anghisi F, Donner CF. Quality of life and chronic respiratory failure. *Monaldi Arch Chest Dis* 1997; 52: 525–528.

104. Carone M, Ambrosino N, Bertolotti G, *et al.* Quality of Life Evaluation and Survival Study: a 3-yr prospective multinational study on patients with chronic respiratory failure. *Monaldi Arch Chest Dis* 2001; 56: 17–22.

105. Clini E, Sturani C, Rossi A, *et al.* Rehabilitation and Chronic Care Study Group - Italian Association of Hospital Pulmonologists (AIPO). The Italian multicentre study on noninvasive ventilation in chronic obstructive pulmonary disease patients. *Eur Respir J* 2002; 20: 529–538.

106. Scientific Advisory Committee of the Medical Outcomes Trust. Assessing health status and quality-of-life instruments: attributes and review criteria. *Qual Life Res* 2002; 11: 193–205.

107. Guyatt GH, Juniper EF, Walter SD, Griffith LE, Goldstein RS. Interpreting treatment effects in randomized trials. *BMJ* 1998; 316: 690–693.

108. Armitage P, Colton T. Encyclopedia of Biostatistics. Chichester, J. Wiley, 1998.

109. Cohen J. Statistical Power Analyses for the Behavioral Sciences. Hillsdale, Erlbaum, 1988.

110. Liang MH. Evaluating measurement responsiveness. *J Rheumatol* 1995; 22: 1191–1192.

111. Espallargues M, Valderas JM, Alonso J. Provision of feedback on perceived health status to health care professionals: a systematic review of its impact. *Med Care* 2000; 38: 175–186.