Minimal Clinically Important Difference for the UCSD Shortness of Breath Questionnaire

David H. Kupferberg, MD, MPH; Robert M. Kaplan, PhD; Donald J. Slymen, PhD; Andrew L. Ries, MD, MPH

- PURPOSE: A minimal clinically important difference (MCID) is important in evaluating clinical measures such as health-related quality of life (HRQOL) instruments. The purpose of this analysis is to evaluate MCID for the UCSD Shortness of Breath Questionnaire (SOBQ).
- METHODS: We examined measures of disease-specific and generic HRQOL in 164 subjects with chronic lung disease before and after pulmonary rehabilitation. Subjects completed 2 disease-specific [SOBQ, Chronic Respiratory Questionnaire (CRQ)], and 2 generic HRQOL measures [RAND-36 and Quality of Well-Being Scale (QWB)]. The MCID was calculated using 3 methods: effect size, standard error of the measurement (SEM), and comparison between the SOBQ and CRQ Dyspnea scores.
- RESULTS: HRQOL measures correlated moderately with measures of maximum exercise tolerance but not with lung function (FEV_1, FVC). HRQOL and exercise capacity improved significantly after pulmonary rehabilitation. A change of 5 units for the SOBQ appears to be a reasonable MCID for this instrument. The calculated MCIDs for the CRQ (0.47/item) and QWB (0.031) were consistent with established change scores.
- CONCLUSIONS: The MCID calculated using an SEM approach for the SOBQ, CRQ, and QWB meets clinical expectations for these instruments. HRQOL measures provide information that is complementary and distinct from physiological measures.

KEY WORDS

chronic lung disease

dyspnea

pulmonary rehabilitation

quality of life

From the Division of Pulmonary and Critical Care Medicine and Department of Family and Preventive Medicine, University of California, San Diego (Dr Kupferberg, Dr Ries); Department of Health Services, University of California, Los Angeles (Dr Kaplan); and the Graduate School of Public Health, San Diego State University, Calif (Dr Slymen).

Supported in part by: Research Training Fellowship, American Lung Association (RT-028-N) and Individual National Research Award NIH/NHLBI (1 F32 HL68358-01) for Dr Kupferberg and NIH Grant R01 HD/HL 30912.

Address correspondence to: Andrew L Ries, MD, MPH, UCSD Medical Center, #8377, 200 W. Arbor Drive, San Diego, CA 92103-8377 (e-mail: aries@ucsd.edu).

Health-related quality of life (HRQOL) has become increasingly important in evaluating health outcomes for patients with chronic lung disease. Interpreting and using such measures is challenging given the variety of questionnaires, each with unique characteristics and scoring methods. Unless one is familiar with a particular instrument, making sense of a change in score is difficult. Estimating the clinical importance of a change in score is relevant to clinicians and patients. The minimal clinically important difference (MCID) has been defined as the smallest difference that patients perceive to be beneficial.[1] This difference may be larger than statistically detectable change. Various approaches have been developed for defining MCID including anchor-based (related to an independent standard) and distribution-based (related to the underlying distribution in a population) methods, but no single approach has been clearly established as best.[2,3] Typically, a variety of methods are applied in an attempt to arrive at a consistent estimate for MCID.

Outcome measures in clinical research in pulmonary diseases have traditionally relied upon physiologic parameters (eg, lung function, exercise tolerance).

Measurement of symptoms such as dyspnea is important, but because of the complex determinants of dyspnea, changes in breathlessness are often independent of physiologic measures.[4–7] In recent years, HRQOL measures have been used more frequently. These questionnaires incorporate multiple dimensions of health into one or a few number of measures. Disease or symptom-specific instruments may be most sensitive to change, have been developed for chronic lung diseases, and are designed to capture changes in key pulmonary symptoms and other health dimensions. Generic HRQOL instruments measure particular generalizable dimensions of health applicable to a variety of disease states, but may not be sensitive to important but modest changes in health, in particular, diseases. In pulmonary rehabilitation, for instance, generic instruments are typically less responsive than disease-specific ones.[8] However, much still needs to be learned about the optimal methods of measuring HRQOL in pulmonary patients.[9]

The overall objectives of this study are to determine an MCID for the UCSD Shortness of Breath Questionnaire (SOBQ) and compare the MCID for the SOBQ to other HRQOL instruments.

Table 1 • BASELINE DEMOGRAPHIC AND PHYSIOLOGIC CHARACTERISTICS OF STUDY POPULATION BEFORE PULMONARY REHABILITATION

Variable	Pre-Rehab
Gender	89M/75F
Age	67.1 ± 8.2
Diagnosis	
Obstructed	143
Mixed obstructed/restricted	17
Restricted	4
Pulmonary function	
FEV_1, L	1.06 ± 0.43
FEV_1, % predicted	45 ± 19
FEV_1/FVC, %	42 ± 13
TLC, L	7.03 ± 1.81
TLC, % predicted	129 ± 23
D_LCO, mL/min/mm Hg	13.1 ± 6.8
D_LCO, % predicted	56 ± 27
Resting PaO_2, mm Hg	74 ± 11
Resting $PaCO_2$, mm Hg	39 ± 6
Resting SaO_2, %	96 ± 2

Values are expressed as mean ± SD.
There was no change after the rehabilitation program significant at $P < .05$.

METHODS

Subjects and Outcome Measures

Data were obtained from 164 subjects with moderate to severe chronic lung disease enrolled in a clinical trial of maintenance after pulmonary rehabilitation.[10] Baseline characteristics are presented in Table 1. Analyses were performed with pooled data from all subjects before and after rehabilitation prior to randomization, since the experimental maintenance intervention was initiated only after the post-rehabilitation evaluation. The pulmonary rehabilitation program involved sixteen 3-hour sessions conducted over 8 weeks and included components of exercise reconditioning, education, physical and respiratory care instruction, and psychosocial support. Physiologic measures included tests of lung function, exercise tolerance, and gas exchange. Five separate HRQOL measures were obtained, including both disease/symptom-specific (SOBQ,[11] Chronic Respiratory Questionnaire [CRQ],[12] Baseline and Transition Dyspnea Index [BDI/TDI][13]) as well as generic (RAND/SF-36,[14] Quality of Well-Being Scale [QWB][15,16]) questionnaires. Psychosocial assessment included measures of depression (CES-D)[17] and self-efficacy for walking.[8]

UCSD Shortness of Breath Questionnaire

The SOBQ asks subjects to indicate severity of shortness of breath on a 6-point scale (0 = Not at all, …, 5 = Maximally or unable to do because of breathlessness) during 21 activities of daily living associated with varying levels of exertion. Three additional questions ask about fear of harm from overexertion, limitations, and fear caused by shortness of breath, for a total of 24 items. If patients do not routinely perform an activity, they are asked to estimate their anticipated shortness of breath. A total sum score ranges from 0 to 120. The SOBQ has been found to have excellent internal consistency, reliability, and moderate-to-strong correlations with measures of exercise tolerance, disease severity, depression, and perceived breathlessness ratings following a 6-minute walk test (6MW).[11,18]

Statistical Analysis

Data before and after rehabilitation were evaluated with descriptive statistics and paired *t* tests for changes. Change scores were calculated from pre- to post-rehabilitation. Pearson correlation coefficients were used to evaluate relationships among variables.

An MCID was evaluated by 3 methods: (1) effect size, (2) standard error of measurement (SEM), and (3) comparison of agreement between the SOBQ and the other 2 dyspnea measures (CRQ Dyspnea and the TDI). In addition, individuals experienced with the instrument were asked to estimate the MCID. The internal consistency of the SOBQ and other measures was calculated

Table 2 • BASELINE HEALTH-RELATED QUALITY-OF-LIFE VARIABLES

Variables	Mean ± SD	Floor (%)	Ceiling (%)	Cronbach α	SEM
SOBQ	55.5 ± 20.8	0	0	0.94	5.0
CRQ					
Dyspnea	3.7 ± 1.0	0	1%	0.71	0.52
Emotional	5.1 ± 1.2	0	2%	0.91	0.35
Fatigue	4.0 ± 1.2	0	1%	0.88	0.42
Mastery	4.9 ± 1.3	0	2%	0.77	0.60
RAND/SF-36					
Physical Component Summary Score	32.8 ± 8.3	0	0	0.92*	2.3
Mental Component Summary Score	51.2 ± 11.1	0	0	0.91*	3.3
Physical Functioning	35.3 ± 20.8	4%	0	0.86	7.6
Bodily Pain	75.3 ± 22.4	0	29%	0.83	10.0
Role-Physical	30.3 ± 36.3	46%	15%	0.83	14.0
General Health	43.9 ± 20.2	1%	0	0.71	11.6
Role-Emotional	69.5 ± 40.0	18%	58%	0.84	14.0
Mental Health	72.7 ± 18.3	1%	4%	0.81	9.0
Social Functioning	69.4 ± 27.0	1%	28%	0.78	14.6
Energy Fatigue	43.3 ± 20.4	2%	1%	0.85	9.5
QWB	0.626 ± 0.098	0	0	0.90*	0.031
CES-D depression	13.1 ± 9.4	2%	1%	0.85	3.5
Self-efficacy, walking	3.6 ± 2.9	14%	12%	N/A	

SOBQ indicates University of California, San Diego Shortness of Breath Questionnaire; CRQ, Chronic Respiratory Questionnaire; RAND/SF-36, RAND Corporation Health Survey (also known as SF-36); QWB, Quality of Well-Being Scale; CES-D, Center for Epidemiologic Studies Depression Questionnaire.
*Cronbach α is from the literature, not this population.

using Cronbach α. For effect size, changes before and after rehabilitation were divided by the standard deviation (SD) of the baseline score. The magnitude of the effect size was judged according to criteria described by Cohen (0.2 = small, 0.5 = moderate, 0.8 = large).[19]

The SEM was estimated according to the following equation: $SEM = \sigma\sqrt{1-r_{xx}}$, where σ = the SD and r_{xx} = its reliability coefficient. The criterion of one SEM was used to define MCID.[20] In order to validate the one SEM criterion, the SOBQ was categorized according to change after rehabilitation by change in SEM (improved = decrease ≥1 SEM; same = <1 SEM change; worse = gain ≥1 SEM) and was cross-classified with the CRQ Dyspnea scores according to previously established MCID.[1] A weighted-Kappa statistic was used to assess the rank agreement between the scores.

Comparison of the change in SOBQ and CRQ Dyspnea score versus the TDI was performed using a receiver-operator curve (ROC) analysis to evaluate the sensitivity (and specificity) of change in SOBQ (and CRQ Dyspnea) in determining a 1-unit change in TDI—by definition a change in symptoms that can be detected by patients.[21]

Prior to data analysis, 3 researchers/clinicians experienced in using the SOBQ were asked individually (without prior discussion) to estimate what change they considered to be clinically significant.

The predetermined α-level was 0.05 for all tests of significance. All analyses were performed using SPSS 12.0 and SAS Statistical Package.

RESULTS

Descriptive Statistics

The baseline psychometric properties of the HRQOL instruments are listed in Table 2. The SOBQ was found to be normally distributed. Most instruments had reliability scores (Cronbach α) greater than 0.80, except the CRQ's dyspnea, mastery, and RAND/SF-36's general health and social functioning domains. As reported in previous studies, about half of the RAND/SF-36 domains (Bodily pain, Role Physical, Role Emotional and Social Functioning) exhibited floor and ceiling effects: 25% or more of the subjects reported the highest or lowest possible score.[22,23] The HRQOL baseline scores are significantly lower than population norms.[24]

Correlations between baseline variables are displayed in Table 3. Correlations were highest between variables in the same categories (spirometry, exercise capacity, and HRQOL). Although HRQOL measures were not significantly correlated with spirometric variables (FEV_1 and FVC), associations with exercise

Table 3 • PEARSON CORRELATIONS BETWEEN BASELINE VARIABLES

	FVC %Pred	FEV_1 %Pred	DLCO %Pred	6MW	METS	SOBQ	CRQ Dys	CRQ EF	CRQ Fat	CRQ Mas	RAND PCS	RAND MCS	QWB
FEV_1 %Pred	0.69†												
DLCO %Pred	0.29†	0.60†											
6MW	0.31†	0.27*	0.42†										
METS	0.24†	0.34†	0.56†	0.67†									
SOBQ	−0.15	0.11	−0.20*	−0.48†	−0.49†								
CRQ Dys	0.04	−0.03	0.00	0.11	0.10	−0.47†							
CRQ EF	0.01	−0.08	−0.07	0.22†	0.11	−0.31†	0.23†						
CRQ Fat	0.04	−0.13	−0.18*	0.13	0.09	−0.42†	0.40†	0.51†					
CRQ Mas	0.14	0.02	−0.03	0.27†	0.25†	−0.44†	0.37†	0.66†	0.49†				
RAND PCS	0.17	0.0	−0.01	0.27†	0.28†	−0.53†	0.34†	0.14	0.52†	0.38†			
RAND MCS	0.04	0.0	−0.06	0.19*	0.12	−0.33†	0.18*†	0.68†	0.40†	0.44†	0.09		
QWB	0.06	−0.12	−0.10	0.19*	0.15	−0.40†	0.36†	0.42†	0.52†	0.38†	0.38†	0.30†	
CESD	−0.02	0.09	0.08	−0.25†	−0.19*	0.38†	−0.24†	−0.77†	−0.50†	−0.60†	−0.20*	−0.71†	−0.44†

DLCO indicates diffusing capacity for carbon monoxide; 6MW, 6-minute walk; METS, maximal treadmill workload as estimated oxygen consumption in metabolic equivalents; SOBQ, University of California, San Diego Shortness of Breath Questionnaire; CRQ, Chronic Respiratory Questionnaire: Domains—Dys = Dyspnea, EF = Emotional Function, Fat = Fatigue, Mas = Mastery; RAND, RAND Corporation Health Survery (also known as SF-36); PCS, Physical Component Summary score; MCS, Mental Component Summary score; QWB, Quality of Well-Being Scale; CES-D, Center for Epidemiologic Studies Depression Questionnaire.

*P value < .05.

†P value < .0001.

capacity (6MW and maximum treadmill workload [$METS_{max}$]) were significant. In particular, the SOBQ correlated strongly with 6MW and $METS_{max}$ (−0.48 and −0.49, respectively), higher than the other HRQOL measures.

Pulmonary Rehabilitation Effect

Changes after pulmonary rehabilitation are presented in Table 4. Although lung function did not change, HRQOL and exercise capacity improved significantly after the 8-week program. The statistically significant changes for the CRQ domains and QWB also met established criteria for clinical significance (0.5 units per question for the CRQ,[25,26] 0.03 for QWB[27]). Although the change in 6MW was statistically significant, it did not reach the generally defined clinically meaningful difference of 54 meters.[28]

Effect Size

Effect sizes are also shown in Table 4. The largest effect sizes were seen with the CRQ domains (range of 0.44 to 0.99). Moderate to large effect sizes (>0.5 and >0.8, respectively, as defined by Cohen[19]) were also seen for the SOBQ and several of the RAND/SF-36 subscales (Physical functioning, Role-physical, Energy/fatigue, and Health change). Most other effect sizes were in the small to moderate range (0.2 to 0.5).

Standard Error of Measurement

As suggested by Wyrwich et al,[29] we considered one SEM as a minimal clinically significant difference score. For the SOBQ, one SEM was 5.0 units (Table 2). This was consistent with the meaningful clinically significant change estimated by individuals familiar with the instrument. The one SEMs (per question) for the individual CRQ domains were as follows: dyspnea 0.52, emotional function 0.35, fatigue 0.42, and mastery 0.60 (Table 2). The average SEM for the 4 CRQ domains was 0.47. This is similar to the established 0.5 change for MCID suggested by Redelmeier.[26] The SEM for the QWB was 0.031, consistent with that instrument's established change score.[27]

SOBQ Versus CRQ Dyspnea

Change after rehabilitation in SOBQ correlated moderately with change in the CRQ Dyspnea domain (−0.43). Neither the SOBQ nor the CRQ Dyspnea change score was correlated with change in 6-minute walk distance or maximum treadmill workload ($METS_{max}$).

The SOBQ and CRQ Dyspnea scores were classified by change after pulmonary rehabilitation according to the already established MCID for CRQ Dyspnea and the suggested 5-unit change for SOBQ. Each subject's change was categorized as improved, worsened, or unchanged after rehabilitation. Weighted kappa of

Table 4 • PULMONARY REHABILITATION TREATMENT EFFECT AND EFFECT SIZE

Variable	Pre-Rehab	Post-Rehab	Effect Size
Maximum Treadmill			
Workload, METS	4.4 ± 2.2	5.4 ± 2.6§	−0.45
6-minute walk			
Distance, m‖	427 ± 105	450 ± 105§	−0.22
Perceived symptom score			
Breathlessness	4.5 ± 1.9	4.0 ± 1.7§	
Muscle fatigue	3.0 ± 2.2	2.6 ± 2.0‡	
Psychosocial measures			
Self-efficacy, walking	3.6 ± 2.9	4.6 ± 2.7§	−0.34
CES-D depression	13.2 ± 8.9	10.2 ± 7.7§	0.34
Score >18, n (%)	41 (25)	24 (15)	
Quality of life measures			
SOBQ	55.5 ± 20.8	45.5 ± 20.3§	0.48
BDI/TDI	5.0 ± 2.0	+2.7 ± 2.3§	NA
QWB	0.626 ± 0.098	0.657 ± 0.114§	−0.31
CRQ			
Dyspnea	3.7 ± 1.0	4.6 ± 1.1§	−0.99
Emotional	5.1 ± 12	5.6 ± 1.0§	−0.44
Fatigue	4.0 ± 1.2	4.9 ± 1.1§	−0.75
Mastery	4.9 ± 1.3	5.8 ± 1.0§	−0.66
RAND/SF-36			
Physical Component Summary Score	32.8 ± 8.3	36.3 ± 9.1§	−0.42
Mental Component Summary Score	51.2 ± 11.1	55.0 ± 8.6§	−0.34
Physical functioning	35.3 ± 20.8	47.1 ± 21.9§	−0.57
Role-physical	30.3 ± 36.3	47.9 ± 41.8§	−0.48
Role-emotional	69.5 ± 40.0	81.7 ± 33.3§	−0.30
Energy/fatigue	43.3 ± 20.4	56.5 ± 19.7§	−0.65
Emotional well-being	72.7 ± 18.3	78.8 ± 15.6§	−0.33
Social functioning	69.4 ± 27.0	78.4 ± 23.3§	−0.33
Pain	75.3 ± 22.4	77.6 ± 24.4	−0.10
General health	43.9 ± 20.2	47.7 ± 21.4‡	−0.19
Health change	41.6 ± 29.0	59.8 ± 31.2§	−0.63
Health status	5.5 ± 1.7	6.1 ± 1.8§	−0.35

Values are expressed as mean ± SD.

METS indicates maximal treadmill workload as estimated oxygen consumption in metabolic equivalents; CES-D, Center for Epidemiologic Studies Depression Questionnaire; SOBQ, University of California, San Diego Shortness of Breath Questionnaire; BDI, Baseline Dyspnea Index: TDI, Transition Dyspnea Index; QWB, Quality of Well-Being Scale; CRQ, Chronic Respiratory Questionnaire; RAND/SF-36, RAND Corporation Health Survey (also known as SF-36); NA, not applicable.

†$P < .05$.

‡$P < .01$.

§$P < .001$.

‖Mean change 23 m. Accepted MCID is 54 m.

agreement was 0.11 (0.06 asymptotic standard error), which confirms that these two instruments statistically had poor general agreement (perfect agreement is 1, no agreement is 0). Table 5 presents the cross-tabulation of changes comparing the SOBQ and CRQ Dyspnea Scale. One hundred three subjects improved according to the SOBQ versus 111 in CRQ Dyspnea; 77 (48%) improved in both. The two instruments agreed in 90 cases (56%). In 13 cases (8%), improvement was noted in the CRQ while the SOBQ worsened. In 2 (1%) cases, the SOBQ score improved when the CRQ Dyspnea score worsened.

Figure 1 presents results of the ROC analysis of changes in the SOBQ and CRQ Dyspnea versus the TDI after pulmonary rehabilitation.[10] Assuming that a threshold of 1-unit change in the TDI represents a noticeable change in symptoms and is a reasonable MCID,[21] this analysis suggests that a change in SOBQ of

Table 5 • AGREEMENT BETWEEN CRQ DYSPNEA AND SOBQ BY MCID IN 162 SUBJECTS

		SOBQ			
		Worse	Same	Improved	Total
CRQ Dyspnea	Worse	2 (1%)	2 (1%)	2 (1%)	6
	Same	10 (6%)	11 (7%)	24 (15%)	45
	Improved	13 (8%)	21 (13%)	77 (48%)	111
	Total	25	34	103	

CRQ indicates Chronic Respiratory Questionnaire; SOBQ, University of California, San Diego Shortness of Breath Questionnaire; MCID, minimal clinically important difference.

−5 units (improvement) is associated with a sensitivity and specificity of approximately 69% and 67%, respectively, in detecting a 1-unit improvement in TDI. Overall, as indicated in Figure 1, a change of 0.5 to 0.6 units in the CRQ Dyspnea domain was comparable to a 4- to 6-unit change in the SOBQ.

DISCUSSION

The results of this study suggest that a change in the SOBQ of 5 units is a reasonable estimate of the MCID for this questionnaire. This MCID was developed by systematic evaluation of several approaches in a patient cohort in response to a pulmonary rehabilitation intervention including comparison with similar instruments (anchor-based), distribution-based analysis, and a priori estimate by experienced users.

Users familiar with a test intuitively know through experience the relative value of a particular result. Statistically significant changes may not indicate clinically meaningful differences such as with small changes in large sample sizes. Unfortunately, there is no gold standard with which to establish MCID for a new instrument.

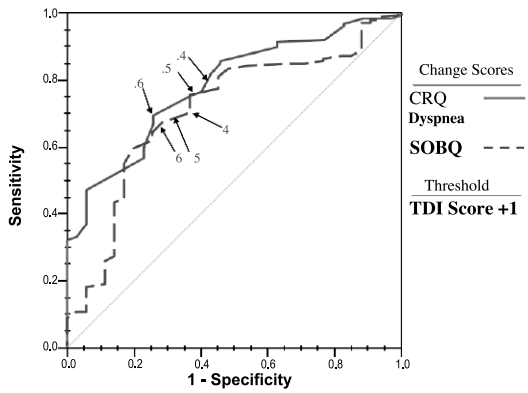

Figure 1. Receiver-operator curve (ROC) comparing changes in the SOBQ and CRQ Dyspnea questionnaires in detecting a 1-unit improvement in the Transition Dyspnea Index (TDI). Arrows indicate threshold values of 4-, 5-, and 6-unit improvement (decrease) for the SOBQ and 0.4-, 0.5-, and 0.6-unit improvement (increase) for the CRQ Dyspnea.

The complexities of methods used to evaluate responsiveness have been reviewed[30] and there does not appear to be a consensus on a single approach to arrive at an MCID. Wyrwich et al[20,29] proposed the one-SEM criterion for MCID. This was successfully cross-validated with the patient-driven MCID in previous studies using the CRQ. Similar findings in an anchor-based method (eg, global change on a 7-point scale) and distribution-based approach (effect size calculation) have been demonstrated.[31,32] Jaeschke et al[1] described a technique comparing a global measure of change (worse, about the same, or better using a 7-point scale) to the CRQ and arrived at an MCID score. Harper et al[23] also used a global question (modified version of item 2 of the SF-36) comparing a subject's health "now" to a specified earlier occasion as a standard for assessing responsiveness and MCID. Redelmeier et al[26] used a method requiring patients to judge themselves relative to others with the same condition and arrived at a similar MCID as the previous study with patients judging themselves according to their own memory approach. Norman et al,[33] though, asserted that global assessments of change can be associated with a recall bias that correlates highly with the present state of health but poorly with a previous one. Other methods include standardized effect sizes such as were calculated for the CRQ, SGRQ, SF-36, and Euroqol by dividing the mean change between assessments by the standard deviation of the change. Effect sizes varied with instrument, domain, and time interactions.[23] The SEM approach is a distributional approach that seems to be valid and yields similar results to other methods. The findings in this study using an SEM-based method yielded similar results for the CRQ and QWB to previously established thresholds.

The categorization of CRQ and SOBQ by significant change produced generally poor agreement, which highlights the difficulty in validating such measures. Although both questionnaires evaluate the broad concept of dyspnea, there are important differences between them. In this study, the CRQ Dyspnea scale was administered by an interviewer and was based upon specific activities selected by each individual. The SOBQ, on the other hand, asks subjects to report dyspnea associated with a standard set of activities. These discrepancies make it difficult to compare scores directly. Most cases of disagreement between the SOBQ and CRQ occurred when one instrument failed to detect change measured by the other. In 15 subjects, the two instruments detected changes in the opposite direction. In the majority of these (13), the CRQ Dyspnea score improved while the SOBQ score worsened, whereas the opposite occurred in only 2 subjects.

Ongoing evaluation of validity is important in developing an outcome measure as it is applied to an increasing variety of patient populations and situations. As with all new tests, these tools should be tested and compared to existing gold standard techniques. HRQOL questionnaires should meet these primary criteria: (1) *validity:* questionnaires should measure what they intend to measure (ie, scores reflect true domains to be evaluated); (2) *reliability:* the extent to which an instrument yields consistent, reproducible results; and (3) *responsiveness:* sensitivity of a questionnaire to measure change when such change has legitimately occurred (eg, after a treatment intervention). Researchers using these tools frequently gain an appreciation from their own experience as to what constitutes a meaningful change. Defining MCID allows those less familiar with the instrument to better interpret results.

The SOBQ, CRQ, and QWB had good psychometric properties in this analysis. These HRQOL measures are relatively easy to administer. They did not exhibit limiting floor or ceiling effects, and had acceptable reliability. Although the discriminative properties of the instruments were not formally tested, HRQOL levels in these patients with moderate to severe COPD were substantially lower than in healthy controls.[24] The RAND/SF-36 also had acceptable psychometrics, but as noted in previous studies, some of the individual domains were skewed at either end of the spectrum.[22,23] This nonnormal distribution might be expected with a generic measure not designed specifically for chronic lung disease. Further, floor or ceiling effects were noted for half of the RAND/SF-36 scales. The summary physical and mental component scores eliminate floor and ceiling effects and limit problems of multiple comparisons while still maintaining validity.[34] Further construct validity was evident with the expected changes in the appropriate direction with improvement after pulmonary rehabilitation.

Pulmonary rehabilitation has been well-established in the management of patients with chronic lung diseases. Benefits include improved exercise tolerance, symptoms, and quality of life with decrease in healthcare expenditures.[35–37] Pulmonary rehabilitation is an ideal intervention to evaluate HRQOL as the treatment and possible outcomes are multidimensional (eg, physiological, psychological, and social). Medical interventions often work through physiological changes to improve patient function. However, a treatment that improves lung function with no effect on symptoms or quality of life may be suspect. Conversely, an intervention with less impact on lung function, but more improvement in symptoms or quality of life, might be more important. Several studies have evaluated HRQOL in pulmonary rehabilitation with disease-specific or general instruments.[35,38–40] More recent studies have used both generic and disease-specific instruments in the same cohort.[10,23,41,42] Although results vary, as expected, generic measures have generally been less sensitive to change after pulmonary rehabilitation.[8,41,43]

In this study, responsiveness was demonstrated in all of the HRQOL measures after pulmonary rehabilitation, with the exception of the RAND/SF-36 Pain subscale. As expected, lung function did not change. This highlights the limitation of using classic physiologic indices as outcome measures in this setting. The responsiveness observed in this study adds to the overall validity of these instruments.

Determining the most sensitive instrument to detect change can be difficult. By design, disease or symptom-specific instruments, such as the CRQ and SOBQ, should be responsive and discriminative in COPD. Both were responsive to change after pulmonary rehabilitation with moderate effect sizes, larger with the CRQ than the SOBQ. The generic instruments were less responsive to change, with the QWB just achieving the estimated MCID.[41] The limitations in responsiveness for such instruments may be offset by gains in comparability between disease states and their use in cost-utility studies. There is also the added advantage that generic instruments may detect either beneficial or deleterious effects from treatment that might not be anticipated or measured with a disease-specific instrument. These findings support the notion that both disease-specific and generic instruments of HRQOL should be used together.

Comparison between different measures of HRQOL is always difficult because no gold standard truly exists. Differences between seemingly similar measures may well be due, in part, to the diverse components of HRQOL being assessed.

In summary, disease-specific and generic measures of HRQOL are important components of health outcome evaluations. Valid and responsive instruments exist for chronic lung disease and pulmonary rehabilitation. The proposed MCID of 5 units for the SOBQ compared favorably to thresholds established for other HRQOL instruments in this patient population under study.

Acknowledgment

The authors thank Cindy Morgan for her expert assistance with data analyses and the preparation of tables and figures for this manuscript.

References

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

2. Guyatt GH, Osoba D, Wu AW, Wyrwich KW, Norman GR, and the Clinical Significance Consensus Meeting Group. Methods to explain the clinical significance of health status measures. *Mayo Clin Proc*. 2002;77:371–383.
3. Make B, Casaburi R, Leidy NK. Interpreting results from clinical trials: understanding minimal clinically important differences in COPD outcomes. *COPD J Chronic Obstr Pulm Dis*. 2005;2:1–5.
4. Mahler DA, Matthay RA, Snyder PE, Pytlik L, Zaret BL, Loke J. Sustained-release theophylline reduces dyspnea in nonreversible obstructive airway disease. *Am Rev Respir Dis*. 1985;131:22–25.
5. van Schayck CP, Rutten-van Molken MP, van Doorslaer K, Folgering H, van Weel C. Two-year bronchodilator treatment in patients with mild airflow obstruction. *Chest*. 1992;102:1384–1391.
6. Tsukino M, Nishimura K, Ikeda A, Koyama H, Mishima M, Izumi T. Physiologic factors that determine the health-related quality of life in patients with COPD. *Chest*. 1996;110:896–903.
7. Jones PW. Issues concerning health-related quality of life in COPD. *Chest*. 1995;107(5 suppl):187–193.
8. 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;122:823–832.
9. Curtis JR, Martin DP, Martin TR. 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.
10. Ries AL, Kaplan RM, Myers R, Prewitt LM. Maintenance after pulmonary rehabilitation in chronic lung disease: a randomized trial. *Am J Respir Crit Care Med*. 2003;167:880–888.
11. Eakin EG, Resnikoff PM, Prewitt LM, Ries AL, Kaplan RM. Validation of a new dyspnea measure: the UCSD Shortness of Breath Questionnaire. *Chest*. 1998;113:619–624.
12. 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.
13. Mahler DA, Weinberger DH, Wells CK, Feinstein AR. The measurement of dyspnea: contents, interobserver agreement, and physiologic correlates of two new clinical indexes. *Chest*. 1984;85:751–758.
14. Hays RD, Sherbourne CD, Mazel RM. The RAND 36-Item Health Survey 1.0. *Health Econ*. 1993;2:217–227.
15. 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.
16. Kaplan RM, Anderson JP. A general health model policy: update and applications. *Health Serv Res*. 1988;23:203–235.
17. Radloff LS. The CES-D Scale: a self-report depression scale for research in the general population. *Appl Psychol Meas*. 1977;1:385–401.
18. Eakin EG, Sassi-Dambron DE, Ries AL, Kaplan RM. Reliability and validity of dyspnea measures in patients with obstructive lung disease. *Int J Behav Med*. 1995;2:118–134.
19. Cohen J. *Statistical Power Analysis for the Behavioral Sciences*. 2nd ed. Hillsdale, NJ: Lawrence Erlbaum Associates; 1988.
20. Wyrwich KW, Tierney WM, Wolinsky FD. Further evidence supporting an SEM-based criterion for identifying meaningful intra-individual changes in health-related quality of life. *J Clin Epidemiol*. 1999;52:861–873.
21. Mahler DA, Witek TJ Jr. The MCID of the Transition Dyspnea Index is a total score of one unit. *COPD J Chronic Obstr Pulm Dis*. 2005:99–103.
22. 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. *J Clin Epidemiol*. 1997;50:463–473.
23. 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.
24. McHorney CA, Ware JE Jr, Lu JF, Sherbourne CD. The MOS 36-item Short-Form Health Survey (SF-36): III. Tests of data quality, scaling assumptions, and reliability across diverse patient groups. *Med Care*. 1994;32:40–66.
25. Juniper EF, Guyatt 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.
26. 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.
27. Kaplan RM. The minimally clinically important difference in generic utility-based measures. *COPD J Chronic Obstr Pulm Dis*. 2005;2:91–97.
28. Redelmeier DA, Bayoumi AM, Goldstein RS, Guyatt GH. Interpreting small differences in functional status: the six minute walk test in chronic lung disease patients. *Am J Respir Crit Care Med*. 1997;155:1278–1282.
29. Wyrwich KW, Nienaber NA, Tierney WM, Wolinsky FD. Linking clinical relevance and statistical significance in evaluating intra-individual changes in health-related quality of life. *Med Care*. 1999;37:469–478.
30. Beaton DE, Bombardier C, Katz JN, Wright JG. A taxonomy for responsiveness. *J Clin Epidemiol*. 2001;54:1204–1217.
31. Norman GR, Sridhar FG, Guyatt GH, Walter SD. Relation of distribution- and anchor-based approaches in interpretation of changes in health-related quality of life. *Med Care*. 2001;39:1039–1047.
32. Samsa G. How should the minimum important difference for a health-related quality-of-life instrument be estimated. *Med Care*. 2001;39:1037–1038.
33. Norman GR, Stratford P, Regehr G. Methodological problems in the retrospective computation of responsiveness to change: the lesson of Cronbach. *J Clin Epidemiol*. 1997;50:869–879.
34. Ware JE Jr, Kosinski M, Bayliss MS, McHorney CA, Rogers WH, Raczek A. Comparison of methods for the scoring and statistical analysis of SF-36 health profile and summary measures: summary of results from the Medical Outcomes Study. *Med Care*. 1995;33(suppl 4):AS264–AS279.
35. ACCP-AACVPR Pulmonary Rehabilitation Guidelines Panel. Pulmonary rehabilitation: Joint ACCP/AACVPR evidence based guidelines. *Chest*. 1997;112:1363–1396.
36. Lacasse Y, Wong E, Guyatt GH, King D, Cook DJ, Goldstein RS. Meta-analysis of respiratory rehabilitation in chronic obstructive pulmonary disease. *Lancet*. 1996;348:1115–1119.
37. American Thoracic Society. Pulmonary rehabilitation—1999. *Am J Respir Crit Care Med*. 1999;159:1666–1682.
38. Goldstein RS, Gort EH, Stubbing D, Avendano M, Guyatt GH. Respiratory rehabilitation improved exercise tolerance and quality of life. *Lancet*. 1994;344:1394–1397.
39. Reardon J, Normandin E, Vale F, Clark B, ZuWallack R. The effect of comprehensive outpatient pulmonary rehabilitation on dyspnea. *Chest*. 1994;105:1046–1052.
40. 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.
41. Guyatt G, King D, Feeny D, Stubbing DG, 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.
42. Singh SJ, Sodergren SC, Hyland ME, Williams J, Morgan MDL. 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. Berry MJ, Rejeski WJ, Adair NE, Zaccaro D. Exercise rehabilitation and chronic obstructive pulmonary disease stage. *Am J Respir Crit Care Med*. 1999;160:1248–1253.