CHEST

Original Research

COPD

Patient-Reported Dyspnea in COPD Reliability and Association With Stage of Disease

Donald A. Mahler, MD, FCCP; Joseph Ward, RCPT; Laurie A. Waterman, BS; Corliss McCusker, RN; Richard ZuWallack, MD, FCCP; and John C. Baird, PhD

Background: **Although questionnaires are used frequently with patients to self-report the severity of dyspnea as related to activities of daily living, the reliability of these instruments has not been established. The two purposes of this study were to examine the test-retest reliability of three widely used dyspnea instruments and to compare dyspnea scores at different stages of disease.**
Methods: **At paired baseline visits, 101 stable patients with COPD were tested; at paired follow-up visits at 3 months, 89 of these patients were tested. At each visit, patients rated dyspnea with three instruments presented in random order and then performed post-bronchodilator therapy lung function tests.**
Results: **Patient-reported dyspnea scores and lung function were similar at baseline (interval, 6 ± 5 days) and follow-up visits (interval, 4 ± 2 days). Intraclass correlation coefficients at baseline and at follow-up were 0.82 and 0.82, respectively, for the modified Medical Research Council scale; 0.90 and 0.84, respectively, for the self-administered computerized versions of the baseline dyspnea index and transition dyspnea indexes; and 0.95 and 0.89 for the University of San Diego Shortness of Breath Questionnaire results. Dyspnea ratings were significantly related to the stage of disease severity based on percent predicted FEV_1 ($p < 0.001$).**
Conclusions: **Test-retest reliability was acceptable for patient-reported dyspnea scores using three clinical instruments at baseline and at the 3-month follow-up. Our results demonstrate for the first time that patient-reported dyspnea ratings are related to the stage of disease severity.**
(CHEST 2009; 136:1473–1479)

Abbreviations: BDI = baseline dyspnea index; IC = inspiratory capacity; ICC = intraclass correlation coefficient; mMRC = modified Medical Research Council; MRC = Medical Research Council; SAC = self-administered computerized; SOBQ = Shortness of Breath Questionnaire; TDI = transition dyspnea index; UCSD = University of California San Diego

Of 18 clinical practice guidelines for COPD[1–4] that have been published since 2000, 78% recommend that "symptoms/dyspnea" be monitored routinely in the care of patients with COPD. Only the monitoring of lung function was recommended more frequently (83%) in the guidelines.[1] A task force on outcomes for COPD pharmacologic trials[5] concluded that dyspnea, along with mortality and health-related quality of life, "remain the most important and robust clinical outcomes in COPD research."

Various questionnaires[6] are available that enable patients to report the impact of daily activities on their breathlessness. Of 33 assessments, Bausewein

Manuscript received April 16, 2009; revision accepted June 18, 2009.
Affiliations: From the Dartmouth-Hitchcock Medical Center (Drs. Mahler and Baird, Mr. Ward, and Ms. Waterman), Lebanon, NH; St. Francis Medical Center (Ms. McCusker and Dr. ZuWallack), Hartford, CT; and Psychological Applications, LLC (Dr. Baird), South Pomfret, VT.
Funding/Support: This work was supported by a National Institutes of Health grant no. R44 HL076888-02, Small Business Innovation Research (John C. Baird, principal investigator).
Correspondence to: Donald A. Mahler, MD, FCCP, Dartmouth-Hitchcock Medical Center, Medicine, CCM 3-D, One Medical Center Dr, Lebanon, NH 03756-0001; e-mail: donald.a.mahler@hitchcock.org

DOI: 10.1378/chest.09-0934

et al[7] found that the Medical Research Council (MRC) scale, the baseline dyspnea index (BDI), and the transition dyspnea index (TDI) were the most widely used in clinical trials. The MRC scale, along with the modified MRC (mMRC) scale, and the BDI can differentiate patients who have more dyspnea from those who have less dyspnea. The MRC scale and the TDI also have been used in clinical trials[8–12] to assess changes in dyspnea with therapy in patients with various respiratory diseases. Despite the recognized importance of dyspnea in patients with respiratory disease and the widespread use of the clinical instruments for patients to report breathlessness, the test-retest reliability of these dyspnea instruments has not been established. Test-retest reliability is an essential criterion of any measurement scale or instrument.[13,14]

The primary purpose of the present study was to investigate the test-retest reliability of the following three patient-reported measures of dyspnea: the mMRC scale[15]; the self-administered computerized (SAC) versions of the BDI and TDI[8,16]; and the University of California San Diego (UCSD) Shortness of Breath Questionnaire (SOBQ).[17] A secondary objective was to examine dyspnea scores based on the stage of COPD. Preliminary results of this investigation have been presented in abstract form.[18,19]

Materials and Methods

This cross-sectional study included a pair of baseline visits and a pair of follow-up visits 3 months later. Paired visits were scheduled 3 to 7 days apart. At the same time of day at each visit, patients completed the dyspnea instruments presented in random order and then performed pulmonary function tests. No changes in maintenance therapy for COPD were made throughout the study.

Subjects

A total of 101 patients were recruited from the outpatient clinics at Dartmouth-Hitchcock Medical Center (Lebanon, NH) [n = 62] and St. Francis Medical Center (Hartford, CT) [n = 39]. The diagnosis of COPD was based on standard criteria.[2] Other inclusion criteria were the ability to read and understand English and the presence of clinically stable disease.

Procedure

The institutional review board at each clinic approved the study, and each patient provided written informed consent.

mMRC scale

The patients read the 5-point mMRC scale presented on a piece of paper and circled the grade (0 to 4) that most closely matched his or her breathlessness.[15] Higher scores represent more breathlessness.

SAC Versions of the BDI and TDI

The SAC versions were presented on a desktop computer. For the BDI (visits 1 and 2), the patient selected grades for each of the three components, which were summed to obtain a total score (0 to 12).[8,16] Lower scores represent more breathlessness. Using the TDI (visits 3 and 4), patients reported changes in breathlessness from baseline for each component by adjusting the length of a bar along a bidirectional visual analog scale.[8,16] The three scores were summed and divided by 2 to obtain a total score (−9 to +9). A negative score indicates deterioration, whereas a positive score indicates improvement.

UCSD SOBQ

Patients circled a number on a 6-point scale to rate shortness of breath for each of 24 items.[17] The scores were summed to obtain a total score (0 to 120). Higher scores represent more breathlessness.

Lung Function

At each visit, the patient performed spirometry and inspiratory capacity (IC) maneuvers using standard equipment (Collins model CPL; Warren E. Collins; Braintree, MA) 20 min after the inhalation of two puffs (180 μg) of albuterol through a metered-dose inhaler. Predicted values for spirometry were taken from Morris et al[20] and were calculated for IC as the predicted total lung capacity minus predicted functional residual capacity from Crapo et al.[21]

Statistical Analysis

Data are presented as the mean ± SD. Paired *t* tests were used to compare results between test sessions, and the intraclass correlation coefficient (ICC) was used to evaluate test-retest reliability.[13,14] A sample size of 100 patients was considered adequate based on an expected ICC of ≥ 0.75 for each dyspnea instrument.[13] The Pearson product-moment correlation was used as a measure of the relatedness among variables. For nonparametric variables, such as the mMRC scale and the SAC BDI, Spearman rank correlation was calculated. Analysis of variance was used to compare results among the different stages of disease. *Post hoc* testing was performed to compare specific stages of disease, using Bonferroni correction.

Differences in dyspnea scores for the mMRC scale and the UCSD SOBQ at visits 3 and 4 were compared with baseline score by subtracting the values obtained at visit 1 for each patient. The SAC TDI is itself a "difference" score that is based on the patient's self-assessment of any change in dyspnea at visits 3 and 4 compared with visit 1.

Results

A total of 110 patients completed testing at visit 1. Nine of these patients did not return for visit 2. Eighty-nine of the initial 101 patients returned for follow-up testing (93 ± 6 days after visit 1) at both visits 3 and 4.

Baseline

The characteristics of patients at visits 1 and 2 are presented in Table 1. The mean interval between

Table 1—*Patient Characteristics at Visit 1 for Each Site and Outcomes at Visits 1 and 2 for all Subjects*

Characteristics	Visit 1 DHMC Patients (n = 62)	Visit 1 SFMC Patients (n = 39)	Visit 1 All Patients (n = 101)	Visit 2 All Patients (n = 101)	ICC (95% CI)
Gender, No.					
Female	32	20	52		
Male	30	19	49		
Age, yr	66 ± 10	67 ± 9	66 ± 9		
Height, cm	167 ± 9	168 ± 11	167 ± 9		
Weight, kg	76 ± 18	83 ± 26	78 ± 21		
FEV_1					
L	1.24 ± 0.41	1.32 ± 0.57	1.27 ± 0.47	1.29 ± 0.50	0.97 (0.95–0.98)
% predicted	52 ± 17	54 ± 16	53 ± 16	53 ± 16	0.96 (0.94–0.97)
FVC					
L	3.03 ± 0.81	2.62 ± 0.88	2.86 ± 0.85	2.93 ± 0.81	0.96 (0.93–0.97)
% predicted	87 ± 16	75 ± 18	82 ± 17	84 ± 16	0.92 (0.88–0.94)
IC					
L	2.06 ± 0.53	1.90 ± 0.65	1.99 ± 0.58	2.02 ± 0.60	0.95 (0.92–0.96)
% predicted	77 ± 19	75 ± 16	76 ± 19	76 ± 19	0.92 (0.89–0.95)
mMRC scale*	2.1 ± 1.0	1.8 ± 0.9	2.0 ± 0.9	1.9 ± 0.9	0.82 (0.74–0.88)
SAC BDI†	5.8 ± 2.2	6.8 ± 2.1	6.2 ± 2.2	6.3 ± 2.2	0.90 (0.85–0.93)
UCSD SOBQ‡	52.5 ± 21.4	46.2 ± 24.6	50.1 ± 23.1	50.6 ± 24.9	0.95 (0.93–0.97)

Values are presented as the mean ± SD unless otherwise indicated. Lung function values are postbronchodilator. $p > 0.05$ for paired *t* tests comparing results of all variables at visits 1 and 2. DHMC = Dartmouth-Hitchcock Medical Center; SFMC = St. Francis Medical Center.
*mMRC scale range, 0 to 4.
†SAC BDI range, 0 to 12.
‡UCSD SOBQ range, 0 to 120.

baseline visits was 6 ± 5 days. Anthropometric status, lung function, and dyspnea scores were similar between the two test sites; therefore, data were combined for the final analysis. As a group, the patients exhibited a wide spectrum of lung impairment (stage II, n = 56; stage III, n = 33; and stage IV, n = 12) and dyspnea associated with activities of daily living.

There were no significant differences for any variable measured at visits 1 and 2 ($p > 0.05$). The ICCs for all measures of lung function were ≥ 0.92 (Table 1). The ICC was 0.82 for the mMRC scale, 0.90 for the SAC BDI, and 0.95 for the UCSD SOBQ. These correlations were consistent across disease severity based on percent predicted FEV_1. Correlations among the dyspnea scores on the three instruments at visits 1 and 2 were −0.61 and −0.65, respectively, for the mMRC scale compared with the SAC BDI; −0.68 and −0.74, respectively, for SAC BDI compared with the UCSD SOBQ; and 0.52 and 0.71, respectively, for the mMRC scale compared with the UCSD SOBQ ($p < 0.001$ for all comparisons).

Among the three stages of COPD (stages II, III, and IV), the percent predicted values for IC ($p = 0.002$) were progressively lower, whereas patients reported more dyspnea ($p < 0.001$ for each instrument) with advanced stages of disease (Table 2). The distributions of dyspnea scores are displayed as tertiles in Figures 1, 2, and 3 for the different stages of disease severity.

Testing at 3 Months

Forty-six men and 43 women (mean age, 67 ± 8 years) completed testing at visits 3 and 4; the mean interval between visits was 4 ± 2 days. There were no significant changes in lung function and in dyspnea scores at visits 3 and 4 compared with baseline values (Table 3). The ICC values were ≥ 0.85 for all measures of lung function, 0.82 for the mMRC scale, 0.84 for the SAC TDI, and 0.89 for the UCSD SOBQ. Correlations among the difference scores for changes in dyspnea at visit 3 compared with visit 1 were −0.12 ($p = 0.28$) for the mMRC scale and the SAC TDI; −0.36 ($p = 0.001$) for the SAC TDI and UCSD SOBQ; and 0.26 ($p = 0.01$) for the mMRC scale and UCSD SOBQ.

Discussion

The major findings of this study in patients with COPD are that test-retest reliability is acceptable at baseline and for any changes in dyspnea at follow-up for all three instruments, and that patients with advanced stages of disease report more dyspnea

Table 2—*Patient Characteristics and Outcomes Based on Stage of Disease*

Variables	Stage II (n = 56)	Stage III (n = 33)	Stage IV (n = 12)	F Statistic (ANOVA)	95% CI	p Value
Gender						
Female	31	16	5			
Male	25	17	7			
FEV_1, L	1.52 ± 0.42*†	0.96 ± 0.22	0.68 ± 0.14	42.3		< 0.001
% predicted	62 ± 8*†	40 ± 5‡	25 ± 2	193.2	32–43* 19–27† 9–21‡	< 0.001
FVC, L	3.03 ± 0.80	2.75 ± 0.96	2.49 ± 0.75	2.3		0.106
% predicted	88 ± 14*†	78 ± 19	65 ± 14	10.8	10–35* 1–19†	< 0.001
IC, L	2.17 ± 0.58*†	1.80 ± 0.52	1.64 ± 0.40	6.9		0.002
% predicted	80 ± 19*	72 ± 14	64 ± 12	5.5	3–31*	0.005
mMRC scale	1.7 ± 0.9*	2.1 ± 0.9‡	3.0 ± 0.8	10.1	0.6–2* 0.2–2‡	< 0.001
SAC BDI	6.6 ± 2.2*	5.9 ± 2.0‡	4.0 ± 2.0	7.0	1–4* 0.1–4‡	< 0.001
UCSD SOBQ	42.3 ± 22.8*†	58.5 ± 16.0	71.8 ± 17.5	12.6	13–46* 5–28†	< 0.001

Values are from visit 1. All lung function values are postbronchodilator. The stages of disease are based on guideline recommendations.[2] ANOVA = analysis of variance; Stage II = FEV_1 50% to 79% predicted; Stage III = FEV_1 30% to 49% predicted; Stage IV = FEV_1 < 30% predicted.
*p < 0.05 between stages II and IV.
†p < 0.05 between stages II and III.
‡p < 0.05 between stages III and IV.

related to activities of daily living, whereas IC is lower. Test-retest reliability indicates the amount of measurement error and is considered a requisite quality for any instrument used in assessing outcomes.[13,14] To be valid, a measurement system also must be reliable.[22] Test-retest reliability is a specific type of reliability that can reflect observations or reports by patients on different occasions separated by an interval of time.[13] Although the original MRC scale was published in 1959,[23] to our knowledge, only one previous report has examined the reliability of questionnaires used to quantify breathlessness. In 1995, Eakin et al[24] reported the responses of 41 patients with either asthma or COPD who completed six different dyspnea questionnaires. Although the initial assessment was performed at the study site, the retest session was conducted during a telephone conversation several days later (mean interval, 2 days). The investigators found test-retest correlations of 0.72 for the mMRC scale, 0.76 for the

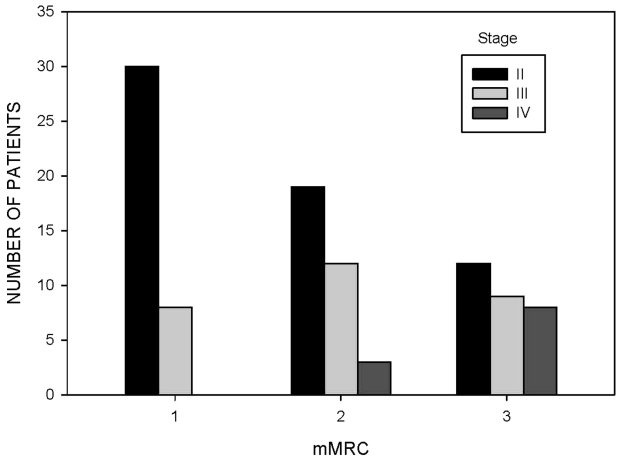

FIGURE 1. Distribution of scores for the mMRC scale in 101 patients with COPD at visit 1. Three categories for mMRC scores were selected based on similarities in activities that provoked breathlessness (category 1, 0 or 1; category 2, 2; category 3, 3 or 4).

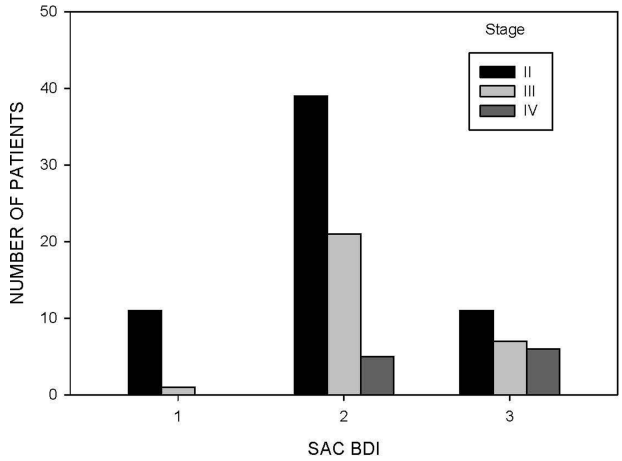

FIGURE 2. Distribution of scores for the SAC BDI in 101 patients with COPD at visit 1. Three categories were selected based on tertiles of possible SAC BDI scores (category 1, 9 to 12; category 2, 5 to 8; category 3, 1 to 4).

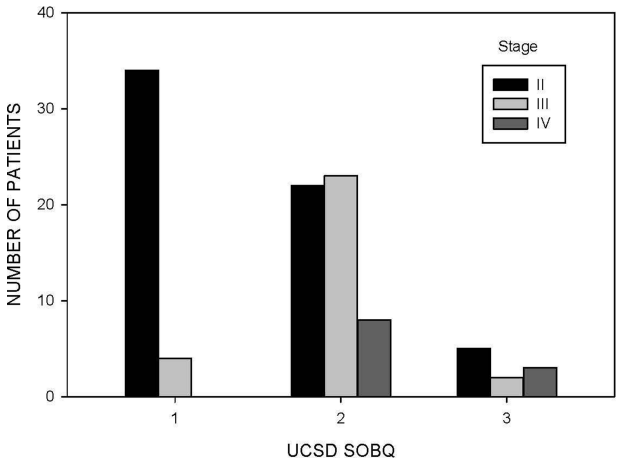

FIGURE 3. Distribution of scores for the UCSD SOBQ in 101 patients with COPD at visit 1. Three categories were selected based on tertiles of possible UCSD SOBQ scores (category 1, 0 to 40; category 2, 41 to 80; category 3, 81 to 120).

interviewer-administered BDI, and 0.94 for the UCSD SOBQ. The limitations of the study were that the instruments were not administered in random order, patients repeated the questionnaires at home with an investigator providing assistance over the telephone, and the stability of disease was not verified by repeat lung function results.

Our patients with COPD exhibited stable lung function at paired baseline visits separated by a mean interval of 6 ± 5 days. The test-retest reliability of the mMRC scale, the SAC BDI, and the UCSD SOBQ was acceptable based on an ICC value of ≥ 0.75.[13] The 95% CIs showed that the ICCs were robust. As expected in this observational study, lung function and dyspnea scores were stable over 3 months. At follow-up visits, patients selected response options for both the mMRC scale and the UCSD SOBQ based on their current status. To assess any changes in dyspnea with these instruments compared with baseline, we subtracted the scores at visit 1 from patient-reported values at visits 3 and 4 (Table 3). In contrast, for the SAC TDI, patients report any change in breathlessness using a bidirectional visual analog scale that compares their initial and baseline conditions.[8] Even with the different approaches used to obtain a measure of change in dyspnea over 3 months, the reliability of each of the three instruments was acceptable at paired follow-up visits (Table 3). These results are the first to examine the reliability of clinical instruments that assess changes in dyspnea related to activities of daily living.

Table 3—*Values for Lung Function and Changes in Dyspnea at Visits 3 and 4*

Variables	Visit 3	Visit 4	ICC (95% CI)
Gender, No.			
Female	46		
Male	43		
Age, yr	67 ± 8		
Height, cm	167 ± 9		
Weight, kg	79 ± 20		
FEV_1			
L	1.26 ± 0.50	1.27 ± 0.48	0.98 (0.97–0.99)
% predicted	53 ± 18	53 ± 17	0.95 (0.92–0.97)
FVC			
L	2.80 ± 0.82	2.85 ± 0.84	0.95 (0.92–0.97)
% predicted	82 ± 17	83 ± 16	0.85 (0.77–0.91)
IC			
L	1.98 ± 0.58	1.98 ± 0.62	0.96 (0.94–0.97)
% predicted	76 ± 20	76 ± 20	0.95 (0.92–0.97)
mMRC scale*	−0.04 ± 0.7	−0.09 ± 0.9	0.82 (0.72–0.88)
SAC TDI	−0.1 ± 2.5	−0.1 ± 1.9	0.84 (0.76–0.90)
UCSD SOBQ*	0.2 ± 14.3	−0.1 ± 14.9	0.89 (0.84–0.93)

Values are presented as the mean ± SD, unless otherwise indicated.
*Differences in dyspnea scores compared with values at visit 1. A negative value for dyspnea scores indicates worse dyspnea.

Most paradigms for classifying the severity of COPD use the percent predicted value for postbronchodilator FEV_1.[2–4] Although the threshold values selected for the different stages of disease were established by expert opinion, subsequent studies[25] have validated this approach. For example, disease severity in COPD patients based on spirometric classification is related to health status, utilization of health-care resources, development of exacerbations, and mortality.[3,26,27] Although cross-sectional studies[8,16,28,29] show modest correlations between measures of lung function and dyspnea ratings, previous investigators[27,30–32] have not shown significant differences in MRC dyspnea scores with different stages of COPD. Our results are the first to demonstrate that patient-reported dyspnea ratings on three different instruments are significantly related to the stage of disease severity based on percent predicted FEV_1 in patients with COPD.

Post hoc testing of dyspnea scores among patients with specific stages of disease demonstrated that differences were not consistent across all stages. For example, differences in scores for the mMRC scale and the SAC BDI were significant between stages II and IV and between stages III and IV, but not between stages II and III. For the UCSD SOBQ, significant differences in scores were observed between stages II and IV and between stages II and III, but not between stages III and IV. We believe that statistical differences would be evident for comparisons among all stages of disease severity with a larger patient population.

In addition, our patients demonstrated lower IC values with progressive airflow obstruction. IC is the inhaled volume of air from the end of exhalation to

total lung capacity and can provide an estimate of the end-exhalation lung volume. As airflow obstruction progresses, patients with COPD are unable to completely exhale the air from their lungs. This process results in an increase in end-expiratory lung volume (*ie*, hyperinflation) with a consequent decrease in IC. Hyperinflation at rest and with physical activity contributes to the breathlessness experienced by patients with COPD.[33] Our findings confirm the previous results of Di Marco et al.[30]

There are some limitations of our study. First, of the 101 patients, only 12 had stage IV disease. We attempted to recruit additional patients with very severe airflow obstruction (*ie*, $FEV_1 < 30\%$ predicted) to participate in this study. However, many patients with stage IV COPD declined to participate. These individuals typically explained that their breathlessness with activities made multiple return visits over 3 months difficult, illustrating the challenge of collecting data repeatedly in patients with advanced lung disease. The second limitation relates to the categorization of COPD severity according to percent predicted post-bronchodilator therapy FEV_1. Although the boundaries or cutoffs of the different stages of COPD were established by expert opinion, gradations in other outcomes provide support for this approach.[3,26] Third, to optimally assess the test-retest reliability of instruments that reflect changes in patient-reported dyspnea, a treatment or intervention would have been incorporated as part of a randomized clinical trial.

Although the classification of COPD severity is based on lung function, guideline recommendations[2,3] suggest that the selection of therapy should be based on the patient's symptoms and clinical presentation. Accordingly, the severity of dyspnea has been recommended[4,27] as an alternative or complementary approach to staging COPD. A panel representing the Canadian Thoracic Society[4] has suggested that a dual stratification system, which includes the severity of dyspnea (*ie*, disability) and the impairment of lung function, be used to categorize the stages of COPD. In addition, Celli et al[34,35] proposed that a multidimensional grading system, which includes BMI, percent predicted FEV_1, the mMRC scale, and 6-min walking distance, be used to assess the comprehensive severity of COPD. At the present time, dyspnea along with exercise performance, frequency of and time to an exacerbation, and health status are considered as important clinical outcomes that have been shown to improve with various treatments for patients with COPD.[5]

In summary, our results demonstrate acceptable and comparable test-retest reliability at baseline and 3-month follow-up for three widely used instruments that reflect the impact of dyspnea on activities of daily living in patients with COPD. These findings are clinically important because reliable instruments that quantify dyspnea are essential for achieving guideline recommendations that symptoms be monitored and that decisions on therapy be guided by the severity of symptoms.[2–4,25,36] We encourage further investigation into the development of a staging system for COPD that includes the severity of dyspnea.

Acknowledgments

Author contributions: Dr. Mahler was responsible for developing the research protocol, supervising data collection, reviewing the analysis, and preparing the manuscript. Dr. ZuWallack, Mr. Ward, Ms. Waterman, and Ms. McCusker were responsible for review and revision of the research protocol, data collection, review of the analysis, and review of the manuscript. Dr. Baird was responsible for developing the research protocol, performing the statistical analyses, and preparing the manuscript.

Financial/nonfinancial disclosures: Dr. Baird is the scientific director for Psychological Applications, which owns the copyright of the SAC versions of the BDI and TDI discussed in this article. Drs. Mahler and ZuWallack, Mr. Ward, Ms. Waterman, and Ms. McCusker have reported to the ACCP that no significant conflicts of interest exist with any companies/organizations whose products or services may be discussed in this article.

References

1 van den Bemt L, Schermer T, van Weel C. Rational monitoring of COPD: where do current clinical guidelines stand? Eur Respir J 2007; 29:1078–1081

2 Rabe KF, Hurd S, Anzueto A, et al. Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease: GOLD executive summary. Am J Respir Crit Care Med 2007; 176:532–555

3 Celli BR, MacNee W, ATS/ERS Task Force. Standards for the diagnosis and treatment of patients with COPD: a summary of the ATS/ERS position paper. Eur Respir J 2004; 23:932–946

4 O'Donnell DE, Aaron S, Bourbeau J, et al. Canadian Thoracic Society recommendations for management of chronic obstructive pulmonary disease: 2007 update. Can Respir J 2007; 14(suppl):5B–32B

5 Cazzola M, MacNee W, Martinez FJ, et al. Outcomes for COPD pharmacological trials: from lung function to biomarkers. Eur Respir J 2008; 31:416–469

6 Mahler DA. Dyspnea: mechanisms, measurement, and management. In: Mahler DA, O'Donnell DE, eds. Measurement of dyspnea: clinical ratings. Boca Raton, FL: Taylor & Francis, 2005; 147–165

7 Bausewein C, Farquhar M, Booth S, et al. Measurement of breathlessness in advanced disease: a systematic review. Respir Med 2007; 101:399–410

8 Mahler DA, Waterman LA, Ward J, et al. Validity and responsiveness of the self-administered computerized versions of the Baseline and Transition Dyspnea Indexes. Chest 2007; 132:1283–1290

9 de Torres JP, Pinto-Plata V, Ingenito E, et al. Power of outcome measurements to detect clinically significant changes in pulmonary rehabilitation of patients with COPD. Chest 2002; 121: 1092–1098

10 Kardos P, Wencker M, Glaab T, et al. Impact of salmeterol/fluticasone propionate versus salmeterol on exacerbations in

severe chronic obstructive pulmonary disease. Am J Respir Crit Care Med 2007; 175:144–149
11 Tashkin DP, Elashoff R, Clements PJ, et al. Cyclophosphamide versus placebo in scleroderma lung disease. N Engl J Med 2006; 354:2655–2666
12 King TE Jr, Behr J, Brown KK, et al. BUILD-1: a randomized placebo-controlled trial of bosentan in idiopathic pulmonary fibrosis. Am J Respir Crit Care Med 2008; 177:75–81
13 Streiner DL, Norman GR. Health measurement scales. In: Anonymous reliability. Oxford, UK: Oxford University Press, 1995; 104–127
14 Bland M. An introduction to medical statistics. In: Anonymous clinical measurement. Oxford, UK: Oxford University Press, 1995; 265–269
15 Brooks SM. Task group on surveillance for respiratory hazards in the occupational setting. ATS News 1982; 8:12–16
16 Mahler DA, Ward J, Fierro-Carrion G, et al. Development of self-administered versions of the modified baseline and transition dyspnea indexes in COPD. J COPD 2004; 1:165–172
17 Eakin EG, Resnikoff PM, Prewitt LM, et al. Validation of a new dyspnea measure: the UCSD Shortness of Breath Questionnaire; University of California, San Diego. Chest 1998; 113:619–624
18 Mahler DA, Ward J, Lentine TF, et al. Reliability of patient-reported dyspnea measures [abstract]. Am J Respir Crit Care Med 2008; 177:A522
19 Waterman LA, Ward J, McCusker C, et al. Reliability of patient-reported changes in dyspnea. [abstract]. Am J Respir Crit Care Med 2009; 179:A2912
20 Morris JF, Koski A, Johnson LC. Spirometric standards for healthy nonsmoking adults. Am Rev Respir Dis 1971; 103: 57–67
21 Crapo RO, Morris AH, Clayton PD, et al. Lung volumes in healthy nonsmoking adults. Bull Eur Physiopathol Respir 1982; 18:419–425
22 Lang TA, Secic M. How to report statistics in medicine. In: Anonymous observing exposures and outcomes concurrently. Philadelphia, PA: American College of Physicians, 2006; 245–246
23 Fletcher CM, Elmes PC, Fairbairn AS, et al. The significance of respiratory symptoms and the diagnosis of chronic bronchitis in a working population. BMJ 1959; 5147:257–266
24 Eakin EG, Sassi-Dambron DE, Ries AL, et al. Reliability and validity of dyspnea measures in patients with obstructive lung disease. Int J Behav Med 1995; 2:118–134
25 Jones PW, Agusti AG. Outcomes and markers in the assessment of chronic obstructive pulmonary disease. Eur Respir J 2006; 27:822–832
26 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
27 Bestall JC, Paul EA, Garrod R, et al. Usefulness of the Medical Research Council (MRC) dyspnoea scale as a measure of disability in patients with chronic obstructive pulmonary disease. Thorax 1999; 54:581–586
28 Mahler DA, Faryniarz K, Tomlinson D, et al. Impact of dyspnea and physiologic function on general health status in patients with chronic obstructive pulmonary disease. Chest 1992; 102:395–401
29 Wegner RE, Jorres RA, Kirsten DK, et al. Factor analysis of exercise capacity, dyspnoea ratings and lung function in patients with severe COPD. Eur Respir J 1994; 7:725–729
30 Di Marco F, Verga M, Reggente M, et al. Anxiety and depression in COPD patients: the roles of gender and disease severity. Respir Med 2006; 100:1767–1774
31 Cardoso F, Tufanin AT, Colucci M, et al. Replacement of the 6-min walk test with maximal oxygen consumption in the BODE Index applied to patients with COPD: an equivalency study. Chest 2007; 132:477–482
32 Nishimura K, Izumi T, Tsukino M, et al. Dyspnea is a better predictor of 5-year survival than airway obstruction in patients with COPD. Chest 2002; 121:1434–1440
33 O'Donnell DE, Banzett RB, Carrieri-Kohlman V, et al. Pathophysiology of dyspnea in chronic obstructive pulmonary disease: a roundtable. Proc Am Thorac Soc 2007; 4:145–168
34 Celli BR, Cote CG, Marin JM, et al. The body-mass index, airflow obstruction, dyspnea, and exercise capacity index in chronic obstructive pulmonary disease. N Engl J Med 2004; 350:1005–1012
35 Celli BR. Update on the management of COPD. Chest 2008; 133:1451–1462
36 Barnes PJ, Chowdhury B, Kharitonov SA, et al. Pulmonary biomarkers in chronic obstructive pulmonary disease. Am J Respir Crit Care Med 2006; 174:6–14