ORIGINAL ARTICLE

A Multidimensional Computer Adaptive Test Approach to Dyspnea Assessment

Anna Norweg, PhD, OTR, Pengsheng Ni, MD, MPH, Eric Garshick, MD, George O'Connor, MD, Kira Wilke, MPH, Alan M. Jette, PT, PhD

ABSTRACT. Norweg A, Ni P, Garshick E, O'Connor G, Wilke K, Jette AM. A multidimensional computer adaptive test approach to dyspnea assessment. Arch Phys Med Rehabil 2011; 92:1561-9.

Objectives: To develop and test a prototype dyspnea computer adaptive test (CAT).

Design: Prospective study.

Setting: Two outpatient medical facilities.

Participants: A convenience sample of adults (N=292) with chronic obstructive pulmonary disease (COPD).

Interventions: Not applicable.

Main Outcome Measure: We developed a modified and expanded item bank and CAT for the Dyspnea Management Questionnaire (DMQ), an outcome measure consisting of 4 dyspnea dimensions: dyspnea intensity, dyspnea anxiety, activity avoidance, and activity self-efficacy.

Results: Factor analyses supported a 4-dimensional model underlying the 71 DMQ items. The DMQ item bank achieved acceptable Rasch model fit statistics, good measurement breadth with minimal floor and ceiling effects, and evidence of high internal consistency reliability (α=.92–.98). With the use of CAT simulation analyses, the DMQ-CAT showed high measurement accuracy compared with the total item pool (r=.83–.97, P<.0001) and evidence of good to excellent concurrent validity (r=–.61 to –.80, P<.0001). All DMQ-CAT domains showed evidence for known-groups validity ($P\leq$.001).

Conclusions: The DMQ-CAT reliably and validly captured 4 distinct dyspnea domains. Multidimensional dyspnea assessment in COPD is needed to better measure the effectiveness of pharmacologic, pulmonary rehabilitation, and psychosocial interventions in not only alleviating the somatic sensation of dyspnea but also reducing dysfunctional emotions, cognitions, and behaviors associated with dyspnea, especially for anxious patients.

Key Words: Dyspnea; Outcome assessment (health care); Pulmonary disease, chronic obstructive; Rehabilitation; Reproducibility of results.

DYSPNEA IS A COMPLEX, multidimensional symptom with sensory, emotional, cognitive, and behavioral components.[1-5] The sensory component is the intensity and quality of the somatic sensation of labored, uncomfortable breathing.[3] The emotional dimension is the affective response to dyspnea, including such emotions as fear, distress, and anxiety.[4-8] The cognitive dimension includes perceptions and interpretations of dyspnea, which may be negative or catastrophic, and coping appraisal such as self-efficacy.[9,10] The behavioral dimension involves the avoidance of activities and hypervigilant behaviors related to dyspnea.[7,11] Because dyspnea is a prevalent yet modifiable symptom of chronic obstructive pulmonary disease (COPD), it is an important symptom to measure and target in health care settings.[12]

COPD research and practice have focused predominately on evaluating and treating the sensory component of dyspnea.[13,14] Scientists have only recently begun to develop multidimensional theoretical models of dyspnea and empirically test them.[13,15] For example, the multidimensional theoretical model of dyspnea proposed by Lansing et al[13] describes dyspnea as consisting of sensory and affective dimensions (immediate unpleasantness and cognitive evaluation/emotional responses). Initial evidence exists that sensory and affective dyspnea dimensions respond differently to treatment.[16]

Given the wide variation in dyspnea experiences[3] and high prevalence of anxiety in COPD,[12] a multidimensional approach to dyspnea assessment is needed to more adequately charac-

List of Abbreviations

CAT	computer adaptive test
CFA	confirmatory factor analysis
CFI	comparative fit index
COPD	chronic obstructive pulmonary disease
CRQ	Chronic Respiratory Disease Questionnaire
CSES	COPD Self-Efficacy Scale
DIF	differential item functioning
DMQ	Dyspnea Management Questionnaire
EFA	exploratory factor analysis
FEV_1	forced expiratory volume in 1 second
FVC	forced vital capacity
HADS	Hospital Anxiety and Depression Scale
IRT	item response theory
MIRT	multidimensional item response theory model
RMSEA	root mean square error of approximation
TLI	Tucker-Lewis Index
UCSD SOBQ	University of California, San Diego Shortness of Breath Questionnaire

From the Health & Disability Research Institute, Boston University School of Public Health, Boston, MA (Norweg, Ni, Wilke, Jette); Pulmonary and Critical Care Medicine Section, Medical Service, VA Boston Healthcare System and Channing Laboratory, Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, MA (Garshick); and Boston University School of Medicine, Boston Medical Center, Boston, MA (O'Connor).

Supported by the National Heart, Lung, and Blood Institute, National Institutes of Health (grant no. 1R21HL091237-01).

No commercial party having a direct financial interest in the results of the research supporting this article has or will confer a benefit on the authors or on any organization with which the authors are associated.

Correspondence to Anna Norweg, PhD, OTR, Visiting Researcher, Health & Disability Research Institute, Boston University School of Public Health, 715 Albany St-T5W, Boston, MA 02118, e-mail: *anorweg@bu.edu*. Reprints are not available from the authors. The DMQ-CAT is available on request.

0003-9993/11/9210-00371$36.00/0
doi:10.1016/j.apmr.2011.05.020

terize dyspnea for each individual with COPD.[13] Multidimensional assessment can identify patients with COPD who experience greater distress and anxiety associated with dyspnea, and therefore, improve therapeutic efficacy by optimizing the match and tailoring of treatment components specifically for these patients.[13,17] A multidimensional assessment approach in COPD could also increase and strengthen the available evidence of how to best minimize the disabling and distressing effects of dyspnea and promote dyspnea self-management[18] and adaptive coping.

Current dyspnea measures tend to focus too narrowly on measuring the sensory dimension of dyspnea, with inadequate measurement of its psychological and behavioral aspects. For example, the University of Cincinnati Dyspnea Questionnaire,[19] the University of California, San Diego Shortness of Breath Questionnaire (UCSD SOBQ),[20] and the Chronic Respiratory Disease Questionnaire (CRQ)[21] dyspnea scale all focus on measuring the sensory component of dyspnea. Single, discrete (categorical) dyspnea scales, such as the visual analog scale[6] and the British Medical Research Council scale,[22] while efficient, are not multidimensional and have high measurement error for comparing dyspnea change after treatment.[13,23] With a single categorical dyspnea scale, the rater cannot be certain of which characteristic of dyspnea is being measured: the sensory, psychological, or behavioral component of dyspnea.[13] While COPD quality-of-life questionnaires are multidimensional, such as the St. George's Respiratory Questionnaire[24] and the Seattle Obstructive Lung Disease Questionnaire,[25] they do not separate dyspnea from other symptoms,[3] and are therefore less sensitive in measuring dyspnea change after treatment.

One challenge to implementing a multidimensional dyspnea approach is administrative and patient burden. To include a sufficient number of items to adequately measure each dimension of dyspnea and to cover the wide spectrum of disability levels among adults with COPD, a traditional fixed-format multidimensional dyspnea instrument would be too long and time-consuming to administer and score. Computer adaptive test (CAT) and item response theory (IRT) techniques are currently being used to develop a new generation of health outcome instruments that enhance usability in a busy clinical or research setting.[26-29] CAT technology uses a simple form of artificial intelligence that selects questions based on a respondent's pattern of responses to previous questions.

This study applied IRT and CAT measurement methods to develop and test a prototype multidimensional dyspnea outcome measure for adults with COPD, the Dyspnea Management Questionnaire (DMQ).[30,31] The fixed-form DMQ consists of 5 theoretically derived dyspnea dimensions: dyspnea intensity, dyspnea anxiety, activity avoidance, activity self-efficacy, and satisfaction with strategy use. The DMQ was developed to measure patient-reported COPD treatment outcomes. Our specific aims were to (1) develop additional items to create an expanded calibrated item pool for the DMQ to improve on its breadth, precision, and conceptual clarity; (2) field test the expanded DMQ item bank to evaluate its dimensionality, scale properties, internal consistency reliability, and validity; (3) develop a prototype DMQ-CAT instrument; and (4) conduct preliminary testing of the accuracy, precision, and validity of the DMQ-CAT as compared with the full-item pool.

METHODS

Instrument

DMQ item bank development. The item pool of the original DMQ[31] (consisting of 56 items) was revised and expanded to a core set of 121 items based on 4 focus groups with multidisciplinary clinicians specializing in pulmonary medicine and people with COPD, 2 cognitive testing groups with adults with COPD, and a comprehensive review of the literature. We applied IRT to develop new and revised DMQ items. The dyspnea intensity items asked patients to rate how much dyspnea they had performing activities. The dyspnea anxiety items asked about emotions, autonomic arousal, and perceptions during breathing difficulty. The wording of the activity avoidance and self-efficacy item stems were changed (appendix 1). The DMQ response scales were changed from 7-point to 6-point Likert-type scales. The response choices for the activity avoidance scale, for example, ranged from "did not avoid it at all" (scored as 6) to "completely avoided it" (scored as 1). The satisfaction with strategy use domain included in the original DMQ[30,31] was not retained for the revised DMQ item pool.

Sample

The sample consisted of 292 adults with COPD from 2 medical centers. Patients were eligible if they had dyspnea with activities, physician-diagnosed COPD, a ratio of forced expiratory volume in 1 second (FEV_1) to forced vital capacity (FVC) of less than .70 (based on postbronchodilator FEV_1), were English speaking, and 40 years or older. Individuals with a neurologic disorder that affected their ability to move or perform daily activities were excluded. This study was approved by the institutional review boards of both cooperating facilities.

Data Collection

We mailed 945 recruitment letters and followed up by a phone call to potential participants to check their interest and to conduct a screening assessment. Participants were interviewed in-person by a trained interviewer in their home or in the clinic. A total of 292 interviews with usable data were conducted and included the DMQ item pool, the COPD Self-Efficacy Scale (CSES),[32] Hospital Anxiety and Depression Scale (HADS)–Anxiety subscale,[33] CRQ,[21] and UCSD SOBQ.[20]

DMQ Structure and Dimensionality

To test the dimensionality of the DMQ, we first used exploratory factor analysis (EFA) with all 121 items by using unweighted least squares estimates based on the polychoric correlation matrix. We then implemented separate confirmatory factor analysis (CFA) in each hypothesized subscale and trimmed the item pool by removing the high residual correlation items and attempting to satisfy the fit indices. Third, we did the EFA on the remaining 71 items to check that the hypothesized structure was maintained. We then compared 3 different IRT models with the 71 items: (1) 1-factor unidimensional, (2) 4-factor multidimensional IRT (MIRT), and (3) bifactor (orthogonal) MIRT by using both fit indices and the likelihood ratio test.[34] The goodness-of-fit indices used to evaluate model fit were the ratio of chi-square to degrees of freedom,[35] the comparative fit index (CFI),[36] the Tucker-Lewis Index (TLI), and the root mean square error of approximation (RMSEA).[37] All the analyses were conducted using MPlus software.[a]

Item Calibrations

We then calibrated the items based on the multidimensional Rasch model and differential item functioning (DIF) analysis[38-41] using Conquest computer software.[b] Item calibrations using a Rasch partial credit model estimated the level of difficulty of each item based on the sample's response pattern.