Reliability and validity of an Arabic version of the dyspnea-12 questionnaire for Saudi nationals with chronic obstructive pulmonary disease

Mohammed M. Alyami1,2, Sue C. Jenkins3,4, Hani Lababidi5, Kylie Hill1,4

Abstract:

BACKGROUND: Dyspnea is a distressing symptom experienced by people with chronic obstructive pulmonary disease (COPD). The dyspnea-12 (D-12) questionnaire comprises of 12 items and assesses the quality of this symptom, its severity and the emotional response. The original (English) version of the D-12 is reliable and valid for the measurement of dyspnea in pulmonary diseases.

AIM: To translate the D-12 into Arabic and determine whether this version is reliable and valid in Saudi nationals with COPD.

METHODS: The D-12 was translated into Arabic version and reviewed by an expert panel before being back-translated into English. The Arabic version was administered to five patients with COPD to test whether it was easily understood after which a final Arabic version was produced. Thereafter, 40 patients with COPD (aged 63 ± 9 years; 33 [82.5%] males; forced expiratory volume in one second (FEV1) 47 ± 16% predicted) completed the D-12, the COPD Assessment Test (CAT) and the Chronic Respiratory Disease Questionnaire (CRDQ). Lung function and 6-minute walk distance were also measured. The D-12 was re-administered two weeks later.

RESULTS: The Arabic version of the D-12 demonstrated good reliability over the two administration (intraclass correlation coefficient = 0.94, P < 0.01). Strong associations were demonstrated between the (1) total score for the D-12 and the CAT, (2) quality sub-score of the D-12 and the CAT and (3) emotional response sub-score of the D-12 and emotional function domain of the CRDQ (r ≥ 0.6, all P < 0.01).

CONCLUSION: The Arabic version of the D-12 is a reliable and valid instrument in Saudi nationals with COPD.

Key words: Arabic, chronic obstructive pulmonary disease (COPD), dyspnea, patient-reported outcome measures, questionnaire

For people with chronic obstructive pulmonary disease (COPD), dyspnea is often the most distressing symptom and the reason for seeking medical assistance.[10] Dyspnea refers to an uncomfortable and inappropriate sensation of breathing that results from the interaction of multiple physiological, psychological, social and environmental factors.[1] The sensation leads to activity limitation and a decrease in health-related quality of life (HRQoL).[2] Similar to pain, the sensation of dyspnea can be described in terms of its intensity, quality and the emotional response that it elicits.[3,4]

The assessment of dyspnea usually focuses exclusively on the severity of the sensation, and this is evaluated using uni-dimensional assessment tools, such as the modified Borg scale[5] and the visual analog scale.[4] However, there is now increasing recognition of the importance of the quality of the sensation.[7-9] Specifically, it appears that the quality of dyspnea varies between people with different cardiopulmonary diseases.[10] For example, the descriptor ‘chest tightness’ is most frequently used by people with asthma,[7,10,11] whereas the descriptors ‘increased work’ and ‘effort of breathing’ are more frequently used by people with COPD.[7,11,12] These differences suggest that the sensations have different pathophysiological origins. For example, in asthma, the sensation of chest tightness appears to arise from bronchoconstriction.[13] In contrast, in COPD, the sensations of ‘increased work’ and ‘effort of breathing’ appear to relate to the deleterious effects of pulmonary hyperinflation, which serves to increase the elastic and threshold loads that are borne by the inspiratory muscles, and thereby reduce their mechanical advantage.[14,15]

In addition to differences in the intensity and quality of the sensation, dyspnea often evokes considerable emotional distress, such as feelings of anxiety and panic.[16] These mood disturbances increase disability and functional impairment,
resulting in a further reduction in the HRQoL.\cite{27} Given the complexity of dyspnea, there is an interest in developing instruments that assess the intensity of dyspnea as well as its quality and emotional responses to this sensation. Recently, Yorke et al.\cite{18} developed a multi-dimensional instrument to quantify these three components of dyspnea, called the Dyspnea-12 questionnaire (D-12). The items for the D-12 were derived following a search of the literature pertaining to the language used to describe dyspnea by people with cardiorespiratory diseases. A pool of 81 items was generated that was then reduced to 12 items using hierarchical methods and Rasch analysis.\cite{18} In people with COPD, the D-12 questionnaire demonstrated good test–retest reliability (intraclass correlation coefficient [ICC] = 0.9, \(P < 0.001\)) and construct validity.\cite{18} Similar measurement properties for the D-12 have been demonstrated in people with asthma, interstitial lung disease (ILD) and pulmonary arterial hypertension (PAH).\cite{23-25,27}

Given the results of these earlier studies and the potential advantages of using a multi-dimensional tool to quantify dyspnea, the aim of this study was to develop an Arabic version of the D-12 and to establish its reliability and validity in quantifying dyspnea in Saudi nationals with COPD.

**Methods**

This study comprised of two parts (A and B). Part A involved translating the D-12 from English to Arabic and part B examined the reliability and validity of the Arabic version of the D-12. The study was approved by the Institutional Review Board at King Fahad Medical City (approval number 12/038) and the Human Research Ethics Committee of Curtin University (approval number HR 106/2012). Participants provided written, informed consent prior to data collection.

**Participants**

Participants for parts A and B were recruited from the outpatient pulmonary clinic at King Fahad Medical City in Riyadh, Saudi Arabia. To meet the inclusion criteria for both parts, participants needed a spirometric diagnosis of COPD, in accordance with the criteria established by the Global Initiative for Chronic Obstructive Lung Disease (GOLD).\cite{22} In addition, participants were required to be clinically stable (defined as no change in use of medication over the past four weeks) and to be able to read and write Arabic version. For part B, participants were excluded if they had a co-morbidity that limited mobility, such as severe musculoskeletal impairment and/or symptomatic cardiovascular disease.

**Protocol**

**Part A**

This part used a process similar to that described previously for translating questionnaires into a different language.\cite{23,24} First, the English version of the D-12 was translated into Arabic by two independent certified professional translators. Second, these two versions of the Arabic D-12 were reviewed by an expert panel that comprised of three respiratory physicians, one respiratory therapist, one respiratory health educator and one physical therapist. The panel compared the two translations with the original D-12 and suggested alternative words for any ambiguous items. A single version of the Arabic D-12 was produced via consensus. Third, the Arabic version of the D-12 was translated back to English by two independent certified translators. This version was compared to the original English version to confirm concordance. Finally, the Arabic version of the D-12 was presented to five participants with COPD to ensure that all items were easily understood.

**Part B**

This part was a cross-sectional study with measures collected during two assessment sessions, completed two weeks apart. During the first session, participants completed three questionnaires. Specifically, they completed Arabic versions of the D-12, the COPD Assessment Test (CAT)\cite{25,26} and the Chronic Respiratory Disease Questionnaire (CRDQ).\cite{27,28} Measures were collected of airflow obstruction and 6-minute walk distance (6MWD). Age, gender, weight and height were recorded. During the second assessment session, participants again completed the Arabic version of the D-12.

**Measures**

**Dyspnea**

Dyspnea was measured using the Arabic D-12, which comprises 12 items: 7 related to the quality of the sensation of dyspnea and 5 related to the emotional response to this sensation. Each item was graded in terms of its intensity using a 4-point scale, with higher scores indicating greater severity.\cite{14}

**Health status**

Health status was measured via the CAT. The CAT consists of eight items and addresses cough, phlegm, chest tightness, breathlessness going up hills/stairs, activity limitation at home, confidence leaving home, sleep and energy. Each item was scored on a 6-point scale, with higher scores indicating poorer health status.\cite{23}

**Disease-Specific health-related quality of life**

Health-related quality of life was measured using the self-administered, standardized version of the CRDQ. This questionnaire consists of 20 items grouped into four domains; dyspnea, fatigue, emotional function and mastery. Each item was scored on a 7-point scale, with higher scores indicating better HRQoL.\cite{27}

**Lung function**

Post-bronchodilator forced expiratory volume in one second (FEV\(_1\)) and forced vital capacity (FVC) were measured using an Easy One Spirometer (NDD Medical Technologies, Massachusetts, USA), according to the guidelines of the American Thoracic Society.\cite{29} Measures were expressed as a percentage of the predicted values previously established in a local population.\cite{30}

**Functional exercise capacity**

Participants performed the 6MWD on a 45-meter straight course within an enclosed corridor, in accordance with the American Thoracic Society guidelines.\cite{31} To account for improvements resulting from familiarization,\cite{32} two tests were conducted, separated by 30 minutes of rest, and the best distance was reported as the test result. The 6MWDs were expressed as a percentage of the predicted values previously established in an international sample.\cite{33}
Sample size

Sample-size calculations were conducted using published data.[18] Specifically, the weakest association reported between D-12 scores and clinical measures was with the FEV₁ (r = 0.30).[18] Therefore, to establish construct validity, we determined that a sample size of 44 participants with COPD was required to detect association of at least this strength between the D-12 and the clinical measures used in this study (α = 0.05, 1-β = 0.8).

Statistical analysis

Statistical analysis was performed using the Statistical Package for Social Sciences (SPSS version 21.0, Armonk, NY, IBM Corp). The distribution of the data was examined using the Shapiro–Wilk test. Test–retest reliability of the D-12 was assessed using an ICC.[34] The weighted kappa coefficient was used to determine the concordance of the responses between the first and second administrations of the D-12. Construct validity of the D-12 scores attained in the first administration was assessed by examining the strength of the associations between the D-12 and the CAT, CRDQ and 6MWD using Pearson’s correlation coefficients. Differences in the D-12 scores, with participants grouped according to their GOLD grades, were compared using analysis of variance (ANOVA). Data are presented as mean ± standard deviation (SD) unless otherwise stated. A probability (P) value less than 0.05 was used to denote statistical significance.

Results

Participants

Four (80%) males and one (20%) female with a mean age of 65 ± 10 years and a mean FEV₁ of 1.45 ± 0.61 L (51 ± 19% predicted) participated in part A. Forty participants (33 [82.5%] males) consented to take part in part B [Table 1]. No participants reported any changes in their symptoms, clinical status or medications use during the study period.

Part A

There were no differences between the back-translated version of the questionnaire and the original English version, demonstrating concordance between these two versions. All participants reported that all items of the Arabic version of the D-12 were easily understood. Figure 1 represents the Arabic version of D-12.

Part B

Reliability

The ICC for the reliability of the D-12 score was 0.94 (P = 0.01). Figure 2 illustrates the mean scores for each item of the D-12 over the two administrations. Excellent agreement in test scores was observed between the two administrations (weighted kappa = 0.83; P = 0.001).

Validity

Table 2 represents the associations between the D-12 total score and its sub-scores with the CAT, CRDQ domains, FEV₁, FVC and 6MWD. Strong associations (r ≥ 0.6) were demonstrated between (1) the total score for the D-12 and the CAT, (2) the quality sub-score of the D-12 and the CAT and (3) the emotional response sub-score of the D-12 and the emotional function domain of the CRDQ. The mean total scores of the D-12 differed between participants in GOLD grades 2 and 4 (mean difference = –4.6; P = 0.01). Figure 3 illustrates mean D-12 scores for participants grouped according to GOLD grades.

Discussion

The study successfully translated the D-12 to the Arabic language and demonstrated that the Arabic version of this questionnaire was reliable and valid in a population of Saudi nationals with COPD. This is consistent with the earlier work.

Table 1: Characteristics of the 40 participants who completed part B

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean±SD</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, yr</td>
<td>63±5</td>
<td>45-80</td>
</tr>
<tr>
<td>BMI, kg/m²</td>
<td>28.0±5</td>
<td>18.0-38.0</td>
</tr>
<tr>
<td>FEV₁, L</td>
<td>1.33±0.47</td>
<td>0.52-2.41</td>
</tr>
<tr>
<td>FEV₁, % predicted</td>
<td>47±15</td>
<td>19-78</td>
</tr>
<tr>
<td>FVC, L</td>
<td>2.53±0.60</td>
<td>1.27-3.52</td>
</tr>
<tr>
<td>FVC, % predicted</td>
<td>72±15</td>
<td>37-96</td>
</tr>
<tr>
<td>D-12 total score</td>
<td>22.4±3.6</td>
<td>14.0-28.0</td>
</tr>
<tr>
<td>Quality sub-score</td>
<td>13.6±2.7</td>
<td>7.0-18.0</td>
</tr>
<tr>
<td>Emotional response sub-score</td>
<td>8.9±1.8</td>
<td>5.0-13.0</td>
</tr>
<tr>
<td>CRDQ total score, ppi</td>
<td>3.7±1.0</td>
<td>1.9-5.5</td>
</tr>
<tr>
<td>Dyspnea, ppi</td>
<td>3.9±1.0</td>
<td>2.0-5.6</td>
</tr>
<tr>
<td>Fatigue, ppi</td>
<td>3.3±1.0</td>
<td>1.2-5.0</td>
</tr>
<tr>
<td>Emotional function, ppi</td>
<td>3.8±1.0</td>
<td>1.8-5.7</td>
</tr>
<tr>
<td>Mastery, ppi</td>
<td>3.7±1.1</td>
<td>1.0-6.7</td>
</tr>
<tr>
<td>CAT</td>
<td>19.1±4.8</td>
<td>8.0-27.0</td>
</tr>
<tr>
<td>6MWD, m</td>
<td>365±78</td>
<td>176-481</td>
</tr>
<tr>
<td>6MWD, % predicted</td>
<td>68±14</td>
<td>36-89</td>
</tr>
<tr>
<td>GOLD grades</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>18 (45)</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>15 (37.5)</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>7 (17.5)</td>
<td></td>
</tr>
</tbody>
</table>

Table 2: Associations between the D-12 Scores with CAT, CRDQ, FEV₁, FVC and 6MWD

<table>
<thead>
<tr>
<th>Measures</th>
<th>D-12 total</th>
<th>D-12 quality</th>
<th>D-12 emotional response</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAT</td>
<td>0.67*</td>
<td>0.61*</td>
<td>0.45*</td>
</tr>
<tr>
<td>CRDQ</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dyspnea, ppi</td>
<td>–0.56*</td>
<td>–0.55*</td>
<td>–0.31†</td>
</tr>
<tr>
<td>Fatigue, ppi</td>
<td>–0.24</td>
<td>–0.03</td>
<td>–0.43*</td>
</tr>
<tr>
<td>Emotional function, ppi</td>
<td>–0.39*</td>
<td>–0.07</td>
<td>–0.67*</td>
</tr>
<tr>
<td>Mastery, ppi</td>
<td>–0.38†</td>
<td>–0.24</td>
<td>–0.40†</td>
</tr>
<tr>
<td>Total, ppi</td>
<td>–0.50*</td>
<td>–0.28</td>
<td>–0.58*</td>
</tr>
<tr>
<td>FEV₁, L</td>
<td>–0.38*</td>
<td>–0.57*</td>
<td>–0.09</td>
</tr>
<tr>
<td>FVC, L</td>
<td>–0.21</td>
<td>–0.39†</td>
<td>–0.16</td>
</tr>
<tr>
<td>6MWD, m</td>
<td>–0.40*</td>
<td>–0.56*</td>
<td>–0.04</td>
</tr>
</tbody>
</table>

CAT: COPD assessment test, CRDQ: Chronic respiratory disease questionnaire, D-12 = Dyspnea-12, FEV₁ = Forced expiratory volume in one second, FVC = Forced vital capacity, GOLD = Global initiative for chronic obstructive lung disease, 6MWD = 6-minute walk distance.
Alyami, et al.: Arabic version of the dyspnea-12 questionnaire


that had translated questionnaires from English to Arabic language. Although the D-12 is likely to be useful in clinical practice for assessing dyspnea in terms of its intensity, quality and emotional responses, further work is required to determine whether this questionnaire is responsive to interventions aimed at reducing dyspnea, such as the use of bronchodilators and pulmonary rehabilitation.

Using the methodology previously described to translate the English version of the CRDQ and the CAT into different languages, we translated the D-12 into the Arabic language. The back-translation demonstrated good concordance with the original version and the Arabic version was easily understood. Similar results were found in a study that translated the D-12 into Korean.

The D-12 scores in this study were reliable over a 2-week period, as demonstrated by the high ICC value and the excellent weighted kappa score. It is noteworthy that the mean scores for 10 of the D-12 items were no different between the two administrations and the mean score for the remaining two D-12 items differed by less than 1 point. In the present study, the ICC value was similar to those values reported in earlier work in which the D-12 was administered twice over the same period in populations with COPD (r = 0.90), asthma (r = 0.93–0.96) and ILD (r = 0.94). The high ICC value indicates that the D-12 provides a stable measure of dyspnea, which serves to increase our confidence that any change in scores following an intervention is unlikely to just reflect variability in responses to the questionnaire items.

Regarding construct validity with questionnaire-based assessments, our data revealed a moderate to strong relationship between the D-12 total score and both the CAT score and the CRDQ total score. In addition, significant relationships were demonstrated with sub-scores of the D-12 and both the CAT and many of the domains of the CRDQ. Earlier work in patients with COPD used the Medical Research Council (MRC) dyspnea scale and the Hospital Anxiety and Depression Scale (HADS) to examine construct validity of the D-12. In contrast, we chose to use the CAT and the CRDQ to establish validity of the Arabic version of the D-12, as both have been validated in the Arabic language and comprise of items related to the impact of dyspnea on the daily life of patients. Despite the different questionnaires used, our data were consistent with a previous study in patients with COPD, revealing that the scores for D-12 were associated with dyspnea during daily life as well as emotional function. Further, the strong association between the D-12 total score and the CAT is
consistent with data in adults with asthma, ILD and PAH showing a strong relationship between the D-12 total score and a measure of health status. Our data confirm that in Saudi nationals with COPD the perception of dyspnea contributes significantly to health status and mood.

Associations of weak to moderate strength were demonstrated between the total score of the D-12 and both the 6MWD and the FEV₁. This is consistent with earlier work undertaken in people with COPD and asthma and suggests that dyspnea worsens with increased functional limitation and disease severity. It is noteworthy that the association between the D-12 total score and both the 6MWD and FEV₁ appears to be driven by the quality sub-score of the questionnaire rather than the emotional response sub-score. This indicates that as COPD increases in severity, and patients experience a greater intensity of various types of dyspnea sensation, but the emotional response shows a less consistent response pattern. This highlights the value of using a multi-dimensional tool to quantify dyspnea in this population.

There are limitations to this study. The participants were only recruited from one center and included a small proportion of female participants. In addition, all participants had a diagnosis of COPD. These factors may limit the generalizability of our results. Further studies are required to investigate the responsiveness of the Arabic version of the D-12 to therapeutic interventions, such as pulmonary rehabilitation.

In summary, the Arabic language version of the D-12 is a reliable and valid instrument for assessing dyspnea in Saudi nationals with COPD and it is likely to be a useful tool for assessing dyspnea in clinical practice.

Acknowledgement

The authors would like to thank all patients and staff at King Fahad Medical City who gave their time for this study.

References


Source of Support: This study was supported by the Ministry of Higher Education in Saudi Arabia, Conflict of Interest: None declared.