A Comparison of the Original Chronic Respiratory Questionnaire With a Standardized Version*

Holger J. Schünemann, MD, PhD; Lauren Griffith, MSc; Roman Jaeschke, MD, MSc; Roger Goldstein, MD, FCCP; David Stubbing, MD, FCCP†; Peggy Austin; and Gordon H. Guyatt, MD, MSc, FCCP

Background and study objectives: The chronic respiratory questionnaire (CRQ), a widely used measure of health-related quality of life (HRQL) in patients with chronic airflow limitation, includes an individualized dyspnea domain (patients identify five important activities, and report the degree of dyspnea on a 7-point scale). Because the individualized domain is unwieldy in multicenter clinical trials, we developed a standardized version and tested its discriminative and evaluative properties.

Methods: We enrolled 51 patients who completed the standardized and individualized CRQ before starting a respiratory rehabilitation program, and again 3 months later. We calculated both cross-sectional and longitudinal correlations between the two versions and a number of other HRQL instruments, and tested the relative ability of the individualized and standardized versions of the CRQ to detect improvement with rehabilitation.

Results: The results of the individualized questions suggested greater dysfunction (lower scores) than did the standardized questions both at baseline (3.18 vs 3.92, p < 0.001) and follow-up (4.62 vs 4.84, p = 0.051). The standardized dyspnea domain showed superior discriminative validity. While both techniques detected important, statistically significant improvement with rehabilitation (individualized domain mean change, 1.44; 95% confidence interval [CI], 1.11 to 1.77 [p < 0.001]; standardized domain mean change, 0.92; 95% CI, 0.61 to 1.24 [p < 0.01]), the difference in effect was substantial and statistically significant (mean difference, 0.52; 95% CI, 0.22 to 0.82; p = 0.001). The two versions showed comparable longitudinal validity.

Conclusions: A standardized version of the CRQ dyspnea domain improves the cross-sectional validity, maintains longitudinal validity, but reduces the responsiveness. By increasing sample size, investigators can use the more efficient standardized version of the CRQ without compromising validity.

(CHEST 2003; 124:1421–1429)

Key words: COPD; health-related quality of life; respiratory questionnaires; responsiveness

Abbreviations: CAL = chronic airflow limitation; CI = confidence interval; CRQ = chronic respiratory questionnaire; FT = feeling thermometer; HRQL = health-related quality of life; HUI3 = health utilities index 3; SF-36 = short form 36; SG = standard gamble; SGRQ = St. George Respiratory Questionnaire; SIP = sickness impact profile; TDI = transition dyspnea index

Health-related quality of life (HRQL) outcomes are gaining importance in clinical trials of patients with chronic airflow limitation (CAL). The chronic respiratory questionnaire (CRQ) is a widely used measure of HRQL in patients with CAL. The original CRQ is specific for patients with CAL...
including COPD. It uses 20 questions to sample four domains: dyspnea (5 questions), fatigue (4 questions), emotional functioning (7 questions), and mastery (4 questions). On all questions, patients rate their experience on a 7-point Likert-type scale ranging from 1 (maximum impairment) to 7 (no impairment). The questions on the dyspnea domain are individualized in that patients identify five important daily activities, and report their degree of dyspnea on those activities.

Although respondents are provided with a list of activities to choose from during completion of the CRQ dyspnea questions, the composition of chosen activities differs from patient to patient. Investigators have criticized the individualized CRQ on a number of counts. Tu et al.\(^7\) stated that the individualized dyspnea domain limits comparison across different settings, because individual patients chose different activities causing dyspnea. Other investigators\(^8\)–\(^10\) have criticized the individualized dyspnea domain because of the time required to elicit activities from participants, and difficulties with ensuring that patients perform activities regularly to allow evaluation at repeated administrations.

In theory, however, individualization of the dyspnea domain optimizes the evaluative properties of the CRQ by increasing the ability to detect change within patients or patient groups (responsiveness); therefore, individualized administration may avoid missing small but important treatment-mediated differences. While the original CRQ is both responsive and valid,\(^11\) the extent to which the individualized questions contribute, or detract from, this responsiveness and validity remains unknown.

Investigators may be interested in the ability of the CRQ dyspnea domain to distinguish between individuals with regard to their shortness of breath in daily activities at a point in time (its discriminative properties) and its ability to measure change in shortness of breath in daily activities over time (its evaluative properties). The aim of this study was to develop a standardized version of the CRQ and test its discriminative and evaluative properties against the individualized instrument. We conducted a study in which we administered the individualized and standardized CRQ before and after a 3-month respiratory rehabilitation program while administering other HRQL instruments used in patients with COPD. The rationale for administering several other disease-specific and generic validation instruments was the comparison of measurement properties of the dyspnea domain across a wide range of correlations.

**Materials and Methods**

**Patients and Data**

We obtained the data for this study from a trial conducted in the respiratory rehabilitation programs of two university centers (McMaster University, Hamilton, ON, Canada, and University of Toronto, ON, Canada).\(^12\),\(^13\) A trained research assistant interviewed patients on the day of entry (baseline) into one of the two standardized rehabilitation programs and after an intensive respiratory rehabilitation program at follow-up 3 months later. The trial had two primary aims. The first aim was the comparison of measurement properties of the informed (previous responses available) and the blind (previous responses not available) administration of both the CRQ and St. George Respiratory Questionnaire (SGRQ).\(^12\) The second aim was the evaluation of a visual analog scale (the feeling thermometer [FT]), as a measure of HRQL in patients with COPD.\(^13\)

We therefore had randomized 85 patients with CAL to the blind or informed CRQ and SGRQ and to the administration of the FT with or without rating of three marker states prior to rating of their own health state using a 2 \(\times\) 2 factorial design.\(^12\),\(^13\) In addition, patients completed a series of other HRQL instruments both at the baseline and at the visit following the standardized rehabilitation program. We accounted for possible small effects of order of administration by randomizing patients to complete the HRQL instruments. At follow-up, the same interviewer administered or supervised the administration of the instruments to each patient in the same order as at baseline. The SGRQ, health utilities index 3 (HU13), transition dyspnea index (TDI), sickness impact profile (SIP), and psychological adjustment to illness scale were self-administered and supervised by the interviewer; all other instruments were interviewer administered. For all instruments, we asked patients to rate their health as it had been in the prior 2 weeks. We excluded patients if they were unable to complete questionnaires due to language or cognitive limitations, or if they had a diagnosis of \(\alpha_1\)-antitrypsin deficiency, silicosis, sarcoidosis, asbestosis, lupus, or cancer.

**HRQL Instruments**

Because we became increasingly aware of the criticism related to the individualized administration of the CRQ after the original trial had commenced, we introduced a standardized set of questions about dyspnea after the initiation of that trial. Thus, of the 85 patients completing the original trial, only 51 patients completed the standardized dyspnea questions in addition to the individualized dyspnea questions.

We obtained the five activities for inclusion in the standardized dyspnea domain from previous trials.\(^14\),\(^15\) These activities were as follows: (1) feeling emotional, such as angry or upset; (2) taking care of basic needs (bathing, showering, eating, or dressing); (3) walking; (4) performing chores (such as housework, shopping, groceries); and (5) participating in social activities (see Appendix). The activities we chose were either those that patients who participated in these previous studies\(^14\),\(^15\) ranked most frequently as among the five important activities causing shortness of breath, or captured important aspects of emotional and social functioning (feeling emotional, such as angry or upset; and participating in social activities).

The individualized dyspnea domain lets patients choose five activities that are most important to them in their daily lives (see Appendix). Patients then rate the degree of dyspnea on these self-selected activities during subsequent administrations of the CRQ. Patients rate their experience on a 7-point scale ranging from 1 (maximum impairment) to 7 (no impairment). The standardized questions were scored in identical fashion, but each
patient was required to rate their dyspnea on the same five activities. We calculated the mean score for the individualized and standardized dyspnea domains by summing the scores for each dyspnea question and dividing it by the number of scored dyspnea questions (which ranged from two to five items).

All patients included in this analysis completed the standardized dyspnea questions in addition to the individualized dyspnea questions of the original CRQ. The order of administration of the individualized and standardized dyspnea questions at baseline was as follows: (1) patients chose the individualized activities from the original CRQ and ranked them according to importance, (2) patients rated their shortness of breath for the five standardized dyspnea questions, and (3) patients rated their shortness of breath for their individualized activities. At the follow-up visit, patients first rated their dyspnea on the five standardized activities and then on the individualized activities they had identified at the baseline visit. At this follow-up visit, for both standardized and individualized items, ratings included those activities for which the patients had provided a rating at the baseline visit.

In addition to the CRQ, we administered the following HRQL instruments: SGRQ, the Short-Form 36 (SF-36),16 standard gamble (SG),17 FT,17 HUI3,18 SIP,19 TDI (administered at follow-up),20 and global ratings of change.21 At the follow-up visit, the same interviewer administered or supervised the administration of the instruments to each patient in the same order as at baseline. The SGRQ, HUI3, TDI, and SIP were self-administered and supervised by the interviewer; all other instruments were interviewer administered.

Statistical Analysis

The primary focus of this study was the comparison of the standardized CRQ with the individualized CRQ. First, we were interested in the distribution of the scores of the dyspnea questions and the number of missing items. We then compared the mean scores of the standardized and individualized CRQ dyspnea questions. We tested for the presence of interaction between dyspnea domain method (individualized or standardized) and CRQ administration technique (blind or informed) and the format of FT administration (with or without hypothetical health states) using analysis of variance. We found no evidence of an interaction (p > 0.1).

Results

Fifty-one of the 85 patients who completed the trial completed both the individualized and standardized dyspnea questions. Table 1 presents the characteristics of the patients who are included in this analysis compared with the other patients in the original trial.12 There was no difference in gender, age, length and type of respiratory diagnosis, smoking history, and employment status in the patients who are included in this analysis compared with those not included in this analysis (the lowest p value for differences between these groups was 0.14).

Table 1—Demographic Information for Patients Completing the Individualized and Standardized Dyspnea Domain

<table>
<thead>
<tr>
<th>Patient Characteristics</th>
<th>Both Individualized and Standardized CRQ Dyspnea (n = 51)</th>
<th>Only Individualized CRQ Dyspnea (n = 34)</th>
<th>p Value for Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female gender, No. (%)</td>
<td>25 (49.0)</td>
<td>14 (41.2)</td>
<td>0.48</td>
</tr>
<tr>
<td>Age, yr (SD)</td>
<td>67.0 (7.0)</td>
<td>66.8 (8.5)</td>
<td>0.91</td>
</tr>
<tr>
<td>Diagnosis, No. (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COPD*</td>
<td>46 (92.0)</td>
<td>33 (96.6)</td>
<td>0.44</td>
</tr>
<tr>
<td>Other†</td>
<td>4 (8.0)</td>
<td>1 (3.4)</td>
<td>0.27</td>
</tr>
<tr>
<td>Time since diagnosis, yr (SD)</td>
<td>10.3 (9.5)</td>
<td>7.5 (4.9)</td>
<td>0.27</td>
</tr>
<tr>
<td>Smoking history, No. (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Never</td>
<td>1 (2.0)</td>
<td>4 (11.8)</td>
<td>0.14</td>
</tr>
<tr>
<td>Current smoker</td>
<td>0 (0.0)</td>
<td>1 (2.9)</td>
<td></td>
</tr>
<tr>
<td>Quit</td>
<td>50 (98.0)</td>
<td>29 (55.3)</td>
<td>0.56</td>
</tr>
<tr>
<td>Living alone, No. (%)</td>
<td>10 (19.6)</td>
<td>5 (14.7)</td>
<td>0.74</td>
</tr>
<tr>
<td>Employed, No. (%)</td>
<td>6 (11.8)</td>
<td>3 (8.8)</td>
<td></td>
</tr>
</tbody>
</table>

*Few patients had missing data for primary diagnosis on the data extraction sheet.
†Other diagnoses include idiopathic pulmonary fibrosis; chronic pulmonary aspergillosis; postpulmonary resection; and bronchiectasis.

Discriminative Properties: We evaluated cross-sectional validity of the standardized and the individualized dyspnea domain using data from the baseline and follow-up visit. For this evaluation, we calculated Pearson correlation coefficients of the individualized and standardized dyspnea questions with the other CRQ domains, the SGRQ, FT, SG, HUI3, SF-36, baseline dyspnea, and the SIP domains. Evaluative Properties: To assess the relative responsiveness of the standardized and individualized dyspnea domains, we calculated the mean change between baseline and follow-up scores and the corresponding 95% confidence limits for each method. To determine if any apparent differences in responsiveness could be explained by chance, we compared the difference of the change in CRQ scores between the standardized and individualized question groups using a paired t-test. To evaluate the longitudinal validity of the standardized and the individualized dyspnea domain, we calculated Pearson correlation coefficients of the change in scores of the two administration modes with the change in the validation instruments.
scores for the individualized questions were lower than for the standardized questions both at baseline ($-0.73; 95\%$ confidence interval [CI], $-0.99$ to $-0.48; p < 0.001$) and follow-up ($-0.22; 95\%$ CI, $-0.44$ to $0.00; p = 0.051$). Figure 3 shows that the number of missing responses is lower for the individualized than for the standardized dyspnea questions at baseline. Because patients only rated those activities on the individualized domain for which they had provided ratings at the baseline visit, we present the number of missing items only for the baseline visit. While almost 90\% of participants answered five of the individualized dyspnea questions, only approximately 50\% answered all five standardized questions ($p < 0.001$). Approximately 15\% of respondents answered to three standardized questions, and only 3\% of respondents answered to just two standardized questions.

Table 2 addresses the discriminative validity of the individualized and standardized CRQ dyspnea domain at baseline. In general, we observed moderate-to-strong correlations of both the individualized and standardized dyspnea domains with other CRQ domains and other HRQL instruments. All correlation

![Figure 1: Distribution of scores for the individualized and standardized dyspnea questions at baseline.](image1)

Compared to the standardized questions, the individualized items tended to produce a narrower distribution of scores and a lower mean score.

![Figure 2: Distribution of scores for the individualized and standardized dyspnea questions at follow-up.](image2)

* $p < 0.001$
coefficients were higher for the standardized domain compared to the individualized domain at baseline. For two comparisons, the difference was statistically significant (CRQ mastery and SF-36 mental score). Table 3 addresses the discriminative validity of the individualized and standardized CRQ dyspnea domain at follow-up. Similar to the baseline visit, we observed moderate-to-strong correlations of the individualized and standardized dyspnea domain with other CRQ domains and other HRQL instruments. As for the baseline visit, correlations were generally higher (11 of the 14 coefficients) for the standardized domain compared with the individualized domain, but only for the SGRQ symptoms domain was the difference statistically significant.

Evaluative Properties

Figure 4 shows the ability of the CRQ individualized and standardized dyspnea domains to detect improvement with the rehabilitation program (responsiveness). Both modes of administration demon-
strated large and statistically significant improvement over the course of rehabilitation; however, the mean change from baseline to follow-up was 0.52 (95% CI, 0.22 to 0.82; p < 0.001) greater for the individualized dyspnea domain, indicating greater responsiveness.

Table 4 addresses the correlations of the change in ratings between the baseline visit and follow-up visit of the individualized and standardized dyspnea CRQ domain with the other measures (longitudinal validity). Most of these correlations were moderate to strong and statistically significant. There was no clear trend for stronger correlations with the individualized or standardized dyspnea domain. Eight of 16 correlations were stronger for the standardized items.

**Discussion**

This study compared the use of standardized vs individualized CRQ dyspnea items. Because the CRQ is widely used in trials involving patients with COPD, our findings have important consequences for clinical studies. In general, standardized items showed consistently strong discriminative and evaluative properties, suggesting that the CRQ works well as a standardized instrument. Compared to the individualized approach, standardized dyspnea items showed stronger discriminative validity (Tables 2, 3); however, we observed similar longitudinal validity and superior responsiveness with the individualized questions (Fig 4).

Compared to the standardized questions, the individualized items tended to produce a narrower distribution of scores and a lower mean score (Figs 1, 2). It is likely that the opportunity to focus on their most important activities led patients to choose individualized items on which they experienced very troublesome dyspnea. The standardized approach does not permit this focus, resulting in items that may make some patients very short of breath, but cause milder degrees of dyspnea in others. Standard-
izing the dyspnea domain improves the discriminative properties of the CRQ dyspnea domain.

These differences in mean and distribution are associated with important differences in measurement properties. The wider distributions may have contributed to higher correlations between standardized items and other HRQL instruments than between individualized items and those same instruments (Tables 2, 3). These higher correlations suggest improved discriminative validity of the standardized approach, an important measurement property for differentiating severity of HRQL impairment between patients.

At the same time, when patients start with a lower mean score at baseline, an intervention can potentially lead to greater improvement and thus greater responsiveness. Several authors3,7–10 have suggested that the individualization of dyspnea items as an important limitation of the CRQ, sometimes without acknowledging the possible benefits in terms of measurement properties. As it turns out, our results show significantly larger change scores, and lower p values, in the individualized compared to the standardized domains. The greater responsiveness of the individualized approach was both substantial, and statistically significant (Table 4).

What is the significance of this difference in responsiveness? In terms of effect size, standardization of the dyspnea questions reduced the effect size by approximately 36% (mean change of 0.92 on the standardized domain divided by a mean change of 1.44 on the individualized domain with a SD of 1.0). Maintaining a power of 0.8 at an α of 0.05, this reduction in responsiveness would lead to an approximate 2.5-fold increase in sample size in a trial comparing a new treatment to a control group with the dyspnea domain as primary outcome measure. For example, to show an improvement of 0.5 (SD of 1) using the standardized questions, would increase the sample size from 63 to 154 per group.

The impact of this decreased responsiveness may be important for clinical trials in which limitation of patient numbers potentially compromises the power of the study. In such studies, investigators should consider using the individualized CRQ dyspnea domain. At the same time, changing to standardized items did not attenuate the correlation between change in the CRQ dyspnea domain and change in other HRQL measures. Thus, if increases in sample size are feasible, and investigators feel the greater convenience of the standardized items is important in their study, they can choose the standardized approach without fearing they are compromising the validity of the CRQ as a measure of change over time. Although we did not measure the time needed for administration of the individualized and the standardized version of the CRQ in this study, we have done so as part of another study (unpublished data). The mean duration for administration of the individualized CRQ without the standardized activity questions was 16 min (SD 6 min) at baseline and 10 min (SD 4 min) at follow-up. For the CRQ using the standardized dyspnea domain, the required time was 8 min (SD 3 min) both at baseline and follow-up.

We found that the standardized approach led to a larger number of missing items. This finding is not surprising because in choosing individualized items, patients are instructed to select items that are both important to them and that they have undertaken in the last week. No such restrictions of either importance or recent performance apply to the standardized items, making it far more likely that respondents will not have undertaken the activity during the time frame specified in the questionnaire. At the same time, patients almost invariably completed at least four standardized dyspnea items (Fig 3). Given prior results suggesting that reducing the number of items from five to two has only a small impact on responsiveness and validity,28 it is not surprising that imputing results on the basis of the items that were completed maintained the measurement properties of the standardized CRQ in terms of validity. It is possible, however, that the missing items contributed to the decreased responsiveness of the standardized approach.

Strengths to the present study include the use of multiple other questionnaires allowing us to thoroughly explore the relative validity of the two formats of measuring the impact of dyspnea on HRQL. We elicited the individualized dyspnea items prior to administering standardized questions. As a result, it is not possible that the standardized item could have influenced the patients’ choice of their most important activities. In addition, administration of the other instruments was highly standardized. Finally, the use of an intervention with known efficacy in HRQOL ensures that we could compare responsiveness between the individualized and standardized techniques.

Our study is limited in that we included only 51 of the 85 patients enrolled in the original trial; however, the included patients were similar to those who were enrolled early in the trial (Table 1) and, thus, selection is unlikely to have influenced the results. Another limitation is that we cannot exclude order effects: responses to the individualized questions might have influenced responses to standardized items. A parallel-group, randomized controlled trial in which patients completed only one version of the dyspnea questions would be required to definitively exclude the possible impact of one approach on the other.
In summary, our study suggests that use of a standardized version of the CRQ dyspnea domain improves discriminative properties of the CRQ in terms of cross-sectional validity, but reduces responsiveness. By increasing sample size, investigators can use the more efficient standardized version of the CRQ without compromising validity or statistical power.

APPENDIX

Standardized Dyspnea Questions (Read by Interviewer)

The CRQ is owned by Gordon H. Guyatt, MD, and is protected by copyright. A license agreement is necessary to use the questionnaire.

I would now like you to describe how much shortness of breath you have experienced during the last 2 weeks while doing the five most important activities you have selected. Please indicate how much shortness of breath you have had during the last 2 weeks while (interviewer will insert activities 1 through 5 from list 1—the question is repeated five times, once for each activity) by choosing one of the following options from the card in front of you:

1. Extremely short of breath
2. Very short of breath
3. Quite a bit short of breath
4. Moderate shortness of breath
5. Some shortness of breath
6. A little shortness of breath
7. Not at all short of breath
Not done

List 1

1. Feeling emotional such as angry or upset
2. Taking care of your basic needs (bathing, showering, eating, or dressing)
3. Walking
4. Performing chores (such as housework, shopping, groceries)
5. Participating in social activities

Individualized Dyspnea Questions (Read by Interviewer)

I would like you to think of the activities that you have done during the last 2 weeks that have made you feel short of breath. These should be activities that you do frequently and that are important to your day-to-day life. Please list as many activities as you can that you have done during the last 2 weeks that have made you feel short of breath.

Can you think of any other activities you have done during the last 2 weeks that have made you feel short of breath? I will now read a list of activities that make some people with lung problems feel short of breath. I will pause after each item long enough for you to tell me if you have felt short of breath doing that activity during the last 2 weeks. If you have not done the activity during the last 2 weeks, just answer “NO.” The activities are as follows:

List 2

1. Being angry or upset
2. Having a bath or shower
3. Bending
4. Carrying, such as carrying groceries
5. Dressing
6. Eating
7. Going for a walk
8. Doing your housework
9. Hurrying
10. Making a bed
11. Mopping or scrubbing the floor
12. Moving furniture
13. Playing with children or grandchildren
14. Playing sports
15. Reaching over your head
16. Running, such as for a bus
17. Shopping
18. While trying to sleep
19. Talking
20. Vacuuming
21. Walking around your own home
22. Walking uphill
23. Walking up stairs
24. Walking with others on level ground
25. Preparing meals

Of the items you have listed, which is the most important to you in your day-to-day life? I will read through the items. When I am finished, I would like you to tell me which is the most important (interviewer will read through all items spontaneously volunteered, and those on the list that the patient mentioned). Which of these items is most important to you in your day-to-day life? (Interviewer will list items on response sheet.) Of the remaining items, which is the most important to you in your day-to-day life? I will read through the items. When I am finished, I would like you to tell me which is the most important (interviewer will read through the remaining items). Which of these items is most important to you in your day-to-day life? (Interviewer will list items on the response sheet. This question is repeated three more times.)

I would now like you to describe how much shortness of breath you have experienced during the last 2 weeks while doing the five most important activities you have selected. Please indicate how much shortness of breath you have had during the last 2 weeks while (interviewer will insert activities 1 through 5 from the response sheet) by choosing one of the following options from the card in front of you:

1. Extremely short of breath
2. Very short of breath
3. Quite a bit short of breath
4. Moderate shortness of breath
5. Some shortness of breath
6. A little shortness of breath
7. Not at all short of breath

ACKNOWLEDGMENT: Dr. Stubbings died after submission of the first version of this manuscript. We have lost a wonderful colleague.

REFERENCES

1 van den Boom G, Rutten-van Molken MP, Molema J, et al. The cost effectiveness of early treatment with fluticasone propionate 250 microg twice a day in subjects with obstructive airway disease: results of the DIMCA program. Am J Respir Crit Care Med 2001; 164:2057-2066
22 Moran LA, Guyatt GH, Norman GR. Establishing the minimal number of items for a responsive, valid, health-related quality of life instrument. J Clin Epidemiol 2001; 54:571-579