The Measurement of Dyspnea*

Contents, Interobserver Agreement, and Physiologic Correlates of Two New Clinical Indexes

Donald A. Mahler, M.D., F.C.C.P.; David H. Weinberg, M.D.; Carolyn K. Wells, M.P.H.; and Alvan R. Feinstein, M.D.

To improve the clinical measurement of dyspnea, we developed a baseline dyspnea index that rated the severity of dyspnea at a single state and a transition dyspnea index that denoted changes from that baseline. The scores in both indexes depend on ratings for three different categories: functional impairment; magnitude of task, and magnitude of effort. At the baseline state, dyspnea was rated in five grades from 0 (severe) to 4 (unimpaired) for each category. The ratings for each of the three categories were added to form a baseline focal score (range, 0 to 12). At the transition period, changes in dyspnea were rated by seven grades, ranging from −3 (major deterioration), to +3 (major improvement). The ratings for each of the three categories were added to form a transition focal score (range, −9 to +9). In 38 patients tested with respiratory disease, interobserver agreement was highly satisfactory for both indexes. The baseline focal score had the highest correlation (r = 0.60; P < 0.001) with the 12-minute walking distance (12 MW), while significant, but lower, correlations existed for lung function. For the transition focal score, there was a significant correlation only with the 12 MW (r = 0.33; p = 0.04). These results indicate that dyspnea can receive a direct clinical rating that provides important information not disclosed by customary physiologic tests.

Dyspnea, or "pathologic" breathlessness, has been defined as an uncomfortable awareness of breathing or an increased respiratory effort that is unpleasant and regarded as inappropriate by the patient. Despite the frequency with which the symptom is encountered in clinical practice, dyspnea has been difficult to evaluate because it is a subjective sensation, and its apparent severity may or may not correlate with physiologic measurements.

Both psychophysical methods and clinical scales have been used to assess breathlessness. Psychophysical testing has involved the measurement of perception of changes in breathing in response to externally added loads. Although this approach has led to greater understanding of respiratory sensations, several factors, including technical aspects and time requirements, limit its application in the care of patients.

Clinical methods used to measure dyspnea have depended primarily on the magnitude of the exertional task that evokes breathlessness. In 1959, Fletcher and colleagues prepared a five-point rating scale based on the patient's history of developing dyspnea while walking distances on the level or climbing stairs. Seven years later, the Medical Research Council (MRC) of Great Britain proposed a similar four-point scale, which has been used for epidemiologic studies for observing the natural history of a disease, for diagnostic evaluation, and in clinical trials. Both of these scales relate primarily to the magnitude of the task that provokes dyspnea, but there is little provision for the associated effort. This consideration is important, since the development of dyspnea during a specific task, such as stair climbing, may vary with the speed of walking. For example, a patient can achieve substantial improvement or decline in exertional capacity by changing the effort with which the task is performed.

A separate disadvantage of these scales is that functional impairment, an important consequence of dyspnea, is not considered. This component should also be included in the evaluation of dyspnea, since a modest reduction in exertional capacity may create no difficulties or symptoms in the sedentary life of a retired, elderly person, but may be disastrous for a younger person who has major occupational demands.

There are additional difficulties with a scale that contains only four or five arbitrarily demarcated grades. When changes are evaluated, the scale may be too coarse to show distinct alterations that may occur. This problem can be circumvented by using the visual analog technique, which consists of a measured line with descriptive phrases at various points along the line. The patient is instructed to place a mark on the line corresponding to the severity of his symptoms.

*From the Departments of Medicine and Epidemiology, Yale University School of Medicine, New Haven; and the Veterans Administration Medical Center, West Haven, Connecticut. Supported by the Medical Research Service, Veterans Administration, The Commonwealth Fund, and the Robert Wood Johnson Foundation grant 6308.

Manuscript received August 10; revision accepted November 29.
Reprint requests: Dr. Mahler, Pulmonary Section/Medicine, Dartmouth-Hitchcock Medical Center, Hanover, New Hampshire 03755.
Since the line is a measured length (usually 100 mm), the location of the mark provides a quantification of the patient's dyspnea. The visual analog scale (VAS) has been used for an "oxygen-cost diagram" in which daily activities are ranked along the 100-mm line in proportion to their associated oxygen cost.2 The patient marks the line at a point above which a task would have to be stopped because of breathlessness.

Although the VAS can provide a dimensional measurement of the severity of dyspnea, it is a "global" rating that does not consider the factors contributing to breathlessness. Furthermore, there are no apparent criteria or standard principles that would allow the scale to be used consistently by different observers. Serial measurements of dyspnea using the VAS might be effective for showing change within a single patient, but this method would not be appropriate for comparing dyspnea in different patients nor would it be satisfactory for summarizing or comparing the conditions of groups of patients.

To improve the clinical rating of dyspnea, we developed two new indexes that include the components of functional impairment and magnitude of effort in addition to the magnitude of task. A baseline dyspnea index was developed to rate the severity of dyspnea at a single state and a transition dyspnea index to denote changes from the baseline condition. This approach eliminates the difficulty of using a single-state index for measuring change. To evaluate the usefulness of the indexes in clinical practice, we prospectively evaluated agreement among multiple observers who used the indexes to grade dyspnea in patients with respiratory disease and examined the relationship between the dyspnea scores and physiologic measurements.

**Methods**

**Construction of the Indexes**

Each index contains an arbitrary rating for three categories: functional impairment, magnitude of task needed to evoke dyspnea, and magnitude of effort needed to evoke dyspnea. At the initial or baseline evaluation, the patient's condition was rated from 0 (severe), to 4 (unimpaired), for each category. The ratings on each of the three categories were added to form a baseline focal score (range, 0 to 12). The lower the total score, the worse the severity of dyspnea.

The transitions or changes in a patient's dyspnea were compared with the baseline state and were rated in seven grades for each category. Scores ranged from −3 (major deterioration), to 0 (unchanged), to +3 (major improvement). The ratings on each of three categories were added to form a transition focal score (range, −9 to +9).

During the construction of the indexes, we made provision for circumstances in which dyspnea could not be rated. For example, when no information was available about the severity, when the information was insufficient, or when the patient's capacity was compromised by other factors, the grading of dyspnea was not possible.

During the early development of the indexes, pilot tests were performed. When there were disagreements between observers for the rating of dyspnea, the observers would discuss their reasons for selecting a particular grade and then reinterview the patient together so that any disparity was clarified. As a result, some changes were made in the wording of the criteria for rating dyspnea. These modifications are included in the baseline and transition dyspnea indexes that are listed in the Appendix.

**Patient Population**

Baseline and transition dyspnea were evaluated in 38 male patients with respiratory disease. Mean age was 62 ± 10 years (range, 42 to 82 years). All patients were ambulatory and were selected from the outpatient clinics at our institution based on the complaint of difficulty breathing for at least one year. Thirty-two had chronic obstructive pulmonary disease (chronic bronchitis and/or emphysema), five had asthma, and one had interstitial fibrosis. According to the predicted values for forced expiratory volume in one second (FEV),14 airflow obstruction was categorized as mild (≥65 percent) in two patients, moderate (50 to 64 percent) in nine, and severe (<50 percent) in 26. One subject had a restrictive ventilatory defect. Most patients were receiving bronchodilator medication at the baseline state, and no attempt was made to alter therapy during the study. All patients gave informed consent to the interviews and physiologic testing.

**Observer Agreement in the Rating of Dyspnea**

For the baseline and transition ratings of dyspnea, each patient was interviewed separately by two observers who were members of a group of four technicians working in the pulmonary function laboratories. For each patient, the same two observers recorded the baseline dyspnea index on the first occasion and the transition dyspnea index on the subsequent visit. Individual members of the interviewing group varied with different patients, but the same two observers interviewed the same patient on both occasions. Each interview was performed separately, without knowledge of either the results of the pulmonary function tests or what had been noted by the other observers. In addition, a pulmonary physician (D.A.M.) performed independent observations on 17 of the 38 patients at baseline and on 14 of the 17 patients at the transition period. For each patient, the interview process took less than five minutes.

All observers were experienced in history taking for respiratory disease and asked open-ended questions concerning the patient's breathlessness. Observers then focused on specific criteria for the severity of breathlessness in each category as outlined in the clinical dyspnea indexes. Based on the patient's responses, the observer was able to grade the degree of impairment related to dyspnea. The intent of this interviewing process was to allow the observer to grade an individual's dyspnea as part of the usual or standard questions asked of a patient.

Interobserver agreement was determined for the baseline and transition focal scores using two methods: the percentage agreement and the weighted kappa (κw) value. Since the dyspnea scores are ranked as integers, such as 0 to 12 or −9 to +9, the percentage of agreement was calculated after weights were assigned to disparities among categories.28 Values for percentage agreement can range from 0 to 100 percent, with higher scores representing better degrees of agreement. The κw statistic provides for disparities in ranks and also adjusts for the amount of agreement that might occur by chance.28,27 The values of κw can range from −1 (total disagreement) to +1 (perfect agreement), with 0 representing the chance expected agreement. Although p values can be calculated for the κw values, the results depend strongly on sample size. With the number of comparisons in this study, the magnitude of κw rather than the p value, was the more appropriate indication of agreement.

**Physiologic Measurements**

Physiologic testing was performed on patients on the same day when dyspnea was graded. Using a pulmonary function testing system (47804A Hewlett-Packard), forced vital capacity (FVC) and
Table 1—Agreement of Observers in Ratings of Dyspnea Indexes

<table>
<thead>
<tr>
<th></th>
<th>Two Technicians</th>
<th></th>
<th>Physician vs Technicians</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No of Ratings</td>
<td>% Agreement</td>
<td>$\kappa$</td>
</tr>
<tr>
<td>Baseline index</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rating for functional impairment</td>
<td>37*</td>
<td>85</td>
<td>.53</td>
</tr>
<tr>
<td>Rating for magnitude of task</td>
<td>38</td>
<td>94</td>
<td>.72</td>
</tr>
<tr>
<td>Rating for magnitude of effort</td>
<td>38</td>
<td>92</td>
<td>.73</td>
</tr>
<tr>
<td>Focal score</td>
<td>37*</td>
<td>92</td>
<td>.70</td>
</tr>
<tr>
<td>Transition index</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rating for functional impairment</td>
<td>38</td>
<td>92</td>
<td>.66</td>
</tr>
<tr>
<td>Rating for magnitude of task</td>
<td>38</td>
<td>90</td>
<td>.57</td>
</tr>
<tr>
<td>Rating of magnitude of effort</td>
<td>38</td>
<td>87</td>
<td>.65</td>
</tr>
<tr>
<td>Focal score</td>
<td>38</td>
<td>90</td>
<td>.63</td>
</tr>
</tbody>
</table>

*One patient, who was a poor historian, was rated as amount uncertain for functional impairment.

FEV1, were measured with the subject in the standing position. Values were selected from the maximal expiratory flow-volume curve with greatest sum of FVC and FEV1. Predicted normal values for FVC and FEV1 were taken from Morris et al. 26

In addition, exercise performance was determined using the 12-minute walking test (12 MW), which was conducted in a hospital corridor 75.5 m long. Patients were instructed to walk as far as they could in 12 minutes. For the baseline measurement, each subject performed the test on two separate days, and the greater distance was selected. At the transition examination, only a single 12 MW test was performed.

Statistical Analyses

The relationships between the dyspnea scores and the results of physiologic tests were determined using Pearson's correlation coefficient (r) 27 To allow a single value for the dyspnea ratings for each patient, an arithmetic mean was taken of the ratings provided by the two or three observers. Assessments of change in physiologic tests were based on the incremental difference in the baseline and subsequent measurements of FVC, FEV1, or 12 MW.

RESULTS

Observer agreement using the dyspnea indexes is shown in Table 1. For the two technicians, agreement was 92 percent with the baseline focal scores and 90 percent with the transition focal scores. The percentage agreement for the three categories of the two indexes ranged from 85 percent to 94 percent. Values for $\kappa$ were .70 for the baseline focal scores and .63 for

Table 2—Relationships of Components of Scores for Dyspnea Indexes*

<table>
<thead>
<tr>
<th>Score for Magnitude of Task</th>
<th>Components of Baseline Scores</th>
<th>Score for Functional Impairment</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Components of Transition Scores

<table>
<thead>
<tr>
<th>Score for Change in Magnitude of Task</th>
<th>Components of Transition Scores</th>
<th>Score for Change in Effort</th>
<th>Score for Change in Functional Impairment</th>
</tr>
</thead>
<tbody>
<tr>
<td>-1</td>
<td>2</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td>2</td>
<td>15</td>
<td>2</td>
</tr>
<tr>
<td>+1</td>
<td>0</td>
<td>4</td>
<td>7</td>
</tr>
<tr>
<td>+2</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
</tbody>
</table>

*Numbers in the cells show the number of patients who had the cited pattern of scores. Boldface numbers show the sites where both scores had similar magnitude.
the transition focal scores; this represented substantial agreement. The $\kappa$ values of agreement for the three categories of the two indexes ranged from moderate to substantial at levels of .53 to .73. Agreement between the physician and the two technicians demonstrated similar, although slightly lower, levels than between the two technicians.

To evaluate the distinctions achieved by utilizing a rating system which included consideration for magnitude of effort and functional impairment, rather than relying solely on the magnitude of task, we compared the relationships for the dyspnea scores of the three categories (Table 2). At baseline, individual scores for effort and functional capacity were often rated lower than the corresponding magnitude of the task. For example, scores for magnitude of task and effort were similar in 28 patients at baseline, while the effort score was lower than the task score in eight patients and higher in one subject. Comparisons between scores for magnitude of task and functional capacity showed identical scores in 23 patients with disagreement in 15 cases.

The relationships between components of transition dyspnea scores were analogous to those observed at baseline. An apparent improvement in performance of tasks was often not accompanied by a corresponding increase in effort or functional capacity. Thus, the task and effort transition scores were similar in 27 patients, but the score for improvement in effort was less than that for the task change in seven patients, and greater in four subjects.

Measurements of lung function and 12 MW distance for the 38 patients at baseline and transition states are shown in Table 3. As noted by the baseline results, there was a wide spectrum of respiratory dysfunction and functional ability.

Individual focal scores for the baseline and transition dyspnea indexes are shown in Figure 1. These results illustrate that the patients had a wide range of breathlessness; baseline index scores ranged from 0 to 9, while transition values ranged from -3 to +6.

Correlations between the baseline focal scores for dyspnea and baseline physiologic measurements are shown in Figure 2. The baseline focal score for dyspnea

![Figure 1](http://journal.publications.chestnet.org/pdfaccess.ashx?url=/data/journals/chest/21413/)

**Figure 1.** Spectrum of baseline and transition focal scores for dyspnea in 38 patients. Each point represents an individual subject.

![Figure 2](http://journal.publications.chestnet.org/pdfaccess.ashx?url=/data/journals/chest/21413/)

**Figure 2.** Correlations between the baseline focal score for dyspnea and baseline physiologic measurements. Each circle shows the particular measurement under consideration. The numbers entered on the lines between the circles show the correlation coefficient ($r$) for each pair of measurements. * indicates $p<0.05$. 

---

**Table 3—Lung Function and 12-Minute Walking Distance at Baseline and Transition States in 38 Patients With Respiratory Disease**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Baseline State</th>
<th>Transition State (Absolute Changes)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean ± SD</td>
<td>Range</td>
</tr>
<tr>
<td>FVC, L</td>
<td>2.63 ± 0.70</td>
<td>1.56 to 4.09</td>
</tr>
<tr>
<td>FEV₁, L</td>
<td>1.22 ± 0.60</td>
<td>0.55 to 3.00</td>
</tr>
<tr>
<td>12 MW, m</td>
<td>822 ± 188</td>
<td>470 to 1300</td>
</tr>
</tbody>
</table>

*FVC = forced vital capacity; FEV₁ = forced expiratory volume in one second; 12 MW = 12-minute walking test.*
had the highest correlation with the 12 MW \( r = 0.60 \); \( p < 0.001 \), while significant but lower correlations existed for FVC \( r = 0.56 \); \( p < 0.001 \) and FEV\(_1\) \( r = 0.41 \); \( p = 0.01 \).

Correlations between the transition focal score for dyspnea and changes in physiologic measurements are illustrated in Figure 3. Although the changes in FEV\(_1\) and FVC correlated well together \( r = 0.64 \); \( p < 0.001 \), there was no relationship between changes in lung function and the change in 12 MW or the transition focal score for dyspnea. However, the transition focal score had a significant correlation \( r = 0.33 \); \( p = 0.04 \) with the change in 12 MW.

**DISCUSSION**

In this study two clinical indexes were developed to measure dyspnea. In contrast to previously described methods, such as the four- or five-point rating scale\(^7\) or the visual analog scale,\(^8\) the baseline and transition dyspnea indexes consisted of three categories which are major factors affecting the development of dyspnea: functional impairment, magnitude of task, and magnitude of effort. Our data demonstrated that a rating based on magnitude of task alone does not indicate the full clinical impact of dyspnea. Many patients may perform certain tasks only by reducing the associated effort, and the successful performance of an individual task may not always be reflected in general functional capacity. Thus, measurement of the specific components of dyspnea is important, and by adding these ratings, a focal or overall dyspnea score was obtained. These clinical indexes, which consider three categories affecting dyspnea, provide a more detailed description than the global ratings of the Medical Research Council scale\(^9\) and the oxygen-cost diagram.\(^10\)

A further advantage of these two new clinical indexes is that changes in dyspnea, including individual components, can be quantified. Because the transition dyspnea index is an instrument for comparing breathlessness to an earlier, baseline state, specific grades can be given to degrees of improvement or deterioration. This approach circumvents the limitations of applying a single-state scale repetitively for determining interval changes in dyspnea.

Both the baseline and transition dyspnea indexes were easy to use, reliable, and required relatively little time after the observers developed familiarity with the criteria for rating dyspnea. Our experience shows that physicians, as well as nonphysician personnel, can successfully use the dyspnea indexes. In fact, the ratings by nonphysicians may be more accurate than those of physicians, since a patient interviewed by someone other than the physician may be less likely to be affected by a perception of the physician's expectations, especially regarding response to therapy.

We observed occasional variation in the dyspnea ratings recorded for the same patient by different observers. However, the disagreements among observers were usually minor, and seldom involved discrepancies of more than one category. In any system of mensuration, minor variations are to be expected, and similar variability has been reported for pulmonary function measurements. For example, the coefficients of variation have ranged from 3 percent to 7 percent in measurements of FVC and from 5 percent to 14 percent in measurements of maximal mid-expiratory flow rates.\(^10,12\)

In our study, dyspnea ratings did not correlate well with results of pulmonary function. Based on previous studies, these findings were not unexpected, since dyspnea is a subjective symptom and has not corresponded with spirometric measurements in previous studies.\(^4\) However, the higher correlations between the dyspnea scores and the 12 MW were considered appropriate because physical exertion frequently precipitates shortness of breath. McGavin et al.\(^8\) also found that exercise performance (12 MW) was significantly related to dyspnea in COPD patients.

Despite the advances in our understanding of breathlessness, the precise mechanism or concept describing the sensation has not been clearly established. In 1961, Campbell and colleagues\(^33\) proposed that dyspnea was due to length-tension or mechanical inappropriateness of the respiratory muscles. Subsequently, Bakers and Tenney\(^34\) and Killian et al.\(^35\) demonstrated that the sensation of breathlessness in normal subjects can be related to an increase in physical
stimulus, such as the addition of resistive or elastic loads. Gottfried et al. showed that the perception of changes in airflow resistance is impaired in patients with COPD. In addition, Burki36 has reported that breathless patients with chronic airway obstruction have a significantly increased inspiratory neuro muscular drive compared to those who are not breathless.

The unique features of the new baseline and transition dyspnea indexes are the inclusion of the categories for functional impairment and magnitude of effort, the use of specific criteria for each rating, and the development of a separate clinical transition index for evaluating changes. This approach contrasts with psychophysical testing of respiratory sensations which has involved the measurement of perception of mechanical loads added to breathing. In addition, the ability to quantify dyspnea and to delineate its changes may contribute improved clinical precision in the investigation of the sensation of breathlessness. 38 Although the new dyspnea indexes may be useful in the routine activities of clinical practice as well as in the assessment of therapeutic agents, further testing is desirable in larger numbers of patients with different cardiorespiratory diseases.

APPENDIX
Baseline Dyspnea Index

**Functional Impairment**

- **Grade 4**: *No Impairment*. Able to carry out usual activities and occupation without shortness of breath.
- **Grade 3**: *Slight Impairment*. Distinct impairment in at least one activity but no activities completely abandoned. Reduction in activity at work or in usual activities, that seems slight or not clearly caused by shortness of breath.
- **Grade 2**: *Moderate Impairment*. Patient has changed jobs and/or has abandoned at least one usual activity due to shortness of breath.
- **Grade 1**: *Severe Impairment*. Patient unable to work or has given up most or all usual activities due to shortness of breath.
- **Grade 0**: *Very Severe Impairment*. Unable to work and has given up most or all usual activities due to shortness of breath.

**W**: *Amount Uncertain*. Patient is impaired due to shortness of breath, but amount cannot be specified. Details are not sufficient to allow impairment to be categorized.

**X**: *Unknown*. Information unavailable regarding limitation of magnitude of task.

**Y**: *Impaired for Reasons Other than Shortness of Breath*. For example, musculoskeletal problem or chest pain.

**Magnitude of Effort**

- **Grade 4**: *Extraordinary*. Becomes short of breath only with the greatest imaginable effort. No shortness of breath with ordinary effort.
- **Grade 3**: *Major*. Becomes short of breath only with such major activities as walking up a steep hill, climbing more than three flights of stairs, or carrying a moderate load on the level.
- **Grade 2**: *Moderate*. Becomes short of breath with moderate or average tasks such as walking up a gradual hill, climbing fewer than three flights of stairs, or carrying a light load on the level.
- **Grade 1**: *Light*. Becomes short of breath with light activities such as walking on the level, washing, or standing.
- **Grade 0**: *No Task*. Becomes short of breath at rest, while sitting, or lying down.

**Transition Dyspnea Index**

**Change in Functional Impairment**

- **3**: *Major Deterioration*. Formerly working and has had to stop working and has completely abandoned some of usual activities due to shortness of breath.
- **2**: *Moderate Deterioration*. Formerly working and has had to stop working or has completely abandoned some of usual activities due to shortness of breath.
- **1**: *Minor Deterioration*. Has changed to a lighter job and/or has reduced activities in number or dura-
tion due to shortness of breath. Any deterioration less than preceding categories.

0: No Change. No change in functional status due to shortness of breath.

+1: Minor Improvement. Able to return to work at reduced pace or has resumed some customary activities with more vigor than previously due to improvement in shortness of breath.

+2: Moderate Improvement. Able to return to work at nearly usual pace and/or able to return to most activities with moderate restriction only.

+3: Major Improvement. Able to return to work at former pace and able to return to full activities with only mild restriction due to improvement of shortness of breath.

Z: Further Impairment for Reasons Other than Shortness of Breath. Patient has stopped working, reduced work, or has given up or reduced other activities for other reasons. For example, other medical problems, being "laid off" from work, etc.

Change in Magnitude of Task

-3: Major Deterioration. Has deteriorated two grades or greater from baseline status.

-2: Moderate Deterioration. Has deteriorated at least one grade but fewer than two grades from baseline status.

-1: Minor Deterioration. Has deteriorated less than one grade from baseline. Patient with distinct deterioration within grade, but has not changed grades.

0: No Change. No change from baseline.

+1: Minor Improvement. Has improved less than one grade from baseline. Patient with distinct improvement within grade, but has not changed grades.

+2: Moderate Improvement. Has improved at least one grade but fewer than two grades from baseline.

+3: Major Improvement. Has improved two grades or greater from baseline.

Z: Further Impairment for Reasons Other than Shortness of Breath. Patient has reduced exertional capacity, but not related to shortness of breath. For example, musculoskeletal problem or chest pain.

Change in Magnitude of Effort

-3: Major Deterioration. Severe decrease in effort from baseline to avoid shortness of breath. Activities now take 50-100% longer to complete than required at baseline.

-2: Moderate Deterioration. Some decrease in effort to avoid shortness of breath, although not as great as preceding category. There is greater pausing with some activities.

-1: Minor Deterioration. Does not require more pauses to avoid shortness of breath, but does things with distinctly less effort than previously to avoid breathlessness.

0: No Change. No change in effort to avoid shortness of breath.

+1: Minor Improvement. Able to do things with distinctly greater effort without shortness of breath. For example, may be able to carry out tasks somewhat more rapidly than previously.

+2: Moderate Improvement: Able to do things with fewer pauses and distinctly greater effort without shortness of breath. Improvement is greater than preceding category, but not of major proportion.

+3: Major Improvement. Able to do things with much greater effort than previously with few, if any, pauses. For example, activities may be performed 50-100% more rapidly than at baseline.

Z: Further Impairment for Reasons Other than Shortness of Breath. Patient has reduced exertional capacity, but not related to shortness of breath. For example, musculoskeletal problem or chest pain.

ACKNOWLEDGMENTS: The writers thank J. Denos, G. Orlowski, D. Papanotth, and K. Weigel for their technical assistance and for applying the indexes to grade dyspnea in this study.

REFERENCES


12 Medical Research Council. Committee on research into chronic bronchitis. Instructions for use of the questionnaire on respiratory symptoms. Devon, WJ, Holman, 1963.


27. Landis RJ, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977; 33:159-74