Comparison of the Responsiveness of Different Disease-Specific Health Status Measures in Patients With Asthma*

Toru Oga, MD; Koichi Nishimura, MD; Mitsuhiro Tsukino, MD; Susumu Sato, MD; Takashi Hajiro, MD; and Michiaki Mishima, MD

Background: Disease-specific health status measures are characterized by higher responsiveness than generic measures and may be preferred in clinical trials. However, comparisons of responsiveness between various disease-specific measures have rarely been performed in asthma studies.

Study objective: We investigated and compared the responsiveness of health status scores in asthmatic patients during treatment using three different disease-specific measures: the Juniper Asthma Quality of Life Questionnaire (AQLQ), the Living with Asthma Questionnaire (LWAQ), and the Airways Questionnaire 20 (AQ20).

Methods: We attempted to follow up 170 patients with newly diagnosed asthma over a 6-month period. Patients underwent treatment with inhaled corticosteroids in accordance with the guideline. A health status evaluation using three disease-specific measures, and pulmonary function tests were performed on the initial visit, and at 3 months and 6 months. The effect size and the standardized response mean were used as responsiveness indexes.

Results: A total of 109 patients completed the 6-month follow-up and were then analyzed. All health status scores and FEV₁ measures improved during the first 3 months (p < 0.001). The total of the AQLQ scores showed high responsiveness indexes ranging from 1.28 to 1.46 between baseline and 3 months, and baseline and 6 months. Spearman correlation coefficients were smaller between the change in FEV₁ and the change in the LWAQ. Although the AQ20 also demonstrated high responsiveness, a ceiling effect was indicated.

Conclusions: The AQLQ was the most responsive measure during asthma treatment. The relationship between the change in airflow limitation and the change in the LWAQ was weaker compared to the AQLQ and the AQ20. Although the AQ20 was also responsive and its simplicity is favorable, the ceiling effect should be considered when using it.

(CHEST 2002; 122:1228–1233)

Key words: asthma; disease-specific measures; health status; responsiveness

Abbreviations: AQLQ = Asthma Quality of Life Questionnaire; ATS = American Thoracic Society; BDP = beclomethasone dipropionate; FP = fluticasone propionate; LWAQ = Living With Asthma Questionnaire; AQ20 = Airways Questionnaire 20

Health status has been evaluated as a primary or secondary outcome in many studies. Disease-specific measures are most appropriate for clinical trials in which disease-specific therapeutic interventions are being evaluated, because they have a higher responsiveness and their outcomes will be directly clinically relevant to the patient. Among the disease-specific measures of asthma, the Juniper Asthma Quality of Life Questionnaire (AQLQ) and the Living with Asthma Questionnaire (LWAQ) are widely used. While the usefulness of the AQLQ has been extensively evaluated, comparisons between the two measures have rarely been performed; with respect to responsiveness, conflicting results have been reported in clinical trials. In order to assess health status in clinical practice using a simple method that removes the problems related to the size and complexity of traditional assessment instruments, the Airways Questionnaire...
20 (AQ20) was developed for patients with asthma and COPD.\textsuperscript{13} The AQ20 contains 20 items, and takes only a few minutes to complete. Validation of the AQ20 as a measurement instrument has been reported in patients with asthma\textsuperscript{13} and COPD.\textsuperscript{14} However, it has been suggested that this short and simple measure with fewer items might be less responsive than more comprehensive measures of assessment.

In the present study, we wanted to evaluate which disease-specific questionnaire (the AQLQ, the LWAQ, or the AQ20) was the most responsive measure. Therefore, we investigated and compared the responsiveness of these three different disease-specific measures during asthma treatment in clinical practice. To undertake statistical comparisons of the responsiveness, the effect size and the standardized response mean were used as described by Harper et al\textsuperscript{15} to compare the responsiveness of four different questionnaires in patients with COPD.

**Materials and Methods**

**Subjects**

A total of 170 outpatients with newly diagnosed asthma as defined by the American Thoracic Society (ATS)\textsuperscript{16} were recruited at the Kyoto University Hospital between September 1997 and December 2000. The analytical criteria for the study were as follows: (1) confirmation of the presence of bronchial hyperresponsiveness during the clinical course, (2) regular attendance for > 6 months; and (3) a best ratio of FEV\textsubscript{1} to FVC of > 0.7 when the subject was a current or former smoker to exclude COPD. Smoking habits were analyzed using patient self-reports. The present study was performed as part of our standard outpatient treatment and care, and verbal informed consent was obtained from all patients.

All subjects had undergone treatment with inhaled beclomethasone dipropionate (BDP) or fluticasone propionate (FP) for their asthma under supervision of an experienced pulmonologist (K.N.) according to the guideline proposed by the British Thoracic Society.\textsuperscript{17} A stepwise therapeutic regimen was used as follows: step 1, short-acting bronchodilators as required; step 2, step 1 plus low-dose inhaled corticosteroids (BDP, 200 to 500 \(\mu\)g/d; or FP, 100 to 400 \(\mu\)g/d); step 3, high-dose inhaled corticosteroids (BDP, 800 to 2,000 \(\mu\)g/d; or FP, 400 to 1,000 \(\mu\)g/d) plus short-acting bronchodilators as required; step 4, step 3 plus regular bronchodilators; and step 5, step 4 plus regular oral prednisolone. All patients warranted treatments at step 3 or greater at the initial visit in the clinical judgment of the physician. When an exacerbation of asthma occurred, a course of oral prednisolone or an increase in the treatment step was employed.

Bronchial hyperresponsiveness was evaluated by the method of Chai et al.\textsuperscript{18} Methacholine chloride was dissolved in phosphate buffered saline solution (pH 7.0). Subjects were then administered serial inhalations of phosphate buffered saline solution and methacholine solution delivered by a dosimeter (Rosenthal French; Baltimore, MD) using a nebulizer (Model 646; DeVilbiss; Somerset, PA) at a pressure of 139 kPa. They inhaled this aerosol during five slow inspiratory capacity maneuvers from functional residual capacity to total lung capacity, followed by a 3- to 5-s breath-hold. Nebulization was manually activated using the dosimeter, and was timed to hold a solenoid valve open for 0.6 s after being triggered. Subjects underwent spirometric testing according to the method recommended by the American Thoracic Society (ATS)\textsuperscript{19} three times before the inhalation, and twice after the inhalation of each concentration of methacholine. The procedure was stopped when FEV\textsubscript{1} fell by > 20% from the baseline value measured after the saline solution inhalation, or when 25 mg/mL methacholine solution was reached. The results were expressed as the dose of methacholine required to produce a 20% fall in FEV\textsubscript{1}. In the present study, the presence of bronchial hyperresponsiveness was defined as the provocative dose of methacholine required to produce a 20% fall in FEV\textsubscript{1} of < 100 cumulative units (one inhalation of 1 mg/mL methacholine solution = 1 cumulative unit).

**Methods**

At the initial, 3-month, and 6-month visits, spirometry and a health status assessment were performed. When an exacerbation requiring a change in the asthma treatment occurred within 4 weeks of the 3-month and 6-month visits, the assessment was postponed until the patient recovered.

The spirometric testing for determining FEV\textsubscript{1} and FVC was performed three times according to the method recommended by the ATS,\textsuperscript{19} using a spirometer (Autospiro AS-600; Minato Medical Science; Osaka, Japan) that was calibrated with a 3.0-L syringe. The largest FEV\textsubscript{1} and FVC values were then analyzed. The predicted values for FEV\textsubscript{1} and FVC were those established by the Japan Society of Chest Diseases.\textsuperscript{20}

Health status was assessed by three disease-specific questionnaires: the AQLQ, the LWAQ,\textsuperscript{5,6} and the AQ20.\textsuperscript{13} The official Japanese versions of the AQLQ, the LWAQ, and the AQ20 have been previously evaluated.\textsuperscript{14,21,22} One of the authors (M.T.) checked all the answers in front of the patient to avoid the possibility of missing values.

**Health Status Measures**

The AQLQ\textsuperscript{4} has 32 items, comprising four domains: symptoms, activity limitations, emotional function, and exposure to environmental stimuli. Five questions related to the activity limitations domain are self-identified by the patient. The self-identified activities at the first visit were retained at the 3-month and 6-month evaluations. Although the original version of the AQLQ was interviewer administered, in this study it was self-administered without informed administration. Patients were asked to recall their experiences during the previous 2 weeks and to score each item using a 7-point scale (1 = maximal impairment, 7 = no impairment). The domain scores were calculated as the mean score from the items forming each domain. In addition, the total score of the AQLQ was calculated as the mean of the sum of all items. Although the original version of the AQLQ did not introduce the idea of a total score, this has been used in other studies\textsuperscript{7-12} to compare results with other questionnaires.

The LWAQ\textsuperscript{5,6} consists of 11 domains, including 68 items in total, and is self-administered. For each item, the score ranges from 0 to 2 (from no impairment to maximum impairment). The total score of the LWAQ was calculated as the mean score of all items except for the items answered as not applicable, and a higher score indicates worse health status. Furthermore, to assess responsiveness, in addition to the total score, the LWAQ was separated into four constructs: avoidance, distress, preoccupation and activities.\textsuperscript{23}

The AQ20 is a unidimensional measure containing 20 items, and has three answers: “yes,” “no,” and “not applicable.”\textsuperscript{13} Yes responses are scored as 1, and the others are scored as 0. The AQ20 scores range from 0 to 20, with a score of 0 indicating no impairment of health status.
Statistical Analysis

All results are presented as mean ± SD. With regard to FEV₁, the significance of differences in the values was determined by a repeated-measures analysis of variance. When a significant difference was observed, the Fisher protected least significant difference method was performed to identify where the differences were significant. Since health status scores and treatment steps were regarded as nonparametric data, the significance between group variability was evaluated by Kruskal-Wallis analysis of variance, and the significance of individual differences was evaluated by the Mann-Whitney U test. The relationship between two sets of data were analyzed by the Spearman rank correlation test. Two widely used responsiveness indexes of the effect size and the standardized response mean were evaluated. The former indicates the ratio of the mean change in the score to the SD of the baseline scores. The latter indicates the ratio of the mean change in the score to the SD of that change. Cohen suggested that effect sizes of 0.2 to 0.5 were regarded as being “small,” 0.5 to 0.8 were “moderate,” and those ≥ 0.8 were “large,” although the standardized response mean is perhaps the closest to the “effect size” of Cohen. A p value of < 0.05 was considered to be statistically significant.

RESULTS

Among the 170 patients enrolled, 61 patients were unavailable for follow-up, and 109 patients completed the 6-month follow-up. These 109 patients were subsequently analyzed. Baseline characteristics of the 109 patients with asthma on their initial visit are shown in Table 1. Sixty-three patients (58%) were never-smokers. FEV₁ percentage of predicted values ranged widely from 18.9 to 116.6%.

Treatment steps, FEV₁, and health status scores during the 6-month period are presented in Table 2. At baseline, patients were undertreated with lower therapeutic steps. FEV₁ and all the health status scores showed significant improvements during the first 3 months (p < 0.001). FEV₁ and patient health status remained statistically unchanged after the first 3 months, except that the emotional function domain of the AQLQ improved significantly from the 3- to 6-month evaluation periods (p = 0.004).

As shown in Table 2, among the total scores of the AQLQ, the LWAQ, and the AQ20, the percentage of the best score was highest on the AQ20 (3.7%, 18.3%, and 30.3% at the initial, 3-month, and 6-month evaluations, respectively). No patients reached the best score on the LWAQ at any evaluation. However, on the total scores of the AQLQ and the LWAQ, no patients rated the possible worst scores at any time point. On the AQ20, the percentages of the worst scores were 0.9%, 0.9%, and 0.0% at each evaluative time point, respectively.

The responsiveness of each health status measure from baseline to 3 months and 6 months was evaluated by the effect size and the standardized response mean (Table 3). Among the total scores of the AQLQ, the LWAQ, and the AQ20, the AQLQ showed the highest responsiveness from baseline to 3 months and 6 months, followed by the AQ20. The responsiveness indexes on the total score of the LWAQ were the lowest, from 0.66 to 0.91. Among the four domains of the AQLQ, the responsiveness indexes on the symptoms domain were the highest, ranging from 1.25 to 1.53. Those of the emotional function domain showed an improvement between the baseline to 3-month evaluation and to 6-month evaluation. With regard to the LWAQ, the responsiveness indexes on the four constructs were lower than those on the total score of the LWAQ.

Spearman correlation coefficients between the changes in FEV₁ and health status scores from baseline to the 3-month evaluation are shown in Table 4. The relationship of the changes in health status to the changes in FEV₁ was the lowest on the LWAQ (r_s = 0.23, p = 0.02). The changes in the AQLQ, the LWAQ and the AQ20 were moderately correlated with each other.

DISCUSSION

We demonstrated that among the three asthma-specific measures used in the present study, the AQLQ showed the highest responsiveness. Furthermore, the AQ20 proved to be a highly responsive measure in spite of its simplicity; however, the ceiling effect was most prominent with the AQ20.

The present study showed that the AQLQ was more responsive to asthma treatment than the LWAQ based on the responsiveness indexes. This observation might have an impact on the problem of which measure to choose in clinical trials, because a highly responsive measure allows them to be performed with fewer patients. Previously, Rutten-van Mölken et al reported the superiority of the AQLQ over the LWAQ, while van der Molen et al reported the opposite. This contradiction may be explained by the fact that “informed administration”
was performed on the AQLQ in the former study,
which may have improved the responsiveness in the
former study. However, in the present study,
although the AQLQ was performed without in-
formed administration, high responsiveness of the
AQLQ was shown. First, better correlation was
found between the change in FEV1 and the change
in the AQLQ than the change in the LWAQ, as
similarly observed by Rutten-van Molken et al.
Therefore, in studies where changes in pulmonary
function would be expected, the AQLQ may be more
responsive than the LWAQ. Second, the 7-point
scale of the AQLQ might originally heighten respon-
siveness compared to the 3-point scale of the
LWAQ. In the present study, although the LWAQ
was separated into four constructs for the analysis of
responsiveness as each construct might change differ-
entially, responsiveness was not higher than that
of the total score of the LWAQ.
The symptoms and activity limitations domains of
the AQLQ showed a higher responsiveness from
baseline to the 3-month evaluation than the emo-
tional and environmental domains of the AQLQ, as
reported by Rowe and Oxman. However, signifi-

Table 2—Treatment Steps, FEV1, and Health Status Scores During the First 6 Months in 109 Patients With Asthma

<table>
<thead>
<tr>
<th>Variables</th>
<th>Baseline</th>
<th>3 mo</th>
<th>6 mo</th>
<th>p Value†</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment steps (1–5)</td>
<td>1.8 ± 1.2</td>
<td>3.2 ± 0.5</td>
<td>3.1 ± 0.5</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>FEV1, % predicted</td>
<td>71.6 ± 21.9</td>
<td>86.0 ± 17.9</td>
<td>85.9 ± 18.2</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>AQLQ (1–7)*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Symptoms</td>
<td>4.24 ± 1.27</td>
<td>6.08 ± 0.89</td>
<td>6.19 ± 0.86</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Activities</td>
<td>4.64 ± 1.25</td>
<td>5.95 ± 0.90</td>
<td>6.08 ± 1.01</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Emotions</td>
<td>4.62 ± 1.28</td>
<td>5.81 ± 1.04</td>
<td>6.13 ± 0.94</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Environment</td>
<td>5.25 ± 1.16</td>
<td>6.21 ± 0.85</td>
<td>6.30 ± 0.88</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Total</td>
<td>4.56 ± 1.09</td>
<td>6.01 ± 0.84</td>
<td>6.16 ± 0.85</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>% best score†</td>
<td>0.0</td>
<td>5.5</td>
<td>9.1</td>
<td></td>
</tr>
<tr>
<td>LWAQ (0–2)*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Avoidance</td>
<td>0.80 ± 0.43</td>
<td>0.57 ± 0.40</td>
<td>0.56 ± 0.43</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Distress</td>
<td>0.90 ± 0.48</td>
<td>0.63 ± 0.40</td>
<td>0.59 ± 0.44</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Preoccupation</td>
<td>0.93 ± 0.42</td>
<td>0.70 ± 0.41</td>
<td>0.65 ± 0.42</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Activities</td>
<td>0.71 ± 0.49</td>
<td>0.49 ± 0.42</td>
<td>0.45 ± 0.45</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Total</td>
<td>0.85 ± 0.38</td>
<td>0.60 ± 0.35</td>
<td>0.57 ± 0.37</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>% best score†</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td></td>
</tr>
<tr>
<td>AQ20 (0–20)*</td>
<td>8.1 ± 4.8</td>
<td>3.5 ± 4.0</td>
<td>3.1 ± 3.8</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>% best score†</td>
<td>3.7</td>
<td>18.3</td>
<td>30.3</td>
<td></td>
</tr>
</tbody>
</table>

*Data in parentheses indicate the theoretical score range.
†p values between the baseline, 3-month, and 6-month evaluations.
‡% best score = percentage of patients with the best score of each measure.
§p < 0.001, compared with the baseline.
||p < 0.01, compared with the 3-month evaluation.

Table 3—Responsiveness From Baseline to 3 Months and 6 Months in 109 Patients With Asthma

<table>
<thead>
<tr>
<th>Variables</th>
<th>Effect Size</th>
<th>Standardized Response Mean</th>
<th>Effect Size</th>
<th>Standardized Response Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>FEV1, % predicted</td>
<td>0.66</td>
<td>0.76</td>
<td>0.66</td>
<td>0.73</td>
</tr>
<tr>
<td>AQLQ Symptoms</td>
<td>1.44</td>
<td>1.25</td>
<td>1.53</td>
<td>1.31</td>
</tr>
<tr>
<td>Activities</td>
<td>1.04</td>
<td>1.15</td>
<td>1.15</td>
<td>1.17</td>
</tr>
<tr>
<td>Emotions</td>
<td>0.93</td>
<td>0.93</td>
<td>1.18</td>
<td>1.17</td>
</tr>
<tr>
<td>Environment</td>
<td>0.83</td>
<td>0.83</td>
<td>0.91</td>
<td>0.83</td>
</tr>
<tr>
<td>Total</td>
<td>1.53</td>
<td>1.28</td>
<td>1.46</td>
<td>1.34</td>
</tr>
<tr>
<td>LWAQ Avoidance</td>
<td>0.53</td>
<td>0.74</td>
<td>0.58</td>
<td>0.74</td>
</tr>
<tr>
<td>Distress</td>
<td>0.56</td>
<td>0.75</td>
<td>0.63</td>
<td>0.77</td>
</tr>
<tr>
<td>Preoccupation</td>
<td>0.55</td>
<td>0.52</td>
<td>0.67</td>
<td>0.62</td>
</tr>
<tr>
<td>Activities</td>
<td>0.47</td>
<td>0.66</td>
<td>0.55</td>
<td>0.66</td>
</tr>
<tr>
<td>Total</td>
<td>0.66</td>
<td>0.86</td>
<td>0.76</td>
<td>0.91</td>
</tr>
<tr>
<td>AQ20</td>
<td>0.96</td>
<td>1.05</td>
<td>1.06</td>
<td>1.13</td>
</tr>
</tbody>
</table>

www.chestjournal.org

CHEST / 122 / 4 / OCTOBER, 2002
significant improvement was found between the 3-month and 6-month evaluations only in the emotional domain of the AQLQ. This shows that even while pulmonary function and asthma symptoms remained stable, the emotional domain could detect small changes. Therefore, although the responsiveness indices of the emotional domain were lower than the other domains of the AQLQ, the factors evaluated by this domain would give us further insight into the effects of asthma treatment on patients.

In the present study, the AQ20 was also highly responsive during asthma treatment. This was understandable given the observation that the changes in the AQ20 correlated well with changes in FEV1, and with changes in the AQLQ and the LWAQ. The AQ20 is characterized by simplicity with 20 items and yes/no responses, but can be widely applied in the assessment of both asthma and COPD. This is in contrast with the two asthma-specific measures of the AQLQ (32 items and a 7-point scale) and the LWAQ (68 items and a 3-point scale). Conversely, merits can cause stronger ceiling effects in the AQ20 especially in patients with well-controlled asthma than on the AQLQ and the LWAQ. In the present study, 18.3% and 30.3% of the patients rated the best score of the AQ20 at 3 months and 6 months, respectively. Therefore, this disadvantage should be taken into consideration when using the AQ20 as an outcome measure in clinical trials. In the present study, patients with newly diagnosed asthma were enrolled at the first evaluation; therefore, the responsiveness of the AQ20 might be overestimated. Further study is needed to investigate which patients with asthma should be targeted for application of the AQ20.

Some limitations of the present study should be mentioned. First, this study did not investigate the test-retest reliability, which might have lowered the accuracy of the comparisons of responsiveness. Second, this was not a randomized, controlled trial. The purpose of the present study, however, was to compare the responsiveness of the different measures of health status, and not to investigate factors related to improvements in health status. Third, as is characteristic of the university hospital, most of our patients belong to a referral population. Therefore, the patient sample of the present study does not necessarily represent all asthmatics. Fourth, dropouts were not included in the analysis. Although health status of this group might differ, these patients would not have a differential response across different disease-specific health status measures. Therefore, selection bias due to dropouts is unlikely to have an important effect on the comparison of the responsiveness of different measures.

In the present study, we demonstrated that among the disease-specific measures of the AQLQ, the LWAQ and the AQ20, the AQLQ was the most responsive to asthma treatment. The AQLQ seems to have become the most popular recent measure of health status in asthma clinical trials, and this choice appears to be justified with regard to responsiveness. The lower responsiveness of the LWAQ relative to the AQLQ was considered to be partly due to the lower correlation of this measure with changes in airflow limitation. Although the AQ20 was also a responsive measure and its simplicity of completion would be more favorable in clinical practice, the stronger ceiling effect should be taken into consideration when using this questionnaire.

References


Table 4—Spearman Correlation Coefficients Between the Changes in the Different Indices From Baseline to the 3-Month Evaluation*

<table>
<thead>
<tr>
<th>Variables</th>
<th>FEV1</th>
<th>AQLQ Total</th>
<th>LWAQ Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>AQ20</td>
<td>0.38</td>
<td>–0.23</td>
<td>–0.62</td>
</tr>
<tr>
<td>AQLQ total</td>
<td>0.62</td>
<td>–0.54</td>
<td>0.60</td>
</tr>
</tbody>
</table>

*All relationships were statistically significant (p < 0.05).
16 American Thoracic Society. Standards for the diagnosis and care of patients with chronic obstructive pulmonary disease (COPD) and asthma: This official statement was adopted by the ATS Board of Directors, November 1986. Am Rev Respir Dis 1987; 136:225–244