Comparison of Clinical Findings with Symptom Assessment Systems (GerdQ and FSSG) for Functional Gastrointestinal Diseases

Mizue Matsuura¹, Masahiko Inamori², Takashi Nonaka¹, Yumi Inoh¹, Kenji Kanoshima¹, Takuma Higurashi¹, Hidenori Ohkubo¹, Yusuke Sekino¹, Hiroshi Iida¹, Koji Fujita¹, Akihiko Kusakabe³, Atsushi Nakajima¹

¹Gastroenterology and Hepatology, Yokohama City University School of Medicine, Yokohama, Japan
²Department of Medical Education, Yokohama City University School of Medicine, Yokohama, Japan

ABSTRACT

Background/Aims: We conducted a trial to evaluate the FSSG questionnaire and GerdQ questionnaire for diagnostic accuracy given correlation with clinical findings.

Methods: One hundred three consecutive patients who were seen in the Yokohama City University hospital with a gastrointestinal symptom in May 2011 were enrolled as potential subjects. Of these, 94 patients underwent upper gastrointestinal endoscopy, and clinical symptoms were evaluated using both the FSSG and GerdQ questionnaires.

Results: There were 94 subjects with GERD and 9 subjects with functional dyspepsia as defined by the Rome III criteria. We investigated the sensitivity and specificity of FSSG and GerdQ for GERD. For FSSG, the sensitivity was 0.564, the specificity was 0.778, and the odds ratio was 4.462. For GerdQ, the sensitivity was 0.298, the specificity was 0.889, and the odds ratio was 3.363. When combining the FSSG and GerdQ, the sensitivity was 0.255, the specificity was 0.888, and the odds ratio was 2.722. When using either the FSSG or GerdQ, the sensitivity was 0.723, the specificity was 0.888, and the odds ratio was 5.307.

Conclusions: We conclude that having patients complete both the FSSG and GerdQ may be more useful in routine medical examinations than completing only one questionnaire.

Key words: GERD, GerdQ, FSSG

INTRODUCTION

Gastroesophageal reflux disease (GERD) is common in the community and seen daily in clinical practice. The prevalence of GERD in the general population of Western countries is approximately 10–20% (1), and this condition accounts for up to 4% of consultations with family physicians. There is evidence that primary care physicians face challenges in making an accurate diagnosis of GERD and in managing it effectively.

GERD occurs when there is reflux of the stomach contents into the esophagus.
Patients present with a variety of symptoms (2,3), which can make an accurate diagnosis difficult in everyday clinical practice. Questionnaires have therefore been developed for this purpose. The initial diagnosis of GERD is based on either the patient’s medical history or questionnaires such as the questionnaire for the diagnosis of reflux disease (QUEST) produced by Carlsson et al. (4), in addition to the findings upon upper gastrointestinal endoscopy.

Other diagnostic modalities for GERD include 24-h esophageal pH monitoring and the proton pump inhibitor (PPI) test, which is a therapeutic diagnostic method (5). Because a diagnosis based on patient history is the simplest and quickest method and places no demands on the patient, it is favored by general practitioners.

Questionnaires are an extremely important aid for rapidly achieving an accurate diagnosis, assisting in the selection of suitable treatment, and monitoring the therapeutic response, without expensive evaluations. In addition, they are an essential component of clinical trials that aim to comprehensively evaluate pharmacotherapies for chronic diseases such as GERD and functional dyspepsia.

The GerdQ and the frequency scale for the symptoms of GERD (FSSG) are self-administered questionnaires. Japanese studies on the clinical signs and symptoms of GERD (6-15) have led to the development of the FSSG questionnaire for use in Japanese patients (16). The FSSG scores the frequency at which patients experience symptoms and has been shown to be useful in the diagnosis and monitoring of GERD. Thus, FSSG is a major diagnostic instrument for GERD in Japan.

The GerdQ has been developed as a tool to support the diagnosis of GERD and assist in the selection of a suitable treatment based on response measurement. This questionnaire was developed on the basis of evidence and information collected from recent high-quality clinical studies as well as from qualitative patient interviews with regard to the preference for easy completion of questionnaires (17).

In this study, we conducted a single-center trial in Japan to compare the FSSG and GerdQ and to examine their diagnostic accuracy correlated with clinical findings.

**MATERIALS AND METHODS**

**Subjects**

A total of 103 consecutive patients who were seen at Yokohama City University Hospital with a gastrointestinal condition in May 2011 qualified for this study. We excluded potential subjects aged less than 20 years, those who did not consent to participate, those who had previously completed either the GerdQ or the FSSG, those who were unable to complete the questionnaires by themselves, and those who had comorbid conditions, including stomach cancer, gastric ulcer, heart failure, renal insufficiency, and liver cirrhosis (table 1).

After an explanation of the study, all of the enrolled subjects gave written consent. This study was conducted with the approval of the Ethics Committee at Yokohama City University Hospital. Ninety-four subjects underwent upper gastrointestinal endoscopy, and clinical symptoms were evaluated using both the FSSG and GerdQ questionnaires.

**FSSG questionnaire**

Japanese studies on the clinical signs and symptoms of GERD (6-15) led to the development of the FSSG questionnaire for use in Japanese patients (16). The FSSG scores the frequency at which patients experience symptoms and has been shown to be useful in the diagnosis and monitoring of GERD. The questions are related to the 12 symptoms that Japanese patients with GERD complain of most often; these include not only ‘heartburn’ and ‘acid taste’ but also dyspeptic symptoms such as ‘heavy stomach’ and ‘feeling full quickly’. An FSSG score ≥8 is considered to indicate probable GERD (16,18,19).

**GerdQ questionnaire**

The GerdQ has been developed as a tool to support the diagnosis of GERD and assist in the selection of a suitable treatment based on response measurement. This questionnaire was developed on the basis of evidence and information collected from recent high-

<table>
<thead>
<tr>
<th>Number of subjects</th>
<th>103</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>65.5 (29-82)</td>
</tr>
<tr>
<td>Sex (male / female)</td>
<td>50 / 53</td>
</tr>
<tr>
<td>With GERD</td>
<td>94</td>
</tr>
<tr>
<td>With FD</td>
<td>9</td>
</tr>
<tr>
<td>With RE</td>
<td>27</td>
</tr>
<tr>
<td>Taking medication</td>
<td>84</td>
</tr>
<tr>
<td>Taking medication (PPI)</td>
<td>68</td>
</tr>
<tr>
<td>Taking medication (H2 blocker)</td>
<td>5</td>
</tr>
<tr>
<td>Taking medication (others)</td>
<td>35</td>
</tr>
<tr>
<td>FSSG score ≥8</td>
<td>55</td>
</tr>
<tr>
<td>FSSG score &lt;8</td>
<td>47</td>
</tr>
</tbody>
</table>

Table 1 - Demographic and clinical characteristics of the subjects. The data are presented as the median (range, when available)
quality clinical studies as well as from qualitative patient interviews with regard to the preference for easy completion of questionnaires. A GerdQ score ≥8 is considered to indicate probable GERD (17).

**Study design**

This randomized, two-way crossover study was conducted in May 2011 at Yokohama City University Hospital. The GerdQ and the FSSG are self-administered questionnaires. None of the subjects in our study had previously completed either the GerdQ or the FSSG. Subjects were randomly assigned to complete either the GerdQ first and the FSSG second, or the FSSG first and the GerdQ second, by themselves. Japanese versions of both questionnaires were used in this study, and we analyzed the questionnaire scores and clinical characteristics.

**Statistical analyses**

Statistical evaluations were performed using the Wilcoxon signed-rank test, the χ² test, the Mann-Whitney U test and Fisher’s exact test. The level of significance was set at p<0.05. Statistical analyses were performed using Stat View Software (SAS Institute, Cary, NC).

**Ethics**

This study was conducted in accordance with the Declaration of Helsinki. Each subject was provided information on the scientific purposes of the study and gave their written informed consent. The study protocol using the GerdQ and the FSSG was approved by the Ethics Committee of Yokohama City University Hospital.

### RESULTS

#### Clinical characteristics of the subjects

All 103 patients completed the study. The group consisted of 50 men and 53 women, and the median age of the patients was 65.5 years (range 29-82 years) (table 1). There were 94 subjects with GERD and 9 subjects with functional dyspepsia (FD), as defined by the Rome III criteria. The presence of GERD was defined as two or more episodes of heartburn per week (table 1). There were 27 subjects with reflux esophagitis (grade A-D) and 67 subjects without reflux esophagitis (table 1). There were 55 subjects with an FSSG score ≥8 and 47 subjects with a FSSG score <8 (table 1).

#### Sensitivity and specificity of GERD diagnosis

We investigated the sensitivity and specificity of the FSSG and GerdQ for GERD. The results are shown in table 2. For the FSSG, the sensitivity was 0.564; the specificity was 0.778; and the odds ratio was 4.462. For the GerdQ, the sensitivity was 0.298; the specificity was 0.889; and the odds ratio was 3.363. When the FSSG and GerdQ were combined, the sensitivity was 0.255; the specificity was 0.888; and the odds ratio was 2.722. When using either the FSSG or GerdQ, the sensitivity was 0.723; the specificity was 0.888; and the odds ratio was 5.307.

#### Sensitivity and specificity for reflux esophagitis (RE)

We also investigated the sensitivity and specificity of the FSSG and GerdQ for RE. The results are shown in table 2. For the FSSG, the sensitivity was 0.741; the

<table>
<thead>
<tr>
<th>FSSG (A)</th>
<th>GerdQ (A)</th>
<th>FSSG (B)</th>
<th>GerdQ (B)</th>
<th>GerdQ (C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sensitivity</td>
<td>0.564</td>
<td>0.298</td>
<td>0.741</td>
<td>0.481</td>
</tr>
<tr>
<td>Specificity</td>
<td>0.778</td>
<td>0.889</td>
<td>0.507</td>
<td>0.776</td>
</tr>
<tr>
<td>Odds ratio</td>
<td>4.462</td>
<td>3.363</td>
<td>3.308</td>
<td>3.434</td>
</tr>
<tr>
<td>p value</td>
<td>0.0783</td>
<td>0.439</td>
<td>0.0142</td>
<td>0.0119</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>FSSG and GerdQ (A)</th>
<th>FSSG or GerdQ (A)</th>
<th>FSSG and GerdQ (B)</th>
<th>FSSG or GerdQ (B)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sensitivity</td>
<td>0.255</td>
<td>0.723</td>
<td>0.444</td>
</tr>
<tr>
<td>Specificity</td>
<td>0.888</td>
<td>0.888</td>
<td>0.82</td>
</tr>
<tr>
<td>Odds ratio</td>
<td>2.722</td>
<td>5.307</td>
<td>3.817</td>
</tr>
<tr>
<td>p value</td>
<td>0.449</td>
<td>0.0356</td>
<td>0.00799</td>
</tr>
</tbody>
</table>
specificity was 0.507; and the odds ratio was 3.308. For the GerdQ, the sensitivity was 0.481; the specificity was 0.776; and the odds ratio was 3.434. When the FSSG and GerdQ were combined, the sensitivity was 0.444; the specificity was 0.82; and the odds ratio was 3.817. When using either the FSSG or GerdQ, the sensitivity was 0.777; the specificity was 0.313; and the odds ratio was 3.459.

**Sensitivity and specificity of the GerdQ with regard to the FSSG score**

We investigated the sensitivity and specificity of GerdQ with regard to the FSSG score, and the results are shown in table 2. The sensitivity of GerdQ was 0.455; the specificity was 0.917; and the odds ratio was 8.973.

*Odds ratio, p value, sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) (gold standard: FSSG score ≥ 8)*

The odds ratio, p value, sensitivity, specificity, PPV and NPV for age (≥65), sex (F), GerdQ (≥8), a diagnosis of GERD and a diagnosis of RE are shown in table 3. The odds ratio, p value, sensitivity, specificity, PPV and NPV were 0.403, 0.0294, 0.436, 0.340, 0.436, and 0.340, respectively, for patients aged ≥65; 2.085, 0.0769, 0.600, 0.583, 0.623, and 0.560, respectively, for female patients; 8.973, <0.001, 0.455, 0.917, 0.862, and 0.595, respectively, for patients with a GerdQ score ≥8; 4.462, 0.0783, 0.964, 0.146, 0.564, and 0.778, respectively, for patients diagnosed with GERD; and 3.308, 0.0142, 0.364, 0.854, 0.741, and 0.539, respectively, for patients diagnosed with RE. The FSSG and GerdQ results were correlated.

**DISCUSSION**

The FSSG was developed as a questionnaire for GERD that identifies symptoms in the order that Japanese patients with GERD most commonly present; specifically, it includes questions related to both acid reflux-associated symptoms and dyspeptic (dysmotility) symptoms. The inclusion of questions directed at both types of symptoms allows the FSSG to identify GERD symptoms more effectively. The FSSG can easily identify GERD symptoms, thereby avoiding the administration of the PPI test in patients (20). The FSSG score shows a good correlation with the extent of endoscopic improvement, indicating that this questionnaire is useful for objectively evaluating the therapeutic response of GERD (16).

The GerdQ has been developed as a tool to support the diagnosis of GERD and assist in selecting a suitable treatment based on response measurement. This questionnaire was developed on the basis of evidence and information collected from recent high-quality clinical studies as well as from qualitative patient interviews with regard to the preference for easy completion of questionnaires. The GerdQ not only assists with diagnosis but also provides the primary care physician with a tool for providing a more structured follow-up of GERD patients, evaluating the treatment response, and selecting patients for referral to secondary care (17). The GerdQ shows correlation with the extent of endoscopic diagnosis, indicating that this questionnaire is useful for objectively evaluating the therapeutic response of GERD (21).

The GSRS (Gastrointestinal Symptom Rating Scale) questionnaire can quantify the strength of digestive symptoms (e.g., heartburn, regurgitation, diarrhea, constipation, stomachache), thus quantifying the patient’s quality of life. The questionnaires described herein may be able to identify more symptoms than a doctor can detect, including unexpected events. Patients are able to answer questionnaires more objectively than when they are asked questions directly by a doctor, and patients can answer the questionnaires in a short time while waiting for examination. It has been reported that there is a difference in GERD severity as

<table>
<thead>
<tr>
<th>Odds ratio</th>
<th>p value</th>
<th>Sensitivity</th>
<th>specificity</th>
<th>PPV</th>
<th>NPV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (≥65)</td>
<td>0.403</td>
<td>0.0294</td>
<td>0.436</td>
<td>0.340</td>
<td>0.436</td>
</tr>
<tr>
<td>Sex (F)</td>
<td>2.085</td>
<td>0.0769</td>
<td>0.600</td>
<td>0.583</td>
<td>0.623</td>
</tr>
<tr>
<td>GerdQ (≥8)</td>
<td>8.973</td>
<td>&lt;0.001</td>
<td>0.455</td>
<td>0.917</td>
<td>0.862</td>
</tr>
<tr>
<td>GERD</td>
<td>4.462</td>
<td>0.0783</td>
<td>0.964</td>
<td>0.146</td>
<td>0.564</td>
</tr>
<tr>
<td>RE</td>
<td>3.308</td>
<td>0.0142</td>
<td>0.364</td>
<td>0.854</td>
<td>0.741</td>
</tr>
</tbody>
</table>
evaluated by a doctor versus a patient, which implies that patients have more severe GERD symptoms than doctors are lead to believe (22). It appears that questionnaires may be useful in closing this gap. Invasive investigations have traditionally been employed to differentiate between erosive and non-erosive GERD and to detect complications, but neither esophagogastrroduodenoscopy nor 24-h pH monitoring presents sufficient sensitivity to be accepted as a diagnostic gold standard (23). The PPI test shows high sensitivity and specificity, but it is an invasive evaluation method that is expensive and time consuming.

One limitation of this study is that it did not take personal medications into account. If the patients were taking any medications, their scores could show variations and errors. Another limitation of this study is that a translated version of the GerDQ was used, and the use of a translated questionnaire altered the nuances of the questions for the studied Japanese population.

The specificity of the GerDQ was higher than that of FSSG. Almost all of the patients with a total GerDQ score of 8 or greater were diagnosed with GERD. Our results indicated that the GerDQ could not distinguish between GERD and other conditions. The GerDQ was inferior to the FSSG in detecting the wide variety of symptoms associated with GERD. As the GerDQ and FSSG are both easy questionnaires to complete, the question of why the specificity of the GerDQ was higher arises. We hypothesize that this difference was due to the question that addresses pain and nausea, as the GerDQ score will be lower if patients are experiencing pain and nausea, whereas the FSSG score will be higher if patients report pain and nausea. Therefore, the GerDQ was inferior to the FSSG in identifying a wider variety of GERD symptoms.

The sensitivity and specificity of using the FSSG or GerDQ were high. The odds ratio obtained for the FSSG or GerDQ was higher compared with either the FSSG or GerDQ alone (table 2). The questionnaires must be able to detect a wide variety of GERD symptoms. Therefore, we propose that patients complete both the FSSG and GerDQ, whose combined results may be more useful in daily medical examinations compared with the application of either questionnaire alone.

Currently, there are no questionnaires that are widely used in daily clinical practice. We anticipate the development of a more suitable questionnaire that combines convenience and accurate identification in the near feature to better diagnose and treat various diseases.

Conflicts of Interest and Source of Funding

There were no conflicts of interest and no funding support in this research.

Acknowledgments

We thank the medical staff of the Yokohama City University Hospital Endoscopy Center.

REFERENCES