Research

Reliability and validity of the Gastrointestinal Symptom Rating Scale (GSRS) and Quality of Life in Reflux and Dyspepsia (QOLRAD) questionnaire in dyspepsia: A six-country study

Károly R Kulich*, Ahmed Madisch2, Franco Pacini3, Jose M Piqué4, Jaroslaw Regula5, Christo J Van Rensburg6, László Újszászy7, Jonas Carlsson1, Katarina Halling1 and Ingela K Wiklund1

Address: 1AstraZeneca R&D, Medical Science, Mölndal, S-431 86, Sweden, 2Medical Department I, Technical University Hospital, Dresden, 01307, Germany, 3Azienda Ospedaliera Careggi, U. O. di Gastroenterologia ed Endoscopia digestiva, Villa Medicea, Viale Pieraccini, 17, 50139, Firenze, Italy, 4Servicio de Gastroenterología, Hospital Clinic de Barcelona, Villarroel 170, 08036, Spain, 5Klinika Gastroenterologii CMKP, Centrum Onkologii, Roentgen Street 5, 02-781, Warszawa, Poland, 6Gastroenterology Unit, Tygerberg Hospital, Cape Town, 7505, South Africa and 7Semmelweis Hospital, Internal Medicine, Csabai Kapu 9-11, 3501, Miskolc, Hungary

Email: Károly R Kulich* - Karoly.Kulich@astrazeneca.com; Ahmed Madisch - ahmed.madisch@uniklinikum-dresden.de; Franco Pacini - pacinif@ao-careggi.toscana.it; Jose M Piqué - jmpique@clinic.ub.es; Jaroslaw Regula - jregula@coi.waw.pl; Christo J Van Rensburg - c.j.vr@telkomsa.net; László Újszászy - lujaszsz.1bel@semmelweis-miskolc.hu; Jonas Carlsson - jonas.carlsson@astrazeneca.com; Katarina Halling - katarina.halling@astrazeneca.com; Ingela K Wiklund - ingela.k.wiklund@gsk.com

* Corresponding author

Abstract

Background: Symptoms of dyspepsia significantly disrupt patients' lives and reliable methods of assessing symptom status are important for patient management. The aim of the current study was to document the psychometric characteristics of the Gastrointestinal Symptom Rating Scale (GSRS) and the Quality of Life in Reflux and Dyspepsia questionnaire (QOLRAD) in Afrikaans, German, Hungarian, Italian, Polish and Spanish patients with dyspepsia.

Methods: 853 patients with symptoms of dyspepsia completed the GSRS, the QOLRAD, the 36-item Short-Form Health Survey (SF-36) and the Hospital Anxiety and Depression scale.

Results: The internal consistency reliability of the GSRS was 0.43–0.87 and of the QOLRAD 0.79–0.95. Test-retest reliability of the GSRS was 0.36–0.75 and of the QOLRAD 0.41–0.82. GSRS Abdominal pain domain correlated significantly with all QOLRAD domains in most language versions, and with SF-36 Bodily pain in all versions. QOLRAD domains correlated significantly with the majority of SF-36 domains in most versions. Both questionnaires were able to differentiate between patients whose health status differed according to symptom frequency and severity.

Conclusion: The psychometric characteristics of the different language versions of the GSRS and QOLRAD were found to be good, with acceptable reliability and validity. The GSRS and QOLRAD were found to be useful for evaluating dyspeptic symptoms and their impact on patients' daily lives in multinational clinical trials.
Background
The definition and characterization of dyspepsia continue to challenge clinical investigators. At the Rome II consensus conference held in 1999, the recommended definition of dyspepsia was 'pain or discomfort centred in the upper abdomen' [1,2]. Yet, symptoms such as heartburn, nausea, post-prandial fullness, early satiety or bloating may also be present [3-7]. Recent guidelines suggest that gastroesophageal reflux disease (GER) is the likely diagnosis if heartburn is the only presenting symptom or the predominant symptom [8-10].

The label 'functional dyspepsia' is often used interchangeably with that of 'dyspepsia', but actually refers to patients whose dyspepsia has been investigated and has no known structural or biochemical cause [8,11]. Studies on the prevalence of dyspepsia in primary care originate mainly from a few developed countries, where the disorder accounts for 1–4% of all primary care consultations [6,12,13]. Estimates of the prevalence of dyspepsia in the adult general population are noticeably higher, ranging from 20–40% [7,11]. The frequency of medical visits for dyspepsia has been found to increase with increasing age [14], and with increasing severity of dyspeptic symptoms [15].

Dyspepsia symptoms impair patients’ daily functioning [16-21]. The symptom-driven nature of patient management places particular importance on a reliable estimation of symptom status. Hence, the assessment of how symptoms of dyspepsia affect patients' lives provides important information about the patients' health status and their perception of the treatment regime [22]. Moreover, this information helps enable clinicians to tailor treatment to the individual patient’s needs.

Despite increasing interest in evaluating the impact of dyspepsia symptoms on patients’ daily lives, measuring this impact in clinical trials can be problematic due to a paucity of validated instruments [23]. Furthermore, clinical trials are increasingly being conducted as multinational studies. This requires the availability of validated non-English language versions of the instrument to be used. Three patient-reported outcomes (PRO) instruments have been validated in patients with dyspepsia in clinical trials: the Severity of Dyspepsia Assessment (SODA) [24], the Nepean Dyspepsia Index (NDI) [25] and the Dyspepsia Symptom Severity Index (DSSI) [26]. However, all three instruments have only been validated as English-language versions. A further concern is that the NDI has not been validated for treatment effects, and the SODA was validated in a non-representative patient population comprising 96% men [24].

To be a viable measure of treatment outcome in clinical trials, PRO instruments need to be extensively documented to meet scientific standards [27,28] and to satisfy regulatory criteria, particularly with regard to claims for labelling and promotion [28,29]. The regulatory criterion is twofold: linguistic cross-cultural adaptation, and psychometric documentation. The US Food and Drug Administration (FDA) has recently issued a draft guidance on PRO measures [28]. In line with these guidelines, instruments need to show reliability, validity and an ability to detect change.

This report aims to document the psychometric characteristics of two PRO instruments, the Gastrointestinal Symptom Rating Scale (GSRS) and the Quality of Life in Reflux and Dyspepsia questionnaire (QOLRAD), in patients with dyspepsia. The GSRS and the QOLRAD were developed in US English. Several translations of both questionnaires have been psychometrically validated previously in patients with GERD [30-35]. In addition, English versions of the questionnaires have been used in dyspeptic populations (AstraZeneca, data on file). For the present report, a series of studies was undertaken in patients with dyspepsia in South Africa, Germany, Hungary, Italy, Poland and Spain. To enable comparisons between the various countries studied, results from all countries are presented here together in the current paper.

Methods
Patients
Consecutive patients with predominant symptoms of dyspepsia attending gastroenterology clinics between December 2000 and August 2003 were included in the study. In Germany and Spain, patients were also included from primary care. Dyspepsia was defined as persistent and/or recurrent pain or discomfort centred in the upper abdomen. Inclusion criteria required a history of episodes of dyspepsia for 6 months or longer, and episodes of dyspepsia on at least 1 day during the previous 7 days. The diagnosis based on the patient reports, and was verified at the discretion of the investigator. Where the diagnosis was uncertain, patients were not included. No physical examinations were performed. The following were exclusion criteria: GERD, irritable bowel syndrome, peptic ulcer disease, dementia or any other significant medical, psychiatric or surgical disease. Patients receiving daily treatment with acetyl salicylic acid or other nonsteroidal anti-inflammatory drugs were also excluded. Patients had to be able to complete the PRO instruments unaided, as no proxy assessment or interpreter was allowed. The administration of the questionnaires took place at visit 1 before any medical procedures were performed. Patients who were judged to be in stable phase were scheduled for a second visit. The second visit occurred 7 days after visit 1. The study was approved by local ethics committees in all...
countries requiring such approval for a non-drug related study. Good Clinical Practice was followed and the patients were free to discontinue participation in the study at any time.

Demographic and clinical variables
Clinicians recorded patient demographics (including age, sex, ethnicity, and marital and employment status), medical history (including history of gastrointestinal diseases) and frequency of dyspepsia symptoms (number of days with episodes during the last 7 days). Investigators also assessed the severity of patients’ dyspepsia symptoms by asking the patients and then recording the answer using a four-point graded scale (0 = none: no symptoms; 1 = mild: awareness of sign or symptom, but easily tolerated; 2 = moderate: discomfort sufficient to cause interference with normal activities; 3 = severe: incapacitating with inability to perform normal activities). All data were recorded in a case-report form.

PRO instruments
Patients completed two disease-specific and two generic PRO instruments: the GRSR, the dyspepsia version of the QOLRAD, the 36-Item Short-Form Health Survey (SF-36), and the Hospital Anxiety and Depression scale (HAD) (except in South Africa, where patients were not given the HAD). Questionnaires were completed in an electronic data capture (EDC) device (Apple Newton Pad, Apple Computers Inc, Cupertino, CA, USA), except in South Africa, where paper and pencil were used. The method of using electronic patient-reported outcomes (ePRO) for recording of symptoms and health-related quality of life has previously been shown to improve the quality of the data and to be well received by patients [36]. Study personnel were trained in how to use the electronic device and how to instruct the patients in a standardized way to minimize bias and enhance compliance.

Gastrointestinal Symptom Rating Scale (GSRS)
The GSRS is a disease-specific instrument of 15 items combined into five symptom clusters depicting Reflux, Abdominal pain, Indigestion, Diarrhoea and Constipation. The GSRS has a seven-point graded Likert-type scale where 1 represents absence of troublesome symptoms and 7 represents very troublesome symptoms. The reliability and validity of the GSRS are well-documented [37], and norm values for a general population are available [38].

Quality of Life in Reflux and Dyspepsia questionnaire (QOLRAD)
The dyspepsia version of the QOLRAD is a disease-specific instrument, including 25 items combined into five dimensions: Emotional distress, Sleep disturbance, Vitality, Food/drink problems and Physical/social functioning. The questions are rated on a seven-point graded Likert scale; lower values indicate a more severe impact on daily functioning. The reliability and validity of the QOLRAD have been documented in patients with dyspepsia [39]. Previous studies have shown that a difference of approximately 0.5 points represents a clinically relevant change [40]. The factor structure of the QOLRAD has been replicated for several language versions [41].

36-item Short-Form Health Survey (SF-36)
The SF-36 is an extensively used generic questionnaire containing 36 items clustered into eight dimensions (Bodily pain, General health, Mental health, Physical functioning, Role – emotional, Role – physical, Social functioning, Vitality). Item scores for each dimension are coded, summed and transformed to a scale from 0 (worst possible health status) to 100 (best possible health status). The SF-36 is well-documented in terms of reliability and validity in all available language versions [42,43]. This study used the acute version of the SF-36 covering a 1-week recall period [44].

Hospital Anxiety and Depression scale (HAD)
The HAD is a screening tool developed for use in medical settings. It consists of 14 items divided into two subscales for anxiety (seven items) and depression (seven items), in which the patient rates each item on a four-point scale. The higher scores indicate the presence of problems. A cut-off of > 11 implies definite cases of anxiety or depression, a cut-off of 8–10 a probable case, and < 7 not a case. The validity and reliability of the HAD have been reported in several studies [45,46]. The HAD was not used in South Africa because the Afrikaans translation was not available at the start of the study.

All instruments have been translated and linguistically validated according to international guidelines [47]. The linguistic validation of a questionnaire is not a literal translation of the original questionnaire, but the production of a translation, which is conceptually equivalent to the original and culturally acceptable in the country in which the translation will be used. This translation process includes forward-and back-translations by different, independent translators. Several translations of the GSRS and the QOLRAD are available, including Afrikaans, German, Hungarian, Italian, Polish and Spanish versions, and all have been psychometrically validated previously in patients with GERD [30-35].

Psychometric evaluation of the GSRS and the QOLRAD
Reliability
*Internal consistency* refers to the extent to which the items within each dimension are interrelated. Cronbach’s alpha is the most widely used measurement for assessing internal consistency reliability [48]. A high alpha coefficient (≥
0.70) suggests that the items within a dimension measure the same construct and supports the construct validity.

Test-retest reliability measures the stability of a score derived from serial administrations of a measure by the same rater. Repeated measurements are made in the same individuals, at a time interval long enough to ensure independence. Here, patients whose symptoms were judged to be stable, and in whom the treatment – not study mandated – remained unchanged, were assessed between visits 1 and 2. An intraclass correlation coefficient (ICC) above 0.70 was considered to be acceptable [49].

Construct validity
Construct validity is concerned with whether the indicator actually measures the underlying attribute. The construct validity was examined by convergent, discriminant and known-groups validity.

Convergent validity consists of showing that a postulated dimension of the instrument correlates appreciably with all other dimensions that theory suggests should be related to it. Here, it was examined by: (a) correlating the GRSR with the QOLRAD; (b) correlating the GRSR and the QOLRAD with the SF-36 dimensions; and (c) correlating the GRSR and the QOLRAD with the HAD (except in South Africa) and with clinician-rated severity of dyspepsia symptoms. Similar dimensions in these instruments were expected to have high correlations with each other as shown by Pearson's product moment correlation. A strong correlation was considered to be over 0.60, a moderate between 0.30 and 0.60 and a low (very low) correlation below 0.30 [50]. Low correlations were expected between those dimensions that are theoretically unrelated constructs, thereby testing the discriminant validity of the instruments.

Known-groups validity consists of showing that an instrument can differentiate between groups of patients whose health status differs according to the characteristics of patients' disease, in this case clinician-rated frequency and severity of dyspepsia symptoms. In about half of patients in Hungary (47.1%) and Spain (64.0%). Full-time employment was highest in Hungary and lowest in South Africa and Spain (overall range: 41.4–64.0%).

The majority (59.3–70.9%) of patients had moderate dyspepsia symptoms over the past week, except in South Africa, where the majority (55%) had severe symptoms. In terms of frequency, the majority (61.0–64.9%) of patients in South Africa, Germany and Poland had symptoms on at least five days in the previous week, as did approximately half of patients in Hungary (47.1%) and Spain (48.4%). In Italy, the majority (53.7%) of patients had symptoms on 3 to 4 days in the previous week.

Psychometric evaluation
Reliability
Cronbach's alpha scores ranged from 0.43 (Abdominal pain, German version) to 0.87 (Constipation, Polish version) for the GRSR, and from 0.79 (Vitality, Afrikaans version) to 0.95 (Emotional distress, Spanish version) for the QOLRAD (Table 2). Scores were high (range: 0.72–0.87) for all language versions of the GRSR in the Indigestion, Diarrhoea and Constipation domains, thus demonstrating high internal consistency. Scores in the Reflux domain were also high for the German (0.72), Hungarian (0.73) and Italian (0.84) versions of the GRSR, but were below 0.7 for the Afrikaans, Polish and Spanish versions. Scores in the Abdominal pain domain were below 0.7 for all six countries. For the QOLRAD, Cronbach's alpha scores were high for all language versions in all domains.

Intraclass correlation coefficients for the two visits spaced about 1 week apart ranged from 0.36 (Abdominal pain, German version) to 0.75 (Indigestion, Hungarian and Spanish versions) for the GRSR, and from 0.41 (Vitality,
Intraclass correlation coefficients were high (> 0.7) for the Hungarian, Polish and Spanish versions of the GSRS in the Indigestion domain, for the Hungarian version in the Diarrhoea domain, and for the Afrikaans, Hungarian, Polish and Spanish versions in the Constipation domain. Intraclass correlation coefficients for the GSRS Reflux and Abdominal pain domains were below 0.7 for all countries. For the QOLRAD, intraclass correlation coefficients were high for all domains in the Hungarian version, and for all domains except Food/drink problems in the German version. Intraclass correlation coefficients were below 0.7 for all QOLRAD domains in the Afrikaans and Polish versions, and for all but one (Food/drink problems) domain in the Spanish version. Test-retest reliability was not assessed in Italy, as a follow-up with the physicians revealed that many patients were not in a stable phase between visits 1 and 2: some patients changed their self-medication between the two visits, and some were endoscoped.

Table 1: Patient demographics and clinical data for all countries (total N = 853). Values are in percentages unless otherwise specified.

<table>
<thead>
<tr>
<th>Variables</th>
<th>South Africa</th>
<th>Germany</th>
<th>Hungary</th>
<th>Italy</th>
<th>Poland</th>
<th>Spain</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of patients</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>visit 1</td>
<td>111</td>
<td>141</td>
<td>136</td>
<td>175</td>
<td>135</td>
<td>155</td>
</tr>
<tr>
<td>visit 2</td>
<td>54</td>
<td>54</td>
<td>82</td>
<td>101</td>
<td>78</td>
<td>72</td>
</tr>
<tr>
<td>Mean age (years), (SD)</td>
<td>42.2 (11.8)</td>
<td>43.8 (14.9)</td>
<td>45.1 (14.2)</td>
<td>45.8 (15.4)</td>
<td>45.0 (14.0)</td>
<td>48.3 (15.9)</td>
</tr>
<tr>
<td>Female</td>
<td>68.5</td>
<td>68.8</td>
<td>72.1</td>
<td>61.1</td>
<td>66.7</td>
<td>67.7</td>
</tr>
<tr>
<td>Caucasian</td>
<td>11</td>
<td>98</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Married</td>
<td>59.5</td>
<td>61.7</td>
<td>69.1</td>
<td>69.7</td>
<td>65.9</td>
<td>69.0</td>
</tr>
<tr>
<td>Full-time employed</td>
<td>41.4</td>
<td>61.0</td>
<td>64.0</td>
<td>48.6</td>
<td>52.6</td>
<td>43.9</td>
</tr>
<tr>
<td>Duration of current episode of dyspepsia symptoms &gt; 1 month</td>
<td>71.2</td>
<td>53.2</td>
<td>72.8</td>
<td>58.3</td>
<td>65.1</td>
<td>72.3</td>
</tr>
<tr>
<td>Duration of disease &gt; 5 years</td>
<td>26.1</td>
<td>41.8</td>
<td>24.3</td>
<td>19.4</td>
<td>40.0</td>
<td>49.7</td>
</tr>
<tr>
<td>Severity of symptoms last 7 days at least moderate</td>
<td>95.5</td>
<td>89.3</td>
<td>73.5</td>
<td>75.4</td>
<td>77.8</td>
<td>69.7</td>
</tr>
<tr>
<td>Symptoms on at least 5 days in previous week</td>
<td>64.9</td>
<td>61.0</td>
<td>47.1</td>
<td>28.6</td>
<td>62.2</td>
<td>48.4</td>
</tr>
<tr>
<td>No comorbidity</td>
<td>59.4</td>
<td>80.8</td>
<td>69.8</td>
<td>84.0</td>
<td>73.3</td>
<td>31.6</td>
</tr>
<tr>
<td>Concomitant medication</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No current medication</td>
<td>33.3</td>
<td>71.6</td>
<td>37.5</td>
<td>58.3</td>
<td>39.2</td>
<td>33.5</td>
</tr>
<tr>
<td>Proton pump inhibitors</td>
<td>21.6</td>
<td>15.6</td>
<td>25.7</td>
<td>6.8</td>
<td>27.4</td>
<td>20.6</td>
</tr>
<tr>
<td>Predominant dyspepsia symptom</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Abdominal pain</td>
<td>70.3</td>
<td>50.4</td>
<td>58.8</td>
<td>38.3</td>
<td>47.4</td>
<td>20.6</td>
</tr>
<tr>
<td>Abdominal discomfort</td>
<td>29.7</td>
<td>46.1</td>
<td>30.1</td>
<td>58.3</td>
<td>51.9</td>
<td>79.4</td>
</tr>
<tr>
<td>Previous peptic ulcer and/or ulcerative reflux esophagitis</td>
<td>14.4</td>
<td>4.3</td>
<td>6.6</td>
<td>9.1</td>
<td>8.1</td>
<td>5.2</td>
</tr>
<tr>
<td>Emotional problems, past 5 years</td>
<td>33.3</td>
<td>2.8</td>
<td>5.1</td>
<td>8.0</td>
<td>13.3</td>
<td>38.7</td>
</tr>
</tbody>
</table>

SD, standard deviation.

Table 2: Cronbach’s alpha at visit 1.

<table>
<thead>
<tr>
<th>Translation</th>
<th>Afrikaans</th>
<th>German</th>
<th>Hungarian</th>
<th>Italian</th>
<th>Polish</th>
<th>Spanish</th>
</tr>
</thead>
<tbody>
<tr>
<td>GSRS</td>
<td>(N = 108)</td>
<td>(N = 132)</td>
<td>(N = 124)</td>
<td>(N = 165)</td>
<td>(N = 134)</td>
<td>(N = 154)</td>
</tr>
<tr>
<td>Reflux</td>
<td>0.61</td>
<td>0.72</td>
<td>0.73</td>
<td>0.84</td>
<td>0.49</td>
<td>0.67</td>
</tr>
<tr>
<td>Abdominal pain</td>
<td>0.60</td>
<td>0.43</td>
<td>0.59</td>
<td>0.65</td>
<td>0.49</td>
<td>0.45</td>
</tr>
<tr>
<td>Indigestion</td>
<td>0.76</td>
<td>0.79</td>
<td>0.74</td>
<td>0.74</td>
<td>0.73</td>
<td>0.78</td>
</tr>
<tr>
<td>Diarrhoea</td>
<td>0.75</td>
<td>0.84</td>
<td>0.79</td>
<td>0.72</td>
<td>0.78</td>
<td>0.82</td>
</tr>
<tr>
<td>Constipation</td>
<td>0.82</td>
<td>0.81</td>
<td>0.79</td>
<td>0.76</td>
<td>0.87</td>
<td>0.75</td>
</tr>
<tr>
<td>QOLRAD</td>
<td>(N = 108)</td>
<td>(N = 123)</td>
<td>(N = 124)</td>
<td>(N = 174)</td>
<td>(N = 134)</td>
<td>(N = 154)</td>
</tr>
<tr>
<td>Emotional distress</td>
<td>0.91</td>
<td>0.94</td>
<td>0.94</td>
<td>0.93</td>
<td>0.92</td>
<td>0.95</td>
</tr>
<tr>
<td>Food/drink problems</td>
<td>0.83</td>
<td>0.90</td>
<td>0.90</td>
<td>0.88</td>
<td>0.86</td>
<td>0.91</td>
</tr>
<tr>
<td>Physical/social functioning</td>
<td>0.84</td>
<td>0.90</td>
<td>0.89</td>
<td>0.89</td>
<td>0.86</td>
<td>0.91</td>
</tr>
<tr>
<td>Sleep disturbance</td>
<td>0.88</td>
<td>0.89</td>
<td>0.94</td>
<td>0.89</td>
<td>0.91</td>
<td>0.93</td>
</tr>
<tr>
<td>Vitality</td>
<td>0.79</td>
<td>0.90</td>
<td>0.87</td>
<td>0.80</td>
<td>0.87</td>
<td>0.89</td>
</tr>
</tbody>
</table>

GSRS, Gastrointestinal Symptom Rating Scale; QOLRAD, Quality of Life in Reflux and Dyspepsia questionnaire.
Construct validity

Correlation of the GRSRS with the QOLRAD
In the Afrikaans versions, correlations (as defined by a Pearson’s product moment correlation = 0.3) were only observed between the GRSRS Abdominal pain domain and the QOLRAD Emotional distress and Physical/social functioning domains. In the German versions, the GRSRS Abdominal pain and Indigestion domains correlated with all QOLRAD domains, but the GRSRS Reflux domain correlated only with the QOLRAD Food/drink problems domain. The relevant GRSRS domains Reflux, Abdominal pain and Indigestion correlated with all QOLRAD domains in the Hungarian version of the questionnaires. This was also the case for the Italian versions of the questionnaires, with the exception of the GRSRS Reflux and Indigestion domains, which did not correlate with the QOLRAD Food/drink problems domain. In the Polish versions of the questionnaires, the GRSRS Reflux and Abdominal pain domains correlated with the QOLRAD Food/drink problems, Physical/social functioning, and Sleep disturbance domains. As was the case with the Italian versions, the GRSRS Reflux, Abdominal pain and Indigestion domains correlated with all QOLRAD domains in the Spanish version of the questionnaires.

Correlation of the GRSRS with the SF-36
Each domain of the GRSRS correlated negatively with each domain of the SF-36 in all language versions of the questionnaires. Significant correlations between GRSRS domains and SF-36 domains (Pearson’s product moment correlation = 0.3) are shown in Table 4.

Correlation of the QOLRAD with the SF-36
QOLRAD domains correlated significantly with the majority of SF-36 domains in most language versions. Significant correlations between QOLRAD domains and SF-36 domains (Pearson’s product moment correlation = 0.3) are shown in Table 4.

Correlation of the GRSRS with the HAD
Correlations between GRSRS domains and the HAD anxiety and depression scores are shown in Table 5. The GRSRS Abdominal pain and Indigestion domains correlated significantly with the HAD anxiety score in the German, Spanish and Hungarian populations. The GRSRS Reflux domain correlated with the HAD depression score only in the Hungarian population. (Note: HAD was not assessed in South Africa.)

Correlation of the QOLRAD with the HAD
Correlations between QOLRAD domains and the HAD anxiety and depression scores are shown in Table 5. All QOLRAD domains correlated with the HAD anxiety scores in the Hungarian, Italian and Spanish versions of the questionnaire, and all except Sleep disturbance correlated with the HAD anxiety scores in the German and Polish versions. Furthermore, all QOLRAD domains correlated with the HAD depression scores in the Italian and Hungarian versions of the questionnaire, and all except Food/drink problems correlated with the HAD depression score in the Spanish version.

Correlation of the GRSRS with physician-assessed severity of symptoms
The relevant GRSRS Abdominal pain domain correlated significantly with physician-assessed severity of symp-
Table 4: Significant correlations between Gastrointestinal Symptom Rating Scale (GSRS), Quality of Life in Reflux and Dyspepsia Questionnaire (QOLRAD) and 36-item Short-Form Health Survey (SF-36) domains (Pearson’s product moment correlation ≥ 3).

<table>
<thead>
<tr>
<th>Questionnaire and domain</th>
<th>SF-36 domain and correlation coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Afrikaans*</td>
</tr>
<tr>
<td>GSRS</td>
<td></td>
</tr>
<tr>
<td>Reflux</td>
<td></td>
</tr>
<tr>
<td>Abdominal pain</td>
<td>BP (0.45), PF (0.33)</td>
</tr>
<tr>
<td>Indigestion</td>
<td>BP (0.39), PF (0.34)</td>
</tr>
<tr>
<td>Diarrhoea</td>
<td>--</td>
</tr>
<tr>
<td>Constipation</td>
<td>BP (0.34)</td>
</tr>
<tr>
<td>QOLRAD</td>
<td></td>
</tr>
<tr>
<td>Emotional distress</td>
<td>BP (0.50), MH (0.35), RE (0.35), RP (0.43), SF (0.36), V (0.39)</td>
</tr>
<tr>
<td>Sleep disturbance</td>
<td>BP (0.39), GH (0.31), MH (0.35), RE (0.31), RP (0.41), SF (0.38), V (0.49)</td>
</tr>
<tr>
<td>Food/drink problems</td>
<td>BP (0.32), GH (0.31), V (0.41)</td>
</tr>
<tr>
<td>Physical/social functioning</td>
<td>BP (0.44), MH (0.39), PF (0.32), RE (0.44), RP (0.47), SF (0.48), V (0.39)</td>
</tr>
<tr>
<td>Vitality</td>
<td>BP (0.46), MH (0.32), RE (0.31), RP (0.35), SF (0.34), V (0.36)</td>
</tr>
</tbody>
</table>

BP, Bodily pain; GH, General Health; MH, Mental Health; PF, Physical functioning; RE, Role – Emotional; RP, Role – Physical; SF, Social Functioning; V, Vitality

The symptoms in the Afrikaans version of the GSRS. The GSRS Abdominal pain and Indigestion domains correlated significantly with physician-assessed severity of symptoms in the Spanish version of the questionnaire, and with physi-
Physician-assessed frequency and severity in the Hungarian version. Taken together, these correlations show convergent validity.

**Correlation of the QOLRAD with physician-assessed severity of symptoms**

All QOLRAD domains correlated with physician-assessed severity and frequency of symptoms in the Hungarian version of the questionnaire, and all domains except Sleep disturbance correlated with physician-assessed severity and frequency in the Spanish version. In the German versions, the QOLRAD emotional distress domain correlated with physician-assessed frequency of symptoms, and the QOLRAD Sleep disturbance domain correlated with physician-assessed severity. The QOLRAD Vitality domain correlated with physician-assessed severity in the Italian version of the questionnaire. In the Polish version, the QOLRAD Physical/social functioning and Vitality domains correlated with physician-assessed frequency and severity of symptoms, and the QOLRAD Emotional distress domain correlated with physician-assessed frequency of symptoms.

**Known-groups validity of the GSRS and the QOLRAD**

All domains of the GSRS and the QOLRAD questionnaires were able to differentiate between groups of patients whose health status differed according to the physician-assessed overall frequency and severity of dyspepsia symptoms, thereby confirming the known-groups validity of the instruments. GSRS scores increased with increasing frequency and severity of dyspepsia symptoms, while QOLRAD scores decreased with increasing frequency and severity of symptoms (representing a decrease in daily functioning) (results not shown).

**Discussion**

The GSRS and the QOLRAD are two of the most established, validated, reliable and responsive disease-specific instruments available for assessing gastrointestinal symptoms and their impact on patients' daily functioning [37,39]. Both questionnaires have been proven to have very good psychometric characteristics when tested in clinical trials in patients with GERD and dyspepsia [20,56,57]. This paper documents the psychometric validation of the Afrikaans, German, Hungarian, Italian, Polish and Spanish translations of the GSRS and the QOLRAD in patients with dyspepsia. The study was conducted...
mainly in gastroenterology centres and, therefore, results are particular to patients referred for gastroenterological investigation. Because of the standardized methodology used, results from all countries are directly comparable.

Questionnaires were completed in an electronic data capture device (ECD), which ensures that all questions are answered in full. Although ECD technology is becoming an increasing part of clinical trials, it is not yet widely available, and certainly not that widespread in clinical practice. Thus, in South Africa, paper versions of the questionnaires were used for the present study, and this may have affected the results obtained in this country, especially as the use of ECD technology has been shown to increase patient compliance and improve the quality of the data [36].

The demographic and clinical characteristics of the patient populations were comparable in the different countries studied in terms of the number of patients recruited, their age, gender, symptom status and medication use. The Afrikaans patient population had the highest incidence of abdominal pain as the predominant dyspeptic symptom as well as the highest proportion of patients with a history of previous peptic ulcer and/or ulcerative reflux esophagitis. This was probably due to the fact that all South African subjects were recruited from a single gastroenterology clinic in a healthcare system that may refer only the most severely affected patients.

In terms of psychometric characteristics, internal consistency was high in all domains of all language versions of the QOLRAD, thus supporting construct validity. Internal consistency was also high in the Indigestion, Diarrhoea and Constipation domains of the GSRS in all language versions. However, although abdominal pain is the typical dyspepsia symptom, internal consistency was only moderate in the Abdominal pain domain of the GSRS. This generally low internal consistency in the Abdominal pain domain probably reflects the difficulty of measuring this often fluctuating symptom, as well as the fact that the GSRS Abdominal pain domain contains only two items. In previous international studies conducted in patients with GERD, internal consistency also tended to be lower in the Abdominal pain domain of the GSRS than in its other domains [30-35].

Test-retest reliability was good for most or all domains in the Hungarian and German translations of the QOLRAD, but only for one domain in the Spanish version and for no domains in the Afrikaans and Polish versions of the QOLRAD. It is not clear why the test-retest reliability was low in these patient populations, although it may be an indication that the frequency, severity and impact of abdominal symptoms were not as stable on a week-to-week basis in these patients as was estimated by the investigators. No test-retest was performed in Italy, so it was not possible to assess test-retest reliability for the Italian version of the QOLRAD. Previous international studies in patients with GERD have yielded acceptable test-retest reliability for the German (ICC: 0.70–0.84), Spanish (ICC: 0.79–0.85) and Afrikaans (ICC: 0.71–0.82) versions of the QOLRAD, but only for some domains of the Polish version (ICC: 0.51–0.74) of the QOLRAD [30-35]. In the GSRS, test-retest reliability was acceptable in the Indigestion and Constipation domains in most language versions. However, reliability was lower in the Abdominal pain domains in all countries assessed. As with internal consistency, the generally low test-retest reliability in the Abdominal pain domain is most likely due to the GSRS Abdominal pain domain containing only two items, as well as the complexity of measuring this variable symptom. It is also likely that at least some of the patients deemed stable by the physicians experienced small changes in their symptoms that were picked up by the PROs.

In terms of construct validity, all GSRS domains correlated with all QOLRAD domains. Correlation was significant for the relevant GSRS Abdominal pain domain and most or all QOLRAD domains in the majority of the different language versions. The GSRS Abdominal pain domain also correlated significantly with the relevant SF-36 Bodily pain domain in all language versions. All QOLRAD domains correlated significantly with the majority of SF-36 domains in most language versions. Both the GSRS and the QOLRAD were able to differentiate between patients whose health status differed with regard to frequency and severity of symptoms, thereby confirming the known-groups validity of the instruments. Known-groups validity has also been confirmed for several translations of the GSRS and the QOLRAD in patients with GERD [30-32,35]. Confirmatory factor analysis has been shown to support the validity of the Scandinavian language versions of the GSRS and the QOLRAD in patients with GERD [41]. Future studies could use this approach to further assess the validity of the additional language versions presented here.

The relevant GSRS Abdominal pain domain correlated with the HAD anxiety score in most of the language versions assessed, and most of the QOLRAD domains correlated with the HAD anxiety score in all language versions assessed. Dyspepsia tends to be associated with anxiety and depression [58]. Previous reports are somewhat contradictory on the role of psychological morbidity in dyspepsia symptoms and healthcare-seeking behaviour [1,58], suggesting that additional research is required on this potentially important aspect of the disorder.
When comparing the frequency and severity of patient-reported symptoms with those assessed by the physician, correlations were only low to moderate. These discrepancies in patient-reported and observer-reported symptom status have been well-documented [59-61]. Our view is that the physician examination needs to be balanced with the patient-reported symptom status when deciding on patient management and outcomes [39,62].

In summary, all language versions of the QOLRAD showed good internal consistency and reliability, although only the German and Hungarian translations were able to demonstrate acceptable test-retest reliability. In general, all language versions of the GSRS were reliable and showed good internal consistency for the Indigestion, Diarrhoea and Constipation domains, but not for the Abdominal pain domain. Overall, the test-retest reliability of the GSRS was not acceptable.

**Conclusion**

In conclusion, the German, Hungarian, Italian, Polish, Spanish and Afrikaans versions of GSRS and QOLRAD are reliable and show good internal consistency for use in clinical trials. However, the test-retest reliability was low in most language versions. By evaluating symptoms and their impact on patients' lives, both these instruments help physicians interpret clinical benefits as outcomes of significance to patients. We believe that these results will further facilitate the use of these instruments in the clinical setting.

**Abbreviations**

DSSI, Dyspepsia Symptom Severity Index

EDC, electronic data capture

ePRO, electronic patient-reported outcomes

FDA, Food and Drug Administration

GERD, gastroesophageal reflux disease

GSRS, Gastrointestinal Symptom Rating Scale

HAD, Hospital Anxiety and Depression scale

ICC, intraclass correlation coefficient

NDI, Nepean Dyspepsia Index

PRO, patient-reported outcomes

QOLRAD, Quality of Life in Reflux and Dyspepsia questionnaire

SD, standard deviation

SF-36, 36-item Short-Form Health Survey

SODA, Severity of Dyspepsia Assessment

**Competing interests**

All authors participated in the present study supported by AstraZeneca R&D, Mölnndal, Sweden. Károly Kulich, Jonas Carlsson and Katarina Halling are employees of AstraZeneca R&D, Mölnndal, Sweden. Ingela Wiklund was an employee of AstraZeneca R&D, Mölnndal, Sweden at the time of the study. Ahmed Madisch, Franco Pacini, Jose Piqué, Jaroslaw Regula, Christoffel Johannes van Rensburg and László Újszászy have no additional financial or other relationships to disclose.

**Authors’ contributions**

All authors contributed to the study design, data analyses and interpretation of results. Ahmed Madisch, Franco Pacini, Jose Piqué, Jaroslaw Regula, Christoffel Johannes van Rensburg and László Újszászy coordinated the patient recruitment and data collection. Jonas Carlsson provided statistical support. All authors provided comments and revisions on a first version of the manuscript drafted by Károly Kulich, Katarina Halling and Ingela Wiklund. All authors have seen and approved the final version of the manuscript.

**Acknowledgements**

This study was funded by AstraZeneca R&D, Mölnndal, Sweden. Dr Anja Becher provided editorial assistance funded by AstraZeneca. The authors would like to acknowledge the investigators who participated in the recruitment of patients.

**Germany:** A Madisch, S Miehlke, Medical Department I, Technical University of Dresden, Dresden; P Malfertheiner, Otto-von-Guericke University, Magdeburg; K Ziegler, Private clinic, Berlin; E Bayerdörffer, Department of Internal Medicine, University Hospital of Marburg, Marburg; J Labenz, Department of Medicine, Jung-Stilling Hospital, Siegen.

**Hungary:** L Újszászy, Semmelweis Hospital, Internal Medicine, Miskolc; TG Tóth, Szent János Hospital, Internal Medicine, Budapest; L Bárány, Megyei Jogú Város Hospital, Internal Medicine, Nagykanizsa.


South Africa: CJ Van Rensburg, E Wilken, CF Kruger, JEC Botha, Gastroenterology Unit, Tygerberg Hospital, South Africa.

References

Food and Drug Administration: Patient-reported outcome measures: Use in medical product development to support labeling claims. [http://www.fda.gov/cder/guidance/5460dft.htm].

Publish with BioMed Central and every
scientist can read your work free of charge

"BioMed Central will be the most significant development for disseminating the results of biomedical research in our lifetime."
Sir Paul Nurse, Cancer Research UK

Your research papers will be:
• available free of charge to the entire biomedical community
• peer reviewed and published immediately upon acceptance
• cited in PubMed and archived on PubMed Central
• yours — you keep the copyright

Submit your manuscript here:
http://www.biomedcentral.com/info/publishing_adv.asp