Digestive Diseases and Sciences, Vol. 49, Nos. 11/12 (November/December 2004), pp. 1822–1829 (© 2004)

Development and Validation of a Cross-Cultural Questionnaire to Evaluate Nonulcer Dyspepsia: The Porto Alegre Dyspeptic Symptoms Questionnaire (PADYQ)

GUILHERME BECKER SANDER, MD, LUIZ EDMUNDO MAZZOLENI, MD, ScD, CARLOS FERNANDO MAGALHÃES FRANCESCONI, MD, ScD, ANDRÉ CASTAGNA WORTMANN, MD, EDUARDO ANDRÉ OTT, MD, ALEXANDRO THEIL, MD, VICENZO DA CRUZ PICCOLI, ÂNGELA CRISTIANE DA SILVA, MD, LEANDRO OLIVEIRA, MD, SIMONE BEHEREGARAY, MD, SIMONE MATIOTI, MD, GUSTAVO SOMM, MD, and JOSÉ ROBERTO GOLDIM, PhD

Despite its high prevalence, nonulcer dyspepsia is still difficult to study, due to the lack of adequate tools to measure significant outcomes. The objective of this study was to develop and validate a symptom-focused, disease-specific questionnaire to evaluate patients with nonulcer dyspepsia. For that, the questionnaire was carefully written following widely accepted terminology, so as to facilitate translation and validation in other languages and cultures. The questionnaire was developed using Rome I terminology for symptoms, which were evaluated according to their intensity, duration, and frequency when applicable. Thirty-one patients with nonulcer dyspepsia, as well as 31 sex-and age-matched volunteers without digestive problems were used to assess the internal consistency, reproducibility, responsiveness, content validity, and discriminant validity of the questionnaire. Another 31 functional dyspeptic patients were enrolled for assessment of criterion validity. Cronbach's α coefficient was 0.82. The intraclass correlation coefficient for the scores obtained 7 days apart was 0.86. The mean score obtained after 3 months of treatment was 16.4, vs. 23.03 at baseline ($P = 0.001$). Two blinded gastroenterologists agreed that the questionnaire adequately evaluated nonulcer dyspepsia. The median symptoms score for controls was 0, vs. 22.5 for dyspeptic patients ($P = 0.001$). An inverse correlation was observed between quality of life and dyspeptic symptoms ($R = -0.28$, $P = 0.026$). The proposed questionnaire has high degrees of both reproducibility and responsiveness. As this questionnaire was based on Rome I International Consensus terminology, it is expected that it will be easy to translate and validate.

KEY WORDS: dyspepsia; questionnaires; questionnaire design; reproducibility of results.

Dyspepsia is a very prevalent condition that affects about 25% of the world population, with an annual incidence of about 1%. It is defined as pain or discomfort located in the upper abdomen (1). The symptoms of dyspepsia include pain, postprandial fullness, early satiety, nausea, vomiting, and bloating. Most affected individuals do not present either structural or biochemical alterations that explain these symptoms and are, thus,

Manuscript received May 23, 2004; accepted July 6, 2004.

From the Graduate Program in Gastroenterology, Universidade Federal do Rio Grande do Sul, Hospital de Clínicas de Porto Alegre, Port Alegre, RS, Brazil.

Address for reprint requests: Dr. Guilherme Becker Sander, Ramiro Barcellos, 1660 apto. 501, CEP 90035–002, Porto Alegre, RS, Brazil; guilhermesander@hotmail.com.

classified as presenting nonulcer dyspepsia or functional dyspepsial (1). The pathophysiology of nonulcer dyspepsia is poorly understood; it may be associated with dysfunction in motility, secretion, and sensitivity of the digestive tract (2–4).

The diagnostic criteria for nonulcer dyspepsia were established by specialists in consensus statements known as the Rome criteria (1, 5). According to these criteria, nonulcer dyspepsia is diagnosed if the following symptoms are present for at least 12 weeks (not necessarily consecutive) during the past 12 months: (i) persistent or recurrent dyspepsia; (ii) no evidence); of organic disease that is likely to explain the symptoms (including endoscopic evidence; and (iii) no evidence that relief of dyspeptic symptoms is associated exclusively with defecation or that dyspepsia is associated with changes in the frequency or stool form.

Despite its great prevalence and the associated socioeconomic cost, nonulcer dyspepsia was not adequately studied until the end of the 20th century. This resulted in part from the dearth of available research tools. Being a disorder without defined structural or pathophysiological anomalies, the outcomes of studies focusing on the efficacy of nonulcer dyspepsia treatments are necessarily subjective. Thus, the development of methods capable of measuring symptoms and patients' perceptions in a reliable and reproducible manner becomes essential. This has been achieved through the development of standardized questionnaires.

When elaborating a questionnaire, it is necessary to select the most representative set of symptoms associated with the disorder to be investigated (6). To be practical, the questionnaire must be planned to enable the largest amount of information to be obtained through the smallest possible number of questions. A questionnaire should also be sufficiently general to maintain its basic characteristics when used in different populations. Concerns about future translations should be kept in mind, as increasingly more multicentric studies are carried out. The challenge is to not create a questionnaire that is easy to translate but, rather, to assure comparability of results between different ethnic and cultural populations (7).

After it is devised, the questionnaire must be validated, that is, it must be tested to ensure that it is useful for clinical research. The properties of clinical measurement (known as clinimetric properties) that are considered essential for the validation of a questionnaire are internal consistency, reproducibility, responsiveness, and validity (8).

It is expected that the several items in a questionnaire will have at least a moderate level of relationship with one another and with the total score—so as to measure different aspects of the same condition, and not different characteristics of diverse entities. This attribute is known as internal consistency and can be estimated using Cronbach's coefficient (9)—a very high coefficient suggests redundancy, whereas a low coefficient suggests that the scale is evaluating more than one construct, the ideal values falling between 0.7 and 0.9 (10). Another tool used to accomplish this is factorial analysis. Factorial analysis has a fundamental role in empirically determining the dimensionality of items. A high degree of correlation between items suggests that all items measure the same concept. After this analysis, items found to have a low correlation with the total score may be excluded from the scale (11, 12).

Reproducibility aims at demonstrating that the data obtained are reliable. This feature can be measured by means of the "test-retest" method (13), according to which similar scores must be obtained in two distinct occasions without any treatment interventions between the two measurements. Testing and retesting are usually performed at an interval that is sufficiently short to ensure that there will be no changes to the underlying disease (14) while, at the same time, sufficiently long to avoid memory bias (patient responding on the second measurement thinking about the answers given the first time around and not really reassessing the symptoms) (15).

Responsiveness expresses the sensitivity of the questionnaire to demonstrate changes if they actually occur[16]. This property is tested by applying the instrument before and after a treatment intervention that has proven effects on the disorder under study.

Validity refers to the ability of an instrument to measure exactly what it is meant to measure (8). It can be assessed in different ways. Content analysis is not a statistical approach but, rather, a judgment by specialists in the field about representativeness and relevance of the items proposed in the scale (11). Another test is the extreme groups analysis—that is, by comparing a population with the disorder with a population without the disorder so as to assess the ability of the instrument to reveal the difference between the two populations. This kind of validity is also known as "discriminant validity" (17, 18). Another way to assess validity is by measuring "criterion validity." In this case, an instrument of proven efficacy is applied in parallel with the instrument to be tested, and a good correlation must be obtained between the results of both questionnaires. If a scientifically established tool is not available, a related construct may be used, although in this case smaller correlation coefficients should be expected (18, 19).

The aim of our study was to develop and validate an easy-to-translate, symptom-focused (unidimensional) questionnaire to assess nonulcer dyspepsia. We chose a unidimensional questionnaire in this context for two

TABLE 1. CHARACTERISTIC SYMPTOMS OF NONULCER DYSPEPSIA ASSESSED BY THE PORTO ALEGRE DYSPEPTIC SYMPTOMS QUESTIONNAIRE

Symptom	*Score*
Pain in upper abdomen	
Intensity	0–5
Duration	0–3
Frequency	0–4
Nausea	
Intensity	0–5
Duration	0–3
Frequency	0–4
Vomiting	
Frequency	0–4
Upper abdominal bloating	
Intensity	0–5
Duration	0–3
Frequency	0–4
Early satiety	
Frequency	4
Total	44

reasons. First, quality-of-life questions are more susceptible to cultural adaptation errors in future translations. Second, the activity of nonulcer dyspepsia is strongly associated with quality of life. The fact that it is often difficult to determine the causal direction between these entities could undermine the use of quality of life as an outcome in nonulcer dyspepsia. Therefore, we developed an English questionnaire based on the symptoms described as characteristic of nonulcer dyspepsia in the Rome I Consensus. To test the translatability and clinimetric properties of the questionnaire, it was translated to Brazilian Portuguese and validated in a group of Brazilian patients with nonulcer dyspepsia.

METHODS

Development

The proposed Porto Alegre Dyspeptic Symptoms Questionnaire (PADYQ) was developed based on the symptoms described as relevant in the Rome I consensus statement (5), resulting in an 11-item form (Table 1). All questionnaires were answered with the assistance of an interviewer. The patients were asked to average their symptoms during the preceding month.

Patients

Factorial analysis was performed taking into consideration all the baseline questionnaires of patients from a study about the effects of *H. pylori* eradication on nonulcer dyspepsia symptoms ($n = 157$). All patients were selected according to Rome I criteria.

To evaluate internal consistency, reproducibility and responsiveness, 31 patients with dyspeptic symptoms according to Rome I criteria were included in the study (Group A). All were submitted to upper digestive endoscopy at Hospital de Clinicas de Porto Alegre and had normal endoscopic results.

To determine discriminant validity, 31 healthy volunteers matched by sex and age to the group of patients described above and who reported having no digestive diseases were selected and interviewed after signing an informed consent form (control group).

To evaluate criterion validity, the control group and an additional group of 31 patients with nonulcer dyspepsia according to Rome I criteria (Group B) answered the Dyspeptic Symptoms Questionnaire and the World Health Organization Quality of Life (WHOQOL-BREF) (20) questionnaire.

The protocol was approved by the Research Ethics Committee at Hospital de Clinicas de Porto Alegre and all the patients and controls signed an informed consent form.

Evaluation of Clinimetric Properties

The clinimetric properties of the proposed PADYQ were evaluated in seven ways, as follows.

Factorial Analysis. Factorial analysis using the principal components method was used to aggregate items in subscales (domains). Components with eigenvalues larger than 1 were selected. After extraction of the components, orthogonal rotation of factors was performed using Kaiser's Varimax method. Items with a factorial load higher than 0.4 on the same component and which did not have values higher than 0.4 on other components were aggregated in domains (21).

Internal Consistency. Internal consistency was calculated using Cronbach's coefficient applied to the baseline questionnaires (TO) of Group A (9). The coefficient was expected to be higher than 0.7 but not higher than 0.9, which would mean redundancy (10, 12).

Reproducibility. The patients included in the study (Group A) answered the PADYQ during their first office visit (appointment TO). A new appointment was scheduled for 7 days later (appointment T1), and patients answered the questionnaire again. The patients were not exposed to any kind of intervention between T0 and T1. For 15 patients, the same interviewer completed the questionnaires at T0 and T1. For 16 patients, a different interviewer completed the two questionnaires, for evaluation of intra-and inter-observer agreement. The correlation between scores was verified by calculating the intraclass correlation coefficient for both the total score and the different domains obtained by factorial analysis.

Responsiveness. Following the two initial evaluations (T0 and T1), the patients in Group A were randomly assigned to one of two groups: one group received 10-day courses of lansoprazole, 30 mg, clarithromycin, 500 mg, and amoxicillin, 1 g, all of them twice daily. The other group received lansoprazole, 30 mg twice a day, plus placebo tablets identical to the antibiotics in appearance and same posology. These groups were considered to be homogeneous, since there is no unequivocal evidence in the literature suggesting that one treatment has more efficacy than the other (22). A decrease of 13 to 73% in the scores was expected, reflecting the response to placebo reported in previous studies (23). A new appointment (T2) was scheduled for 3 months after the beginning of treatment, and the PADYQ was again applied by one of the investigators. Responsiveness was calculated using the *t* test for paired samples, comparing the results of T1 and T2 for the total score and for the different domains assessed by the questionnaire.

Content Validity. Content validity was assessed by two experienced gastroenterologists blinded to the purpose of the questionnaire. They were asked to determine what construct the questionnaire was supposed to measure. Then they were asked to confirm that the questionnaire sampled the full range of symptoms of nonulcer dyspepsia.

Discriminant Validity. The baseline scores (T0) obtained for dyspeptic symptoms were compared with the scores obtained for the control group. This analysis was performed using the Wilcoxon test.

Criterion Validity. The PADYQ was applied in parallel to the WHOQOL-BREF (20) in 31 controls and in 31 patients with nonulcer dyspepsia (Group B). The results were correlated using Spearman' nonparametric coefficient of correlation.

Sample Size The size of the sample required to test reproducibility was calculated expecting an intraclass correlation of 0.90 (considering an α error of 0.05 and aiming at confidence intervals of ± 0.10). According to Streiner (24), the number of subjects needed for this purpose was 31. Sample size for responsiveness was calculated to allow the use of paired t test comparison, considering an α error of 0.05, a power of 90%, an expected reduction of 20% in scores from an initial score of 25, a standard deviation of 8 at both visits, and a correlation of 0.5 among variables. The estimated sample size for responsiveness was 27.

For factorial analysis, it is recommended to analyze at least five times more questionnaires than the number of items in the scale and suggested to analyze a minimum of 100 forms (21).

For other clinimetric properties, sample size is difficult to estimate, so we used 31 patients as calculated for reproducibility and responsiveness, in agreement with other studies of questionnaire validation that used similar samples (25, 26).

RESULTS

Sample Characteristics

Of the 31 patients select, 24 were women. Sex-matching in the control group was perfect. The mean age of patients with dyspepsia was 43.1 years (Group A), and the mean age of controls was 43.0 (Pearson's correlation coefficient ≥ 0.99). Schooling was higher in controls than in patients with dyspepsia: 67% of the patients in the control group had finished high school, while 58% of the dyspeptic patients had less than 8 years of schooling.

Factorial Analysis

Among 11 items, factorial analysis identified 3 components with an eigenvalue higher than 1. The first component, with an eigenvalue of 3.077 and responsible for 24% of the total variance for the questionnaire, included the intensity, duration, and frequency of bloating, as well as the frequency of early satiety (factorial loads after Varimax rotation: 0.937, 0.876, 0.910, and 0.433, respectively). The second component, with an eigenvalue of 2.4 and responsible for 24.3% of the total variance, included the intensity, duration, and frequency of nausea and frequency of vomiting (factorial load after Varimax rotation: 0.912, 0.847, 0.873, and 0.577). The third component, with an eigenvalue of 2.224 and responsible for 21.1% of the total variance, included the intensity, duration, and frequency of upper abdominal pain (factorial load after Varimax rotation: 0.905, 0.841, and 0.867). None of the items had a factorial load higher than 0.2 on two components simultaneously (Table 2 and Figure 1). Thus, the first component was called the Bloating/Satiety domain; the second, the Nausea/Vomiting domain; and the third, the Pain domain. The sum of the variance for the three domains accounted for 70% of the variance for the questionnaire.

TABLE 2. FACTORIAL LOADS OF THE QUESTIONNAIRE ITEMS AFTER VARIMAX ROTATION

	Component		
	1	*2*	*3*
Upper abdominal pain intensity	0.090	0.032	**0.905**
Upper abdominal pain duration	−0.036	0.070	**0.841**
Upper abdominal pain frequency	0.052	0.027	**0.867**
Nausea intensity	0.034	**0.912**	0.003
Nausea duration	−0.009	**0.847**	−0.025
Nausea frequency	0.072	**0.873**	0.021
Vomiting frequency	0.142	**0.577**	0.141
Upper abdominal bloating intensity	**0.937**	0.038	0.037
Upper abdominal bloating duration	**0.876**	0.004	−0.011
Upper abdominal bloating frequency	**0.910**	0.034	−0.074
Early satiety frequency	**0.433**	0.155	0.144

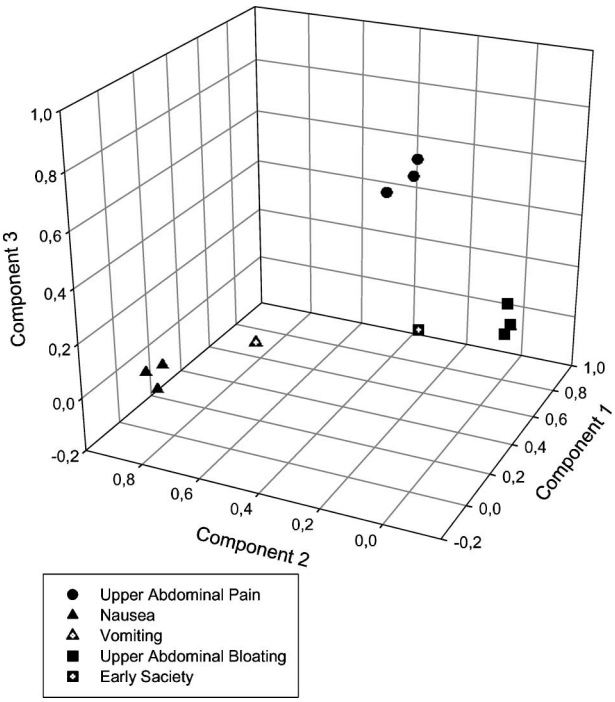

Fig 1. Factorial loads of the questionnaire items plotted after Varimax rotation.

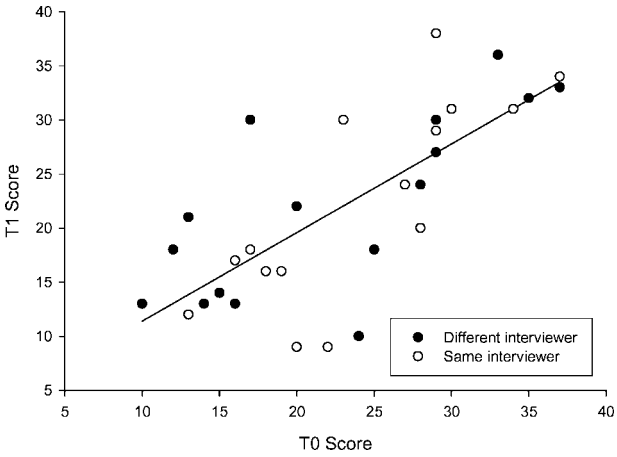

Fig 2. Correlation between scores obtained by applying the questionnaire at the T0 and T1 visits.

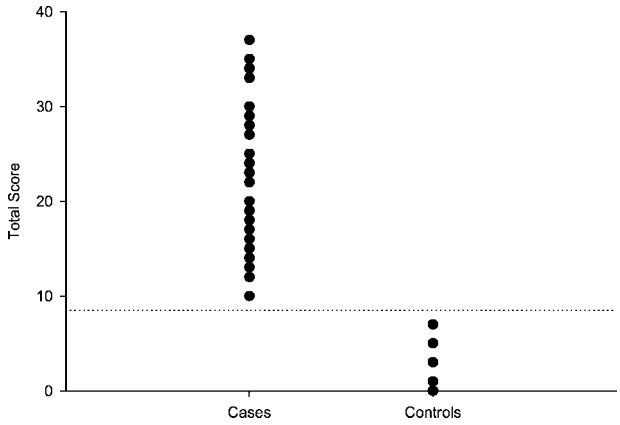

Fig 3. Total score obtained with the questionnaire for patients with dyspepsia and controls.

Internal Consistency

Cronbach's α coefficient for the 11 items on the Porto Alegre Dyspeptic Symptoms Questionnaire (PADYQ) answered by 31 patients with nonulcer dyspepsia (Group A) during appointment T1 was 0.82. Cronbach's α coefficient was 0.75 for the domain Bloating/Satiety, 0.90 for Nausea/Vomiting, and 0.83 for the Pain domain.

Reproducibility

The intraclass correlation coefficient for the total scores obtained at appointments T0 and T1 was 0.86 (Figure 2). The coefficient for interviews carried out by the same interviewer was 0.86; for interviews performed by different interviewers the coefficient was 0.87. The intraclass correlation coefficient was 0.87 for Bloating/Satiety, 0.88 for Nausea/Vomiting, and 0.68 for Pain. The mean of the total scores was 22.2 (SD, 1.56) for appointment T0 and 23.2 (SD, 1.42) for appointment T1 ($P \geq 0.2$).

Responsiveness

As stated above, the mean T1 (pretreatment) score was 23.2. After drug treatment, at the 3-month follow-up visit, the mean score was 16.0 ($P \leq 0.001$). In patients receiving lansoprazole plus placebo ($n = 10$), the mean score was 18.8 (vs. a basal score of 25.9; $P = 0.038$). In patients receiving antibiotics ($n = 21$), the mean score after 3 months was 14.67 (vs. a basal score of 20.48; $P = 0.015$). The pretreatment score for Bloating/Satiety was 9.5, vs. 6.9 after 3 months ($P \leq 0.001$), for Nausea/Vomiting the mean score was 5.36 vs. 3.6 ($P = 0.082$), and for Pain the mean score was 7.4 vs. 5.5 ($P = 0.016$).

Content Validity

The two blinded gastroenterologists determined that the questionnaire evaluated nonulcer dyspepsia. They thought the items sampled the full range of symptoms of nonulcer dyspepsia and were relevant to this disease. The clarity of questions also was considered to be adequate.

Discriminant Validity

The mean Dyspeptic Symptoms score for controls was 0.9, with a median of 0 (minimum score = 0, maximum = 7) (Figure 3). The mean score for dyspeptic patients (Group A) was 22.2, with a median of 21.0 (minimum score = 9, maximum = 38). The difference between controls and dyspeptic patients was significant ($P \leq 0.001$). A Value of $P \leq 0.01$ was obtained for each of the comparisons between the results of the three domains in the group of controls and dyspeptic patients.

Criterion Validity

There was a significant, although weak, correlation for the 31 patients with dyspepsia (Group B) and controls in terms of the scores obtained for the PADYQ and the WHOQOL-BREF ($R = -0.288$, $P = 0.023$). The correlation was negative: the higher the score obtained on the PADYQ, the lower the score on the WHOQOL-BREF. Concerning the domains of the PADYQ, the only domain presenting a significant correlation with the quality-of-life questionnaire was Bloating/Satiety ($R = -0.352$, $P = 0.005$).

DISCUSSION

The questionnaire showed excellent clinimetric properties. Cronbach's alpha, which measures internal consistency (that is, the extent to which each item is related

to other items and to the total score), was observed to be within the range considered as ideal (0.7–0.9) (15), both when the questionnaire's 11 items were analyzed together and when each domain was analyzed separately. The Porto Alegre Dyspeptic Symptoms Questionnaire (PADYQ) was also shown to be reproducible when submitted to the test–retest procedure, both by the same interviewer and by different interviewers. Its responsiveness was shown to be adequate for clinical trials, since in groups as small as 10 patients the questionnaire was capable of detecting changes in symptoms resulting from the placebo effect. The PADYQ was also capable of adequately measuring the construct it aims to evaluate, since it was capable of markedly differentiating a group of patients without digestive diseases from a group with chronic symptoms in the upper abdomen. Content validity, evaluated through qualitative interviews, showed good item clarity and relevance.

Dyspepsia is a very prevalent problem, whose treatment still lacks efficacy (27). One of the problems that prevents the identification of efficacious treatments is the difficulty in measuring the benefits achieved when a new treatment modality is proposed. The Rome II consensus statement emphasizes the need to use assessment instruments that have been validated in clinical assays evaluating functional disorders of the digestive tract (1). The application of questionnaires that were created and validated in developed countries to different socioeconomic and cultural backgrounds in the developing world is questionable, especially when taking into account quality-of-life questions and measures of access to health-care facilities (28, 29).

Some existing questionnaries, in our opinion, present methodological flaws. The questionnaire developed by van Zanten *et al*. (25) addresses symptoms not currently recognized as dyspeptic, such as heartburn, "sour taste," bad breath, and eructation. A second important limitation of the validation process for that questionnaire was the performance of an upper digestive endoscopy between the first and the second applications of the questionnaire, during the period when reproducibility was being evaluated. The strong placebo effect of endoscopies, especially when the result is normal, is well known (22).

The questionnaire validated by Buckley *et al*. in 1997 (26) is also flawed in terms of the choice of symptoms reflecting dyspepsia and, also, includes heartburn and eructation, while other important dyspeptic symptoms such as nausea are not included. However, that questionnaire has the great advantage of detailing symptoms, focusing not only on the presence or absence of symptoms, but also on features such as duration, intensity, and frequency, which allows a more efficient quantification of nonulcer dyspepsia symptoms.

Another scale that has been used in large trials of nonulcer dyspepsia is the Gastrointestinal Symptom Rating Scale (GSRS). The GSRS uses a 7-point ordinal scale to measure a huge range of symptoms, not all considered part of nonulcer dyspepsia, such as abdominal pain, reflux, indigestion, diarrhea, and constipation. In fact, the scale was initially developed to be used with irritable bowel syndrome and peptic ulcer patients, not with nonulcer dyspepsia (30).

One of the most commonly used outcome measures in large studies of nonulcer dyspepsia is the Seven-Point Likert Scale (29). It is a self-completed instrument based on only one question: "Please state each day if you have experienced pain or discomfort in the stomach." As such, the Likert Scale may be more liable to different interpretations by each patient, a shortcoming reflected in a lower than expected reliability coefficient (29, 31). This is especially true if the scale is applied to populations whose schooling level is different from that of the original population in which the questionnaire was validated. Also, this outcome measure evaluates the symptoms occurring within a 7-day interval, which may be too short to assess a chronic disease whose symptoms vary in intensity (23). Finally, with the scoring system ranging only from 1 to 7, the discriminant power is very limited (31).

Other questionnaires (Glasgow score, for example) focus on the effect of dyspepsia on the life of patients, with little emphasis on dyspeptic symptoms (28). The Glasgow questionnaire was developed in Scotland, within an organized health system, and includes questions about the number of medical visits to the patient's home, number of medical consultations in the past 6 months, and use of nonprescription and prescription medication, which are likely to have a much smaller discriminant effect in other populations. Another problem is that responsiveness was determined in a study evaluating patients with peptic ulcer, and therefore, to our knowledge, the validation of the Glasgow score for functional dyspeptic patients has not been fully demonstrated (23).

When the PADYQ was created, the Rome II Consensus statement had not been released, so the questionnaire was based on Rome I criteria. In fact, the Rome II Consensus introduced few changes concerning the diagnostic criteria and range of symptoms considered to be associated with nonulcer dyspepsia. One notable exception is vomiting, present as one of the main dyspeptic symptoms in Rome I but not in Rome II. However, we found a strong correlation between vomiting and nausea in functional dyspeptic patients, corroborating the idea that vomiting should

be considered a dyspeptic symptom. Additionally, factorial analysis findings could help in the understanding and further study of the different pathophysiologic mechanisms involved in nonulcer dyspepsia.

Taking into consideration the multiple causes of nonulcer dyspepsia, as well as the several aspects that influence quality of life, it was expected that dyspeptic symptoms would not be closely related to quality of life. However, the negative and statistically significant correlation between the PADYQ and the WHOQOL-BREF supports the notion that the questionnaire is an adequate instrument, since there have been several publications that report an interrelation (although partial) between nonulcer dyspepsia and quality of life. Another interesting finding, which should be further investigated, is that the domain that most affects quality of life in nonulcer dyspepsia patients is that of Bloating/Satiety.

CONCLUSIONS

The Porto Alegre Dyspeptic Symptoms Questionnaire (PADYQ) has adequate clinimetric features for use in scientific investigation. This questionnaire was shown to have adequate internal consistency and reproducibility when assessed by both the same and different investigators. It was also shown to be sensitive to changes following drug therapy and capable of measuring the correct construct. Therefore, it is an important research tool for the investigation of nonulcer dyspepsia. In addition, since it was based on internationally accepted terminology and on information that is not dependent on sociocultural background, we believe that the PADYQ is easy to translate and that it will maintain its clinimetric properties in other languages.

APPENDIX: PORTO ALEGRE DYSPEPTIC SYMPTOMS QUESTIONNAIRE

Concerning the past 30 days:

PAIN

How do you describe the intensity of upper abdominal pain on most of the days during this period? (□)
0. Absent
1. Very mild
2. Mild
3. Moderate
4. Severe
5. Very severe

What has been the duration of pain on most of the days during this period? (□)
0. Does not apply
1. A few minutes (less than 30 minutes)
2. Less than 2 hours
3. More than 2 hours

How often did you feel upper abdominal pain in the past 30 days? (□)
0. Does not apply
1. Seldom
2. 1 to 2 days per week
3. Almost daily
4. Daily

Total Score for Upper Abdominal Pain: —— (maximum 12 points)

NAUSEA

How intense was your nausea on most days during this period?(□)
0. Absent
1. Very mild
2. Mild
3. moderate
4. Severe
5. Very severe

What has the duration of most nausea episodes been? (□)
0. Does not apply
1. A few minutes (less than 30 minutes)
2. Less than 2 hours
3. More than 2 hours

How often did you feel nausea in the past 30 days?(□)
0. Does not apply
1. Seldom
2. 1 to 2 days per week
3. Almost daily
4. Daily

VOMITING

How often did you vomit in the past 30 days?(□)
0. Does not apply
1. Seldom
2. 1 to 2 days per week
3. Almost daily
4. Daily

Total Score for Nausea/Vomiting: —— (maximum 16 points)

UPPER ABDOMINAL BLOATING

How intense has your feeling of upper abdominal bloating (distension sensation/fullness) been in the past 30 days? (□)
0. Absent
1. Very mild
2. Mild
3. Moderate
4. Severe
5. Very severe

How long have these episodes lasted during this period?(□)
0. Does not apply
1. A few minutes (less than 30 minutes)
2. Less than 2 hours
3. More than 2 hours

How often have you experienced upper abdominal bloating/fullness in the past 30 days?(□)
0. Does not apply
1. Seldom
2. 1 to 2 days per week

3. Almost daily
4. Daily

EARLY SATIETY

In the past 30 days, how often have you felt that your stomach is full right after you start eating?(□)

0. Does not apply
1. Seldom
2. 1 to 2 days per week
3. Almost daily
4. Daily

Total Score for Upper Abdominal Bloating/Early Satiety: —— (maximum 16 points)

TOTAL SCORE FOR DYSPEPTIC SYMPTOMS: —— (maximum 44 points)

ACKNOWLEDGMENTS

We thank the following individuals for their contribution to this paper: Sérgio Gabriel Barros, MD, PhD, Júlio Pereira Lima, MD, PhD, Carisi Polanczyk, MD, ScD, Maria Helena Itaqui Lopes, MD, ScD, and Cleber Kruel, MD, PhD.

REFERENCES

1. Talley N, Stanghellini V, Heading RC, Koch KL, Nalagelada J, Tytgat G. Functional gastroduodenal disorders. In *The Functional Gastrointestinal Disorders*, 2nd ed. Drossman DA, Corazziari E, Talley N, Thompson WG, Whitehead WE (eds). Lawrence, Degnon Associates, 2000 pp. 299–350
2. Holtmann G, Gschossmann J, Neufang-Huber J, Gerken G, Talley NJ: Differences in gastric mechanosensory function after repeated ramp distensions in non-consulters with dyspepsia and healthy controls. Gut 47(3):332–336, 2000
3. Mearin F, Cucala M, Azpiroz F, Malagelada JR: The origin of symptoms on the brain-gut axis in functional dyspepsia. Gastroenterology 101(4):999–1006, 1991
4. Fennerty MB: Is nonulcer dyspepsia related to Helicobacter pylori infection? Semin Gastrointest Dis 12(3):180–185, 2001
5. Talley NJ, Colin-Jones D, Koch K, Koch M, Nyren O, Stanghellini V: Functional dyspesia: A classification with guidelines for diagnosis and management. Gastroenterol Int 4:145–160, 1991
6. Kirschner B, Guyatt G: A methodological framework for assessing health indices. J Chronic Dis 38(1):27–36, 1985
7. Sperber AD: Translation and validation of study instruments for cross-cultural research. Gastroenterology 126:s12.4–128, 2004
8. Streiner DL, Norman GR: *Health Measurement Scales*, 2nd ed. Oxford, Oxford Medical, 1995
9. Cronbach L: Coefficient alpha and the internal structure of tests. Psychometrika 16(3):297–334, 1951
10. Streiner DL, Norman GR: Selecting the items. In *Health Measurement Scales*, 2nd ed. Oxford, Oxford Medical, Publications, 1995.
11. Fachel J, Carney S: Avlição psicométrica: A qualidade das medidas e o entendimento dos dados. In: Cunha J (ed). *Psicodiagnstico V.* Porto Alegre, Artmed Editora, 2000
12. Shaw M, Talley NJ, Adlis S, Beebe T, Tomshine P, Healey M. Development of a digestive health status instrument: Tests of scaling assumptions, structur and reliability in a primary care population. Aliment Pharmacol Ther 12(11):1067–1078, 1998
13. O'Leary MP, Barry MJ, Fowler FJ Jr: Hard measures of subjective outcomes: Validating symptom indexes in urology. J Urol 148(5):1546–1548, 1992
14. Talley NJ, Verlinden M, Jones M: Quality of life in functional dyspepsia: Responsiveness of the Nepean Dyspepsia Index and development of a new 10-item short form. Aliment Pharmacol Ther 15(2):207–216, 2001
15. Streiner DL, Norman GR: Health Measurement Scales, 2nd ed. Oxford, Oxford Medical, 1995
16. Liang MH: Evaluating measurement reponsiveness. J Rheumatol 22(6):1191–1192, 1995
17. Fleck M, Louzada S, Xavier M, Chachamovich E, Vieira RR, Santos L, Pinzon V: Aplicação da versão em português do instrumento de avaliação de qualidade de vida da Organização Mundial da Saúde (WHOQOL-100). Rev Saude Publica 33(2):198–205, 1999
18. Chassany O, Marquis P, Scherrer B, Read NW, Finger T, Bergmann JF, Fraitag B, Geneve J, Caulin C: Validation of a specific quality of life questionnaire for functional digestive disorders. Gut 44(4):527–533, 1999
19. Streiner DL, Norman GR: Validity: In Health Measurement Scales, 2nd ed. Oxford, Oxford Medical, 1995
20. Fleck MP, Louzada S, Xavier M, Chachamovich E, Vieira G, Santos L, Pinzon V: [Application of the Portuguese version of the abbreviated instrument of quality life WHOQOL-bref]. Rev Saude Publica 34(2):178–183, 2000
21. Almeida TL: Anlise fatorial. http://lula.dmat.furg.b/∼taba/posanafat.htm. Accessed 6 Aug 2003
22. van Zanten SJ, Talley NJ, Blum AL, Bolling-Sternevald E, Sundin M, Junghard O: Combined analysis of the ORCHID and OCAY studies: Does eradication of Helicobacter pylori lead to sustained improvement in functional dyspepsia symptoms? Gut 50 (Suppl 4):iv26–iv30, 2002
23. van Zanten SJ: Assessment of outcome in dyspepsia: Has progress been made? Gut 50 (Suppl 4):iv23–iv25, 2002
24. Streiner DL: Sample size formulae for parameter estimation. Percept Mot Skills 78:275–284, 1994
25. van Zanten SJ, Tytgat KM, Pollak PT, Goldie J, Goodacre RL, Riddell RH, Hunt RH: Can severity of symptoms be used as an outcome measure in trials of nonulcer dyspepsia and Helicobacter pylori associated gastritis? J Clin Epidemiol 46(3):273–279, 1993
26. Buckley MJ, Scanlon C, McGurgan P, O'Morain CA: A validated dyspepsia symptom score. Ital J Gastroenterol Hepatol 29(6):495–500, 1997
27. Drossman DA: Chronic abdominal pain (with emphasis on functional abdominal pain syndrome). In *Feldman, Sleisenger & Fordtran's Gastrointestinal and liver disease*, 7th ed. Feldman M, Scharschimidt BF, Sleisenger MH (eds). Philadelphia, W. B. Saunders, 2002
28. El Omar EM, Banerjee S, Wirz A, McColl K: The Glasgow Dyspepsia Severity Score—A tool for the global measurement of dyspepsia. Eur J Gastroenterol Hepatol 8(10):967–971, 1996
29. Junghard O, Lauritsen K, Talley NJ, Wiklund IK: Validation of seven graded diary cards for severity of dyspeptic symptoms in patients with non ulcer dyspepsia. Eur J Surg Suppl 583:106–111, 1998
30. Svedlund J, Sjdin I, Dotevall G: GSRS—A clinical rating scale for gastrointestinal symptoms in patients with irritable bowel syndrome and peptic ulcer disease. Dig Dis Sci 33:129–134, 1988
31. Jaakkimainen RL, Boyle E, Tudiver F: Is *Helicobacter pylori* associated with nonulcer dyspepsia and will eradication improve symptoms? A meta-analysis. BMJ 319:1040–1044, 1999