Dig Dis Sci (2006) 51:1509–1515
DOI 10.1007/s10620-006-9120-6

ORIGINAL ARTICLE

Evaluation of a Gastrointestinal Symptoms Questionnaire

H. J. Bovenschen · M. J. R. Janssen · M. G. H. van Oijen · R. J. F. Laheij · L. G. M. van Rossum · J. B. M. J. Jansen

Received: 20 September 2005 / Accepted: 27 October 2005 / Published online: 22 August 2006
© Springer Science+Business Media, Inc. 2006

Abstract Questionnaires are widely used instruments to monitor gastrointestinal (GI) symptoms. However, few of these questionnaires have been formally evaluated. We sought to evaluate our GI symptoms questionnaire in terms of clarity and reproducibility. Primary care patients referred for open access *Helicobacter pylori* urea breath testing reported GI symptoms (type + severity) and demographic information by written questionnaire. In an interview, patients gave a personal description of the meaning of the GI symptoms on the questionnaire. Patients' descriptions of GI symptoms were compared with current definitions. Symptom severity scores were compared before and after, interview versus questionnaire. Of the 45 patients included, 19 (42%) described all symptoms correctly, whereas 17 (38%) described one symptom incorrectly. None of the patients made more than three mistakes. Regurgitation was the most common incorrectly described symptom (16 patients [36%]), whereas the other individual symptoms were well explained. Symptom severities before the interview, after the interview and reported by questionnaire (mean value ± SEM) were 2.1 ± 0.2, 2.1 ± 0.2, and 1.5 ± 0.2 points on a 7-point Likert scale (0–6), respectively. Mean severity reported by interview (95% CI) was 1.4 (1.3–1.5) times higher than reported by questionnaire ($P < .05$). In conclusion, the GI symptom questionnaire is understandable and has good reproducibility for measuring the presence of GI symptoms, although symptom severity is consistently rated higher when reported by interview.

H. J. Bovenschen (✉) · M. J. R. Janssen · M. G. H. van Oijen · R. J. F. Laheij · L. G. M. van Rossum · J. B. M. J. Jansen
Department of Gastroenterology and Hepatology, Radboud University Nijmegen Medical Centre,
P.O. Box 9101, NL-6500 HB, Nijmegen, The Netherlands
e-mail: j.bovenschen@derma.umcn.nl

Keywords Dyspepsia · Evaluation · Gastrointestinal symptoms · Questionnaire

Introduction

Dyspepsia refers to upper abdominal pain or discomfort, which is thought to arise in the upper gastrointestinal (GI) tract [1–3]. Besides epigastric pain, patients often complain of other upper GI symptoms, such as belching, regurgitation, heartburn, and bloating. In Western Europe and the United States about 25–40% of the general population suffers from dyspeptic symptoms at least once a year [1]. It is a common reason for consultation in general practice, with estimates ranging from 1–4% of all consultations. Of all outpatient visits, 6% concern complaints deriving from the GI tract. These symptoms increase healthcare resource utilization and are altogether a major health care problem in the community [1, 4–6].

Although dyspeptic symptoms are poor predictors of findings at upper GI endoscopy [7], questionnaires have been developed to measure the presence and severity of GI symptoms [8–14]. These questionnaires can be used to measure the symptomatic response to treatment more objectively. Furthermore, these questionnaires are widely used in trials to describe the prevalence and severity of dyspeptic symptoms in the study population as well as to evaluate treatment outcome. Hence, it is important to have reliable and reproducible GI symptom questionnaires. However, although many different questionnaires have been developed, few of them have been formally evaluated or validated. Unfortunately, this is also the case for the questionnaire used in our department.

The way in which these instruments are evaluated is controversial. Often invasive procedures, such as endoscopy, are

used as a gold standard to validate questionnaires, but this seems to be of little importance for dyspeptic symptoms, because it has already been established that dyspeptic symptoms and findings at endoscopy are poorly correlated [5–7]. In fact, in the majority of patients with dyspeptic symptoms no pathology is detected at endoscopy at all. On the other hand, reliably measuring dyspeptic symptoms is very important, because achieving adequate symptom reduction is the treatment goal in patients with dyspepsia. Therefore, numerous statistical methods have been developed for the evaluation of questionnaires in terms of reliability, reproducibility, and validity. We chose to evaluate our GI symptom questionnaire by comparing its outcomes to the results of an oral interview and by evaluating patients' comprehension of terms, such as heartburn, regurgitation, bloating, belching, abdominal pain, and epigastric pain.

The evaluated GI symptoms questionnaire is completely adapted to the Dutch language and culture. The instrument is used often and has proven to be well suited as "postal questionnaire," which is only sparsely appraised. The aim of the present study was to scrutinize our GI symptoms questionnaire in terms of clarity and reproducibility, to create an optimally reliable tool for measuring the presence and severity of GI symptoms.

Methods

From May to July 2003, consecutive patients with upper GI symptoms who visited a general practitioner's practice for an open access *Helicobacter pylori* urea breath test, were included. Patients who did not understand or did not speak the Dutch language properly were excluded.

Questionnaire

At home, all patients filled out a Dutch questionnaire. The English equivalent is depicted in Fig. 1. This questionnaire included questions about the severity of GI symptoms during the last 4 weeks, rated 0–6, where 0 meant "no complaints" and 6 represented the worst imaginable severity of that symptom. Furthermore, demographic variables, duration of symptoms, overall severity (on a visual analogue scale) and the use of medication, coffee, alcohol, and smoking were reported. Finally, *H pylori* urea breath test results were obtained.

Interview

After *H pylori* urea breath testing, all patients had a semistructured interview, conducted by a single physician (H.J.B.). For this study, six upper GI symptoms were chosen from the questionnaire, namely, abdominal pain, epigastric pain, heartburn, regurgitation, belching, and bloating. These symptoms were believed to be difficult to be described/interpreted by the patients in comparison to more straightforward GI symptoms in the list, such as nausea, vomiting, and hematemesis. First, patients were asked to report the presence and severity of these symptoms. Then, they were asked to give their own description of these symptoms. Finally, after the interview, they were asked to score the severity of these symptoms again. Patients' descriptions of the symptoms were compared with current definitions [3] as described in this section. Furthermore, symptom severities, as rated in the questionnaire and in the interview, were compared. Additionally, severity scores rated before and after the interview were compared. Ratios between symptom severity rated before the interview and severity rated by questionnaire were calculated for each symptom.

Definitions

The following current definitions were used for comparison with the patients' symptom descriptions [3]. *Abdominal pain* was considered as a broad, nonspecific term, concerning a feeling of discomfort or pain located in the upper or lower abdominal area. *Epigastric pain* was defined as severe burning pain or gnawing pain in the midepigastric area, often accompanied by other upper GI symptoms and possibly by radiation to the back. The used definition of *heartburn* was retrosternal burning pain owing to gastric acid, often progressing to the throat and becoming worse with bending over. *Regurgitation* was defined as spontaneous reflux of gastric acid or gastric contents into the esophagus and sometimes into the mouth. *Belching* was considered as the, often noisy, release of air from the stomach through the mouth. And *bloating*, or *abdominal distension*, was described as a condition in which the abdomen feels full and tight, not only after eating, and often interpreted as excessive intestinal gas. If patients' descriptions did not match these definitions at all, as judged by the interviewer, they were considered incorrect.

Statistics

All analyses were performed using SAS statistical software (SAS Institute Inc., Cary, NC) version 8.0. Paired *t*-tests, univariate, and multivariate analyses were performed to study differences in severity, compare symptom description correctness, and study confounding variables. A difference with a *P*-value of $<.05$ was considered statistically significant. Calculated means and relative risks were presented with standard errors of the mean (SEM) or 95% confidence intervals (95% CI).

Number _____ T 00 D 00

Abdominal or Epigastric symptoms

*Did you experience during the **last 4 weeks**:*		none	mild	moderate	quite a lot	severe	very severe	unbearable
1. Abdominal pain	in common	○	○	○	○	○	○	○
	postprandial	○	○	○	○	○	○	○
	fasting	○	○	○	○	○	○	○
	doesn't decline after defecation	○	○	○	○	○	○	○
2. Epigastric pain	in common	○	○	○	○	○	○	○
	during daytime	○	○	○	○	○	○	○
	at night / asleep	○	○	○	○	○	○	○
3. Heartburn		○	○	○	○	○	○	○
4. Regurgitation		○	○	○	○	○	○	○
5. Abdominal rumbling		○	○	○	○	○	○	○
6. Bloating		○	○	○	○	○	○	○
7. Empty feeling		○	○	○	○	○	○	○
8. Nausea		○	○	○	○	○	○	○
9. Vomiting		○	○	○	○	○	○	○
10. Loss of appetite		○	○	○	○	○	○	○
11. Postprandial fullness		○	○	○	○	○	○	○
12. Belching		○	○	○	○	○	○	○
13. Flatulence		○	○	○	○	○	○	○
14. Haematemesis		○	○	○	○	○	○	○
15. Dysphagia	liquid food	○	○	○	○	○	○	○
	solid food	○	○	○	○	○	○	○
16. Stools	melaena	○	○	○	○	○	○	○
	bloody	○	○	○	○	○	○	○
	mucous	○	○	○	○	○	○	○
	frequent hard	○	○	○	○	○	○	○
	diarrhea	○	○	○	○	○	○	○
	alternately solid or loose	○	○	○	○	○	○	○
	constipation	○	○	○	○	○	○	○
	frequently with pain	○	○	○	○	○	○	○
	urging stools	○	○	○	○	○	○	○
	incomplete	○	○	○	○	○	○	○
	steatorrhoea	○	○	○	○	○	○	○

18. Describe your abdominal or epigastric pain during the last 4 weeks by marking the line below with a cross. No text.

no complaints _____ unbearable symptoms

0000000000000000000 3330049989

Fig. 1 A copy of the evaluated GI symptoms questionnaire (English language)

Table 1 Patient characteristics ($n = 45$)

Male gender (%)	25 (56)
Mean age in years (SEM)	45 (2)
Dutch nationality (%)	36 (80)
General practitioner (%)	19 (42)
H pylori positive test result (%)	14 (31)
Months symptom duration (SEM)	65 (17)
Current smoker (%)	21 (47)
Alcohol use (%)	25 (56)
Coffee use (%)	40 (89)
PPI, H2RA, or antacid use (%)	24 (53)
NSAID or aspirin use (%)	6 (13)

Note. *NSAID* nonsteroidal anti-inflammatory drug, *PPI* proton pump inhibitor.

Results

Patient population

During 3 months, 45 consecutive patients were included in this study. All patients completed the written questionnaire. Twenty-five patients were male (56%). Overall, 36 (80%) patients were born in The Netherlands. Average age (± SEM) was 45 ± 2 years and mean symptom duration (± SEM) was 65 ± 17 months. Current smoking, the use of coffee, alcoholic drinks, nonsteroidal anti-inflammatory drugs (NSAIDs) or aspirin, and drugs to control GI symptoms were reported in 21 (47%), 40 (89%), 25 (56%), 6 (13%), and 24 (53%) of patients, respectively. Finally, 14 (31%) patients were *H pylori* positive, as measured by the urea breath test (Table 1).

Symptom descriptions

Overall, 19 (42%) patients correctly described all symptoms. Of those who did not, 17 (38%) described five correct, 7 (16%) made four correct descriptions, and 2 (4%) patients three correct descriptions. No patients made more than three mistakes (Fig. 2). Figure 3 shows that, in general, symptoms were well understood. Regurgitation was the most common incorrectly described symptom (16 patients [36%]).

Questionnaire versus interview

Severity score for all symptoms reported by questionnaire (mean ± SEM) was 1.5 ± 0.2. Mean severity score reported by interview, before the dialogue, was 2.1 ± 0.2. The ratio of severity (95% CI) reported by interview or by questionnaire was 1.4 (1.3–1.5), which was a statistically significant difference ($P < .01$) (Table 2; Fig. 4). Differences for the individual symptoms are also shown in Table 2.

Symptom severity before and after interview

Overall, 69% of patients had no differences in severity before and after interview. The mean severity score was 2.1 ± 0.2, both before and after the interview ($P > .05$). For the individual symptoms, there were only slight differences between reported severity before and after the interview and none of them was of statistical significance ($P > .05$) (Table 2, Fig. 4).

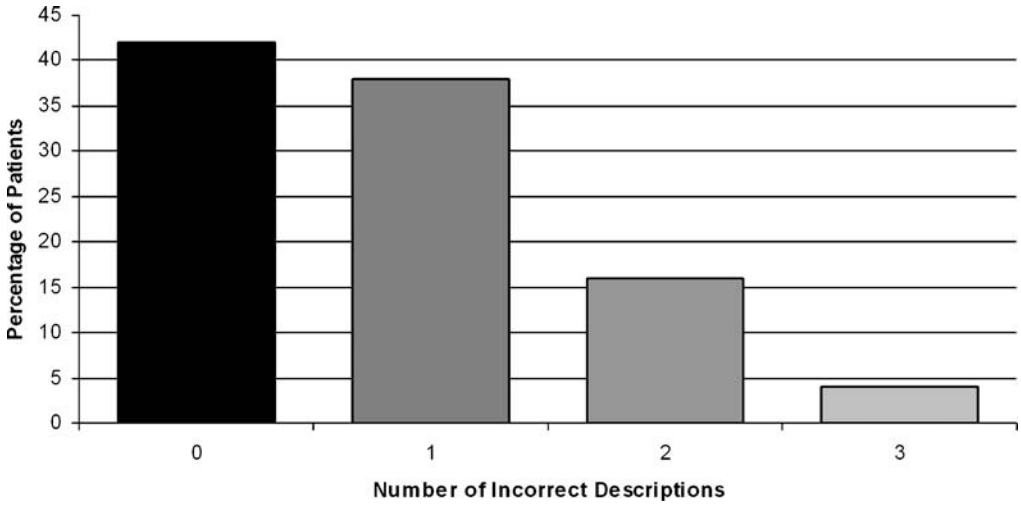

Fig. 2 Percentage of patients with incorrect descriptions of the meaning of the GI symptoms on the questionnaire

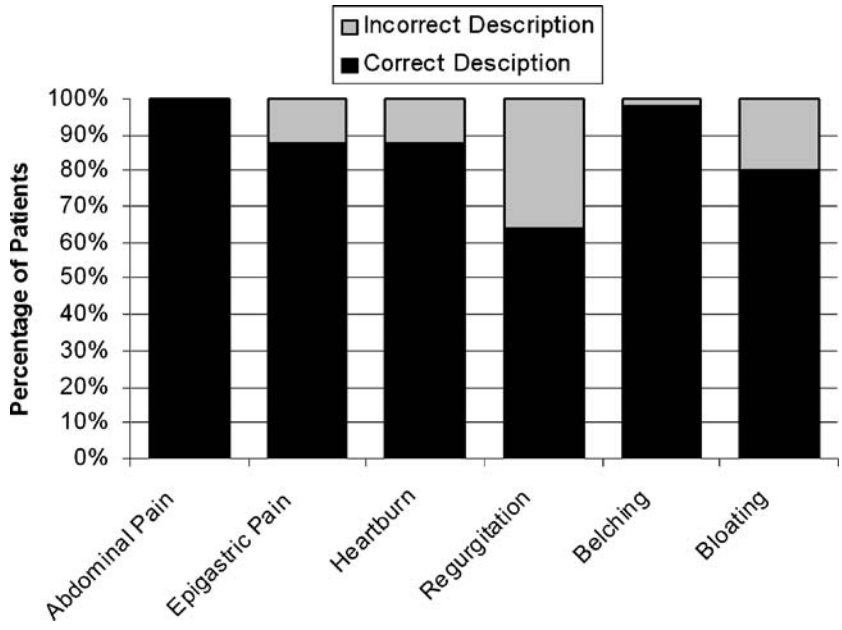

Fig. 3 Percentage of correct and incorrect symptom descriptions per individual symptom

Risk factors and associations

The following factors were scrutinized with univariate and multivariate analyses to identify potential confounders: gender, age, nationality, *H pylori* status, use of NSAIDs, use of GI medication, intoxicants (coffee, alcohol, smoking), and the duration of symptoms. In summary, coffee drinkers report higher symptom severities on our questionnaire, whereas patients with a more than 2-year history of GI symptoms report lower symptom severities ($P < .05$) (Table 3).

Discussion

Symptoms of dyspepsia are very common and have great impact on patients' well-being as well as on community health care utilization [1, 4–6]. Whereas questionnaires, being cheap and easy to use, are extensively used to measure these symptoms, but few have been evaluated [8–15].

The results of the present evaluation of our questionnaire showed that patients have acceptable understanding of the meaning of symptoms such as regurgitation, bloating, heartburn, epigastric pain, belching, and abdominal pain. Only one fifth of patients did not know the meaning of two or more items on the questionnaire. *Regurgitation* was especially hard to describe. Therefore, we will change the description of this symptom in the future. Concerning symptom severity, patients scored 37% higher when reporting symptom severity by interview than reported by questionnaire, whereas symptom severities before and after the interview were exactly the same. Hence, our questionnaire is clear and has good reproducibility, but there may be a difference in rating symptom severity by postal questionnaire or interview.

A valid questionnaire needs to meet several criteria. First, it must be understandable and easy to fill out by the population of interest. Second, it must be reproducible when administered on two separate occasions in a clinically stable

Table 2 Comparison of symptom severity reported in an interview and by questionnaire on a seven point Likert scale (0–6)

Symptom	Mean severity reported by interview and by questionnaire			Severity ratio
	Before interview	After interview	Questionnaire	Oral/written
Abdominal pain	1.3	1.2	0.9	1.4
Epigastric pain	2.7	2.9	1.8	1.5
Heartburn	2.3	2.3	1.8	1.3
Regurgitation	1.8	1.7	1.4	1.3
Belching	2.1	2.1	1.6	1.3
Bloating	2.1	2.1	1.6	1.3
Mean severity (SEM)	2.1 (0.2)	2.1 (0.2)	1.5 (0.2)	1.4 (0.1)*

*Statistically significant ($P < .01$).

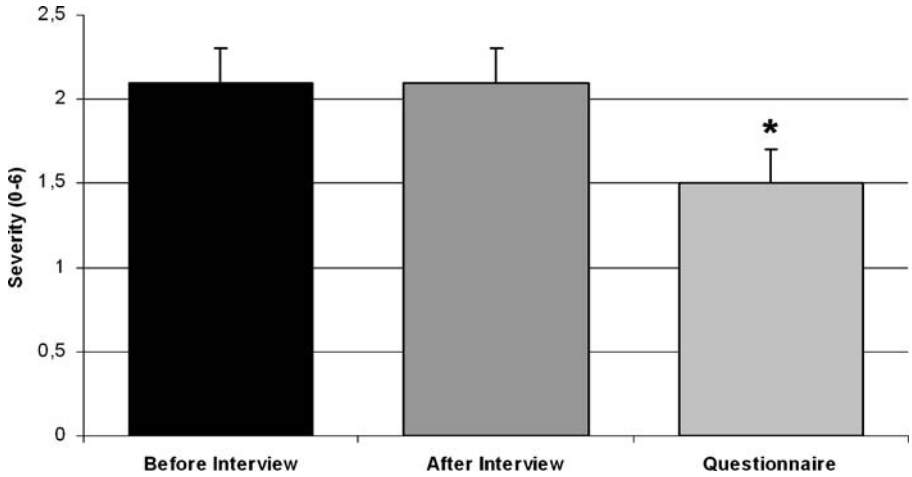

Fig. 4 Mean overall severity of GI symptoms before the interview, after the interview and on the questionnaire. *Statistically significant ($P < .01$)

patient. Third, it must discriminate for the disorder under study and it give a true picture of the situation. Finally, the questionnaire should be sensitive to changes in the grade of dysfunction [14, 15].

In the present study, our main focus was to assess clarity and reproducibility of our translated questionnaire to get a base to further expand our activities to fully validate the instrument.

With respect to the results, the following was notable: our questionnaire was easily completed, as it was self-administered at home in less than 10 min and had a 100% response rate. The questionnaire was understandable, according to the patients, because patients filled out all items. In contrast, symptom descriptions were not always correct as shown in the present study, but most of them were quite accurate. For example, *regurgitation* was hard to describe. Most patients who described this symptom incorrectly regarded regurgitation the same as belching. Because our questionnaire has constantly been revised over several years, this symptom could be better described on the questionnaire itself to receive more reliable data.

Before and after interviewing the patients, severity scores were equal. Patients tended to rate symptom severity higher during the interview than on the written self-administered questionnaire. This effect was seen for all individual symptoms. It seems to be a general tendency in this study, and probably beyond this study. It is of note that the time between filling out the questionnaire and the interview was longer than the time between the two separate ratings during the interview. This might have influenced the mentioned difference. Reproducibility of the orally reported symptom severity was excellent. Factors that could possibly interfere with the results have been found to have only minor influence on the outcomes. Coffee use and symptom duration have

Table 3 Factors of importance for correct descriptions and severity outcomes

	Description correctness*		Severity by interview*		Severity by questionnaire*	
Factor	Relative risk (95% CI)***	*P*-value	Mean change**	*P*-value	Mean change**	*P*-value
Male gender	1.0 (0.1–6.4)	.97	0.6	.37	0.4	.38
Age >50 y	1.7 (0.2–16.5)	.65	− 1.3	.18	− 0.8	.18
Dutch nationality	4.1 (0.4–39.7)	.86	− 0.5	.50	− 0.7	.20
H. pylori positivity	0.9 (0.2–4.3)	.22	0.8	.20	− 0.3	.49
Use of NSAIDs	2.3 (0.1–64.1)	.61	1.3	.27	1.1	.15
Use of GI medication	0.2 (0–1.0)	*<.05*	0.7	.21	0.7	.06
Alcohol use	0.7 (0.1–4.7)	.75	− 0.9	.18	− 0.6	.17
Coffee use	0.3 (0.0–3.9)	.38	0.9	.38	1.4	*<.05*
Current smoking	2.2 (0.4–11.9)	.38	0.6	.39	0.4	.35
Symptom duration >2 years	4.6 (0.7–29.1)	.11	− 1.3	.07	− 1.0	*<.05*

Note. *GI* gastrointestinal, *NSAID* nonsteroidal anti-inflammatory drug.

*After adjustment for all variables presented in this table.

**Mean change in points on a 7-point Likert scale (0–6) when the factor mentioned in column 1 is present.

***Relative risk for describing the symptoms correctly (description correctness).

already been described to influence GI symptom severity [16, 17].

In the present study, a start was made to fully validate our GI symptoms questionnaire. To assess the instrument completely further studies with a larger patient group and complementary methods are needed. The questionnaire should be given to other researchers and other patient groups, including multiple countries and different languages. All symptoms have to be assessed, including lower GI symptoms. Furthermore, patient follow-up data could well serve to assess our questionnaire's sensitivity to changes and reproducibility in the near future. This should comprise testing the questionnaire's intrapatient results at two different moments in time, for example, before and after treatment. This should elucidate sensitivity to changes. A gold standard, such as endoscopy, is less important, because questionnaires are only modestly predictive of underlying pathology. Monitoring symptom and symptom severity reduction during active treatment is the main purpose using our questionnaire in patients with dyspepsia.

Although small number of patients was included, no higher numbers were needed for this evaluative study because of paired sampling. The questionnaire was already tested for understandability in Dutch language. In the present study, we chose for one single evaluator, to avoid interobserver variability. It might be interesting in the future to take into account multiple independent judgments. Also, the definitions are evidently debatable, but used in a standardized manner to obtain the most reliable results. In general, our method of evaluation of the instrument is innovative and provides further insight in patients' understanding of medical language.

In conclusion, our GI symptoms questionnaire is understandable and has good reproducibility for measuring GI symptoms. Further studies are needed to completely validate the questionnaire.

References

1. Knill-Jones RP: Geographical differences in the prevalence of dyspepsia. Scand J Gastroenterol 26 (Suppl 182):17–24 1991
2. Garatt AM, Ruta DA, Russell I, MacLeod K, Brunt P, McKinlay A, Mowat A, Sinclair T: Developing a condition-specific measure of health for patients with dyspepsia and ulcer-related symptoms. J Clin Epidemiol 49:565–571 1996
3. Malfertheiner P: Current concepts in dyspepsia: a world perspective. Eur J Gastroenterol Hepatol 11 (Suppl 1):S25–29 1999
4. Guillemot F, Ducrotte P, Bueno L: Prevalence of functional gastrointestinal disorders in a population of subjects consulting for gastroesophageal reflux disease in general practice. Gastroenterol Clin Biol 29:243–246 2005
5. Wiklund I, Glise H, Jerndal P, Carlsson J, Talley NJ: Does endoscopy have a positive impact on quality of life in dyspepsia? Gastrointest Endosc 47:449–454 1998
6. Dimenäs E, Glise H, Hallerbäck B, Hernqvist H, Svedlund J, Wiklund I: Well-being and gastrointestinal symptoms among patients referred to endoscopy owing to suspected duodenal ulcer. Scand J Gastroenterol 30:1046–1052 1995
7. Thomson AB, Barkun AN, Armstrong D, Chiba N, White RJ, Daniels A, Escobedo S, Chakraborty B, Sinclair P, Veldhuyzen van Zanten SJO: The prevalence of clinically significant endoscopic findings in primary care patients with uninvestigated dyspepsia: the Canadian adult dyspepsia empiric treatment—prompt endoscopy (CADET-PE) study. Aliment Pharmacol Ther 17:1481–1491 2003
8. Ofman JJ, Shaw M, Sadik K, Grogg A, Emery K, Lee J, Reyes E, Fullerton S: Identifying patients with gastroesophageal reflux disease: validation of a practical screening tool. Dig Dis Sci 47:1863–1869 2002
9. Bardhan KD, Stanghellini V, Armstrong D, Berghöfer P, Gatz G, Mönnikes H: International validation of ReQuest™ in patients with endoscopy-negative gastro-oesophageal reflux disease. Aliment Pharmacol Ther 20:891–898 2004
10. Bardhan KD, Stanghellini V, Armstrong D, Berghöfer P, Gatz G, Mönnikes H: Evaluation of GERD symptoms during therapy. Part I. Development of the new GERD questionnaire ReQuest™. Digestion 69:229–237 2004
11. Kuykendall DH, Rabeneck L, Campbell JM, Wray NP: Dyspepsia: how should we measure it? J Clin Epidemiol 51:99–106 1998
12. Kennedy T, Jones R: Development of a postal health status questionnaire to identify people with dyspepsia in the general population. Scand J Prim Health Care 13:243–249 1995
13. Agréus L, Svärdsudd K, Nyrén O, Tibblin G: Reproducibility and validity of a postal questionnaire. The abdominal symptom study. Scand J Prim Health Care 11:252–262 1993
14. Österberg A, Graf W, Karlbom U, Pahlman L: Evaluation of a questionnaire in the assessment of patients with faecal incontinence and constipation. Scand J Gastroenterol 31:575–580 1996
15. Talley NJ, Phillips SF, Melton L3rd , Wiltgen C, Zinsmeister AR: A patient questionnaire to identify bowel disease. Ann Intern Med 111:671–674 1989
16. Ruth M, Finizia C, Lundell L: Occurrence and future history of oesophageal symptoms in an urban Swedish population: Results of a questionnaire-based, ten-year follow-up study. Scand J Gastroentrerol 40:629–635 2005
17. DiBaise JK: A randomized, double-blind comparison of two different coffee-roasting processes on development of heartburn and dyspepsia in coffee-sensitive individuals. Dig Dis Sci 48:652–656 2003