Aliment Pharmacol Ther 1998; 12: 1257–1262.

The Leeds Dyspepsia Questionnaire: a valid tool for measuring the presence and severity of dyspepsia

P. MOAYYEDI*, S. DUFFETT*, D. BRAUNHOLTZ†, S. MASON‡, I. D. G. RICHARDS†, A. C. DOWELL† & A. T. R. AXON*

**Gastroenterology Unit, Centre for Digestive Diseases, General Infirmary at Leeds, Leeds, UK; †Epidemiology and Primary Care, University of Leeds, Leeds, UK; and ‡Clinical Trials and Research Unit, Arthington House, Cookridge, Leeds, UK*

Accepted for publication 21 June 1998

SUMMARY

Background: There is currently no validated questionnaire that assesses both the presence and severity of dyspepsia.

Aim: To develop the Leeds Dyspepsia Questionnaire (LDQ) as a measure of the presence and severity of dyspepsia, and to assess the validity, reliability and responsiveness of this instrument.

Methods: Unselected patients attending either a hospital dyspepsia clinic or a general practice surgery were interviewed by a trained gastroenterologist or a general practitioner on the presence and severity of dyspepsia. This opinion was compared with the results of the nurse-administered LDQ. Test–retest reliability was assessed by the same research nurse re-administering the LDQ 4–7 days after the initial visit in a subgroup of hospital patients. In a further subgroup of patients one researcher interviewed the patients and a second researcher re-administered the LDQ within 30 min to evaluate inter-rater reliability. The responsiveness of the LDQ was measured by repeating it in patients with endoscopically proven peptic ulcer or oesophagitis 1 month after receiving appropriate therapy.

Results: The LDQ was administered to 99 general practice and 215 hospital patients. In the GP population 41/98 (42%) had dyspepsia according to the GP and the LDQ had a sensitivity of 80% (95% CI: 65–91%) and a specificity of 79% (95% CI: 66–89%). The weighted kappa statistic for the agreement between the LDQ and the clinician for the severity of dyspepsia was 0.58 in the GP population and 0.49 in hospital patients. The kappa statistic for test–retest reliability was 0.83 in 107 patients. The LDQ had excellent inter-rater reliability with a kappa statistic of 0.90 in 42 patients. The median LDQ score fell from 22.5 (range 9–36) to 4.5 (range 0–27) in 12 patients 1 month after receiving appropriate therapy (Wilcoxon signed rank test, $P < 0.0001$).

Conclusion: The LDQ is a valid, reliable and responsive instrument for measuring the presence and severity of dyspepsia.

INTRODUCTION

Dyspepsia is a common problem affecting 30% of the population and accounting for 3–4% of all general practice consultations.[1] The investigation and treatment of dyspepsia is a significant burden on health service resources with antisecretory drugs being the single most expensive item on the NHS pharmacy budget.[2] Strategies have been suggested that might reduce the investigation and treatment of this condition but results have been inconclusive.[3, 4] One of the reasons for this is the lack of validated questionnaires for assessing dyspepsia. A dyspepsia questionnaire should accurately identify patients with dyspepsia so

Correspondence to: Dr P. Moayyedi, Gastroenterology Department, General Infirmary at Leeds, Great George Street, Leeds LS1 3EX, UK.
E-mail: paulmo@ulth.northy.nhs.uk

that appropriate subjects are enrolled into studies. It should also assess dyspepsia severity so that the effects of investigation and treatment can be measured.

Questionnaires have been developed that either measure the presence of dyspepsia[5, 6] or the severity of dyspepsia[7, 8] but to our knowledge a well validated questionnaire that measures both these parameters has not been described. Furthermore, dyspepsia questionnaires are often used in population studies but are only validated in patients attending hospital. We have developed the Leeds Dyspepsia Questionnaire (LDQ) which measures both the presence and severity of dyspepsia and have validated this tool in both secondary and primary care patients.

An instrument designed to measure dyspepsia should be validated against a 'gold standard' test. Unfortunately there is no ideal test for dyspepsia and we have therefore used the opinion of an experienced clinician as a quasi 'gold standard'. Questionnaires should also be reliable, with minimal test–retest and inter-rater variation. Items in the questionnaire should have internal consistency in that questions should relate to each other in producing an overall score. Finally, a dyspepsia questionnaire should be responsive to change if it is being used to establish the effect of interventions on symptoms.[9] We have assessed the LDQ against all these criteria.

METHODS

Development of the LDQ

A preliminary dyspepsia questionnaire was prepared by reviewing the instruments used by other investigators.[5–7, 10–18] Questions from these instruments were redefined and new items added to form a preliminary dyspepsia questionnaire. This initial questionnaire contained items relating to dyspeptic symptoms as well as to tests and treatments received. The questionnaire was administered to 20 dyspeptic patients attending general practice. The items relating to tests and treatments received for dyspepsia had poor validity compared with information from general practice note review. We therefore revised the questionnaire to include only questions concerning dyspeptic symptoms. We then circulated this to senior and junior gastroenterologists for their comments on the appropriateness and clarity of the items included. This questionnaire was validated in 50 patients attending a hospital dyspepsia clinic and questions which exhibited poor test–retest reliability or did not contribute to the evaluation of dyspepsia were excluded. The final Leeds Dyspepsia Questionnaire contained eight items, each with two stems, relating to the frequency and severity of dyspeptic symptoms over the previous 6 months, and one item on the most troublesome symptom experienced by the subject. The LDQ gave a range of scores from 0 to 40 and contained questions on epigastric pain, retrosternal pain, regurgitation, nausea, vomiting, belching, early satiety and dysphagia. The first five questions were used to determine the presence of dyspepsia whilst all eight questions were required to measure the severity of dyspepsia. The questionnaire is available on request from the corresponding author.

Evaluation of the LDQ

Unselected patients attending a hospital dyspepsia clinic were interviewed by a trained nurse using the LDQ. Patients were then interviewed by one of five gastroenterologists, each having at least 3 years of specialist experience. The gastroenterologists were informed that dyspepsia referred to symptoms attributable to the upper gastrointestinal tract occurring more than once a month for at least 6 months. They were asked to state whether the patient had dyspepsia and if this was present to grade the severity on a 5-point Likert scale from very mild to very severe. The gastroenterologists were reminded that severity referred to what extent symptoms interfered with daily activities rather than to the likelihood of underlying organic pathology. Unselected patients attending a particular general practice surgery were also interviewed by a research nurse using the LDQ. The patient was then interviewed by one of three qualified general practitioners who were given the same advice as the gastroenterologists. The diagnosis reached by the LDQ on the presence and severity of dyspepsia was compared with the opinion of the clinician.

Test–retest reliability was assessed in a subgroup of hospital patients with dyspepsia by the same research nurse administering the LDQ on two separate occasions 4–7 days apart. The research nurse was blinded to the LDQ score at the first interview and the two scores were compared. The inter-rater reliability was evaluated in a separate group of hospital dyspeptic patients. Subjects were interviewed by one researcher and the LDQ was then re-administered within 30 min by a second researcher blinded to the initial score. Internal

consistency of the LDQ was measured using Cronbach's alpha coefficient.

The responsiveness of the LDQ to the effects of treatment was assessed in patients who had oesophagitis or peptic ulcer disease at endoscopy. The LDQ was administered before endoscopy and again 1 month later, after the patient had received proton pump inhibitors and/or *H. pylori* eradication therapy.

Statistical analyses

Validity, test–retest and inter-rater reliability were evaluated using the kappa statistic. Responsiveness to change was assessed using the Wilcoxon signed rank test. Statistical analyses were carried out using Stata version 3, SPSS for Windows version 6.1 and CIA.

RESULTS

Presence of dyspepsia

The LDQ was administered to 99 general practice patients (median age 47.5 years, range 15–84; 36% male) and 215 hospital patients (median age 45 years, range 24–76; 51% male). The questionnaire was usually completed within 5 min, although mean completion times were not formally assessed. The doctor's opinion on the presence and severity of dyspepsia was missing in nine cases (one from the GP group and eight from the hospital group). In the general practice population 41/98 (42%) had dyspepsia according to the GP and the LDQ gave eight false negative and 12 false positive results compared with this gold standard (80% sensitivity, 95% CI: 65–91%; 79% specificity, 95% CI: 66–89%: kappa statistic = 0.59]. As expected for a hospital population attending a dyspepsia clinic, a large proportion (193/207, 93%) had dyspepsia according to the gastroenterologist. The LDQ gave two false negative results (99% sensitivity, 95% CI: 96–100%) and seven false positive results (53% specificity, 95% CI: 26–79%) giving a kappa statistic of 0.62.

Severity of dyspepsia

The LDQ scores were calculated by summing the severity responses of the eight different symptoms. The distribution of dyspepsia scores was positively skewed in the GP population with a large proportion of subjects having little or no dyspepsia (LDQ score 2.9 ± 3.9 (mean ± standard deviation), Figure 1). The scores were more normally distributed in the hospital dyspepsia clinic population (LDQ score 11.5 ± 6.2, Figure 2). The LDQ scores were categorized according to bands that best fitted the dyspepsia severity judged by the doctor. Those that did not have dyspepsia according to questions 1–5 scored 0. LDQ scores 1–4 were classified as very mild dyspepsia, 5–8 as mild dyspepsia, 9–15 as moderate dyspepsia, > 15 as severe or very severe dyspepsia. The kappa statistic was then used to compare the agreement between doctor and LDQ on the severity of dyspepsia using these categories. The kappa statistic was weighted as disagreements that were only one cell apart were less important than discrepancies that were two or more cells apart (weighted kappa used was as follows: no discrepancy = 1.0, one cell discrepancy = 0.9, two cell discrepancy = 0.1, three cell discrepancy = 0. The weighted kappa statistic for the GP population (Table 1) was 0.58 and for the hospital population (Table 2) was 0.49, suggesting moderately good agreement between the clinician and the LDQ.

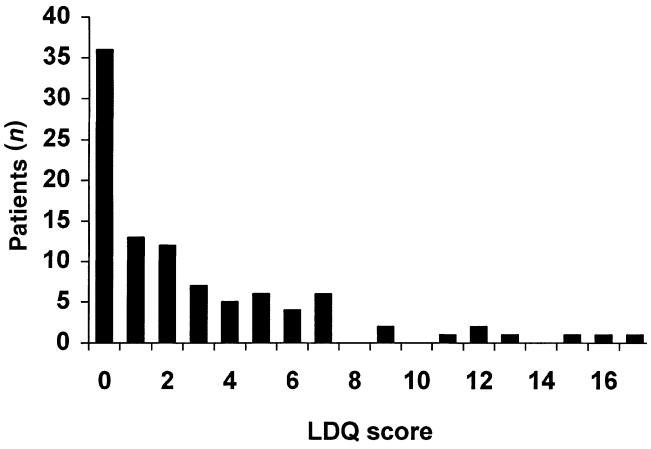

Figure 1. Range of LDQ scores in an unselected GP population.

Test–retest reliability

The last 111 hospital patients (median age 45 years, range 24–76; 43% male) that completed the LDQ were invited to return within 4–7 days to complete a further questionnaire. The LDQ was administered by the same researcher on both occasions and 107 (96%) returned to complete the LDQ for a second time. The two LDQ scores were compared and discrepancies were symmetrically distributed around zero (difference in LDQ scores 0.2 ± 3.6, Figure 3). The scores were also categorized

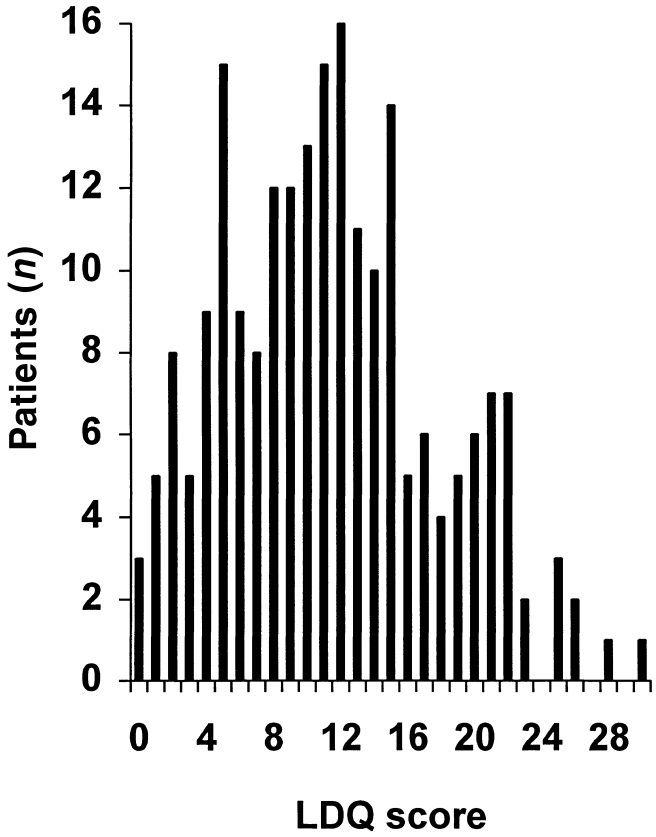

Figure 2. Range of LDQ scores in those attending an out-patient dyspepsia clinic.

from no dyspepsia to severe/very severe dyspepsia as described above and evaluated using a weighted kappa statistic. The LDQ showed excellent test–retest reliability (Table 3) with a weighted kappa score of 0.83.

Inter-rater reliability

A separate group of 42 patients was interviewed by one researcher and the LDQ was then re-administered by a second researcher within 30 min. The two LDQ scores were compared and there was excellent inter-rater reliability (Table 4) with a weighted kappa score of 0.90.

Table 1. Validity of the LDQ in diagnosing the severity of dyspepsia in an unselected general practice population

	GP population				
LDQ	no dys.	v. mild	mild	mod.	severe +
no dys.	45	5	1	2	0
v. mild	8	6	4	3	0
mild	3	4	4	4	0
mod.	1	1	1	3	1
severe +	0	0	1	1	0

Table 2. Validity of the LDQ in diagnosing the severity of dyspepsia in those attending an out-patient dyspepsia clinic

	Gastroenterologist diagnosis				
LDQ	no dys.	v. mild	mild	mod.	severe +
no dys.	8	0	2	0	0
v. mild	0	3	9	6	0
mild	0	0	22	15	6
mod.	5	1	12	50	19
severe +	2	2	3	22	21

Internal consistency

The responses of both the GP and hospital patients were pooled and the extent to which questions within the LDQ contributed to the overall score were measured using Cronbach's alpha coefficient. The LDQ had a coefficient of 0.69 which represents excellent internal reliability for a questionnaire used for research purposes.

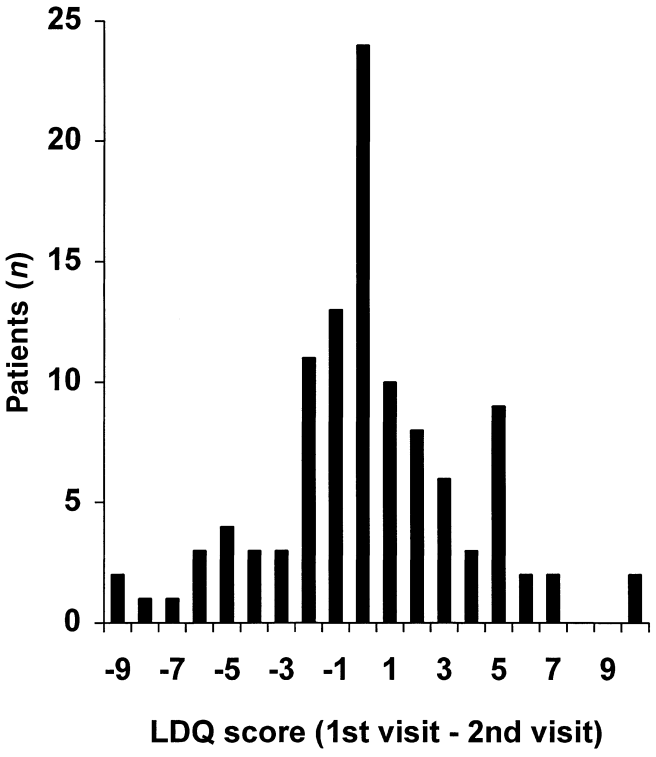

Figure 3. Test–retest variability in the LDQ.

Table 3. Test–retest reliability of the LDQ

	2nd visit				
1st visit	no dys.	v. mild	mild	mod.	severe +
no dys.	6	0	0	0	0
v. mild	0	7	0	1	0
mild	1	3	12	6	0
mod.	2	0	4	31	7
severe +	0	0	0	4	23

Responsiveness to change

Twenty-seven patients completing the LDQ had a peptic ulcer or oesophagitis at endoscopy; 12/27 (44%) of these returned at 1 month to complete a further LDQ after receiving appropriate treatment. The median LDQ score fell from 22.5 (range 9–36) to 4.5 (range 0–27) $P < 0.0001$ (Wilcoxon signed rank test) indicating that this instrument is responsive.

DISCUSSION

Dyspepsia trials rarely use a validated questionnaire to assess the impact of therapeutic interventions.[19] The data from this study suggest that the LDQ is a useful instrument for assessing the presence and severity of dyspepsia. The LDQ had a high degree of sensitivity in detecting dyspepsia compared with the opinion of a gastroenterologist or GP as a gold standard. The specificity of the LDQ observed in hospital patients was relatively low but the 95% confidence intervals are wide because only a small number of patients did not have dyspepsia. In the GP population the specificity of the LDQ was adequate, suggesting that it is a valid instrument for detecting the presence of dyspepsia. Agreement between the LDQ and clinician for the severity of dyspepsia was moderately good, with a kappa statistic similar to that reported by other disease-specific health measures.[9] Given the disagreement between individual clinicians in this area,[8] we believe these data suggest that the LDQ score is an acceptable measure of the severity of dyspepsia. This is supported by the skew towards zero in the scores recorded in the general practice subjects which would be anticipated in this population, compared with a more normal distribution in those attending a dyspepsia clinic.

Table 4. Inter-rater reliability of the LDQ

	2nd researcher				
1st researcher	no dys.	v. mild	mild	mod	severe +
no dys.	0	1	0	0	0
v. mild	0	2	0	0	0
mild	0	0	5	1	0
mod.	0	0	4	14	4
severe +	0	0	0	0	11

Dyspepsia is a term that encompasses a variety of upper gastrointestinal symptoms and it has been suggested that these can be divided into clinically important subgroups. The LDQ generates a global dyspepsia score and also provides information on the patient's worst dyspeptic symptom. In theory this questionnaire could also generate a symptom profile categorizing patients into those with ulcer-like, reflux-like, dysmotility-type or mixed dyspepsia. The value of this classification, however, has been questioned[20] and the accuracy of the LDQ in recognizing these subgroups has not been assessed in this study.

The questionnaire was designed to be administered at a face to face interview so that any questions of which the patient is uncertain can be clarified. This is more time-consuming and expensive than a self-administered questionnaire, but this is a minor consideration in interventional trials where patients are being interviewed anyway as part of the study protocol. The LDQ is simple to complete, however, and could be modified to become a self-administered questionnaire with further validation.

The inter-rater reliability of the LDQ was excellent with a kappa statistic of 0.86, which is similar to that achieved by other health measures.[9] This questionnaire can therefore be used with confidence by other researchers. The kappa statistic of 0.9 for test–retest reliability suggests that this measure can be used to make comparisons over time. This is supported by the responsiveness of the questionnaire when patients were reassessed after receiving treatment for organic lesions at endoscopy.

This questionnaire could therefore be used in epidemiological surveys assessing the frequency of dyspepsia in populations and also in clinical trials assessing the response of dyspepsia to treatment. Others have also reported rigorously validated dyspepsia questionnaires that can be used in such trials.[7, 8] To our knowledge, however, the LDQ is the first questionnaire to be validated in primary as well as secondary care populations and so is suitable for general practice dyspepsia studies.

ACKNOWLEDGEMENTS

Dr P. Moayyedi is currently funded by a training fellowship from the UK Medical Research Council. This study was supported by a Northern and Yorkshire Research and Development Grant.

We wish to thank Dr I. M. Drake, Dr D. A. F. Lynch, Dr J. G. O'Mahony, Dr F. Hassan, Dr S. D. Humphris, Dr A. D. Newbound, Miss K. McConnell and Mrs. C. Chell for their assistance in validating the questionnaire. We are grateful to Andrea Scales for her help in the typing of this manuscript.

REFERENCES

1 Jones R, Lydeard S. Prevalence of symptoms of dyspepsia in the community. Br Med J 1989; 298: 30–2.
2 H.M.S.O. Statistics of prescriptions dispensed in the Family Health Service Authorities: England 1984–94. Statistical Bulletin 9515. London: 1995.
3 Talley NJ. *Helicobacter pylori* and non-ulcer dyspepsia. Scand J Gastroenterol Suppl 1996; 220: 19–22
4 Veldhuyzen van Zanten SJ. A systematic overview (meta-analysis) of outcome measures in *Helicobacter pylori* gastritis trials and functional dyspepsia. Scand J Gastroenterol Suppl 1993; 199: 40–3.
5 Jones R. Dyspeptic symptoms in the community. Gut 1989; 30: 893–8.
6 Talley NJ, Phillips SF, Melton LJ III, Wiltgen C, Zinsmeister AR. A patient questionnaire to identify bowel disease. Ann Intern Med 1989; 111: 671–4.
7 Veldhuyzen van Zanten SJO, Tytgat KMAJ, Pollack PT, *et al.* Can severity of symptoms be used as an outcome measure in trials of non-ulcer dyspepsia and *Helicobacter pylori* associated gastritis. J Clin Epidemiol 1993; 46: 273–9.
8 El-Omar EM, Banerjee S, Wirz A, McColl EL. The Glasgow Dyspepsia Severity Score—a tool for the global measurement of dyspepsia. Eur J Gastroenterol Hepatol 1996; 8: 967–71.
9 McDowell I, Newell C. Measuring Health. A Guide to Rating Scales and Questionnaires. Oxford: Oxford University Press, 1987.
10 Sjoden PO, Bates S, Nyren O. Continuous self-recording of epigastric pain with two rating scales: compliance, authenticity, reliability and sensitivity. J Behav Assess 1983; 5: 327–44.
11 Holdstock G, Harman M, Machin D, Patel C, Lloyd RS. Prospective testing and scoring system designed to improve case selection for upper gastrointestinal investigation. Gastroenterology 1986; 90: 1164–9.
12 Spiegelhalter DJ, Crean GP, Holden R, Knill-Jones RP. Taking a calculated risk: predictive scoring system in dyspepsia. Scand J Gastroenterol 1982; 79(Suppl.): 152–60.
13 Crean GP, Card WI, Beattie AD, *et al.* ''Ulcer-like dyspepsia''. Scand J Gastroenterol 1982; 79(Suppl.): 9–15.
14 Johannessen T, Peterson H, Kleveland PM, *et al.* The predictive value of history in dyspepsia. Scand J Gastroenterol 1990; 25: 689–97.
15 Talley NJ, McNeil D, Piper DW. Discriminant value of dyspeptic symptoms: a study of the clinical presentation of 221 patients with dyspepsia of unknown cause, peptic ulceration, and colelithiasis. Gut 1987; 28: 40–6.
16 Davenport PM, Morgan AG, Darnborough A, De Dombal FT. Can preliminary screening of dyspeptic patients allow more effective use of investigational techniques? Br Med J 1985; 290: 217–20.
17 Horrocks JC, De Dombal FT. Clinical presentation of patients with ''dyspepsia'': detailed symptomatic study of 360 patients. Gut 1978; 19: 19–26.
18 Edwards FC, Coghill NF. Clinical manifestations in patients with chronic atrophic gastritis, gastric ulcer and duodenal ulcer. Q J Med 1968; 37: 337–60.
19 Veldhuyzen van Zanten SJ, Tytgat KM, Jalali S, Goodacre RL, Hunt RH. Can gastritis symptoms be evaluated in clinical trials? An overview of treatment of gastritis, non-ulcer dyspepsia and *Campylobacter*-associated gastritis. J Clin Gastroenterol 1989; 11: 496–501.
20 Talley NJ, Zinsmeister AR, Schleck CD, Melton LJ III. Dyspepsia and dyspepsia subgroups: a population-based study. Gastroenterology 1992; 102: 1259–68.