

EDITORIAL

Measuring Gastroesophageal Reflux Symptoms: Musings from Marrakech

Paul Moayyedi and Brendan C. Delaney

Gastroenterology Division, McMaster University, Hamilton, Ontario; and Department of General Practice and Primary Care, University of Birmingham, Edgbaston, Birmingham, United Kingdom

The optimum approach for evaluation of reflux symptoms is uncertain. Trials have used a variety of methods, and it is important that a more consistent approach is taken. One piece of the puzzle is whether symptoms should be assessed by a clinician or directly by a patient completed questionnaire. McColl *et al.* suggest that clinicians tend to underestimate the severity of symptoms and that a more patient-centered approach should be used to elicit symptoms. This has important implications for the development of future questionnaires in gastroesophageal reflux and other gastrointestinal diseases.

(Am J Gastroenterol 2005;100:19–20)

Gastroesophageal reflux disease (GERD) is a common problem but research has been hampered by lack of clarity on how to evaluate symptoms. An approach that is often taken in this situation is to set up a working group, usually in a pleasant and exotic location. Professor John Dent duly organized an international group of experts, and we were honored to be included. The working party met for two days in September 2002 in Marrakech, Morocco, and reviewed the evidence that addressed symptom evaluation in GERD (1). The group made a number of recommendations both for research and clinical practice. These included the observation that symptoms occurring on two or more days per week were usually associated with reduced quality of life and that this frequency could be used as a guide when gastroesophageal reflux becomes a disease (1). Controversially, it was suggested that dysphagia alone should not be considered an alarm symptom unless its pattern and duration give cause for concern. A systematic review of GERD randomized trials reported that heartburn and regurgitation should be assessed over a defined time period (usually 1 wk) and that absence of symptoms or minimal symptoms (mild symptoms occurring less than twice a week) should be used as a definition of successful therapy (2). These recommendations may be useful but the working group would have failed if all it achieved was the ossification rather than the development of research. In an attempt to avoid this, the report also pointed out areas that needed further investigation. These included evaluating the diagnostic accuracy of symptoms using more sophisticated statistical techniques, such as latent class and Bayesian analysis (3), and devoting more effort to adequately assess patient satisfaction. It was generally agreed that diary cards were not widely useful in clinical practice but that they may be helpful in the research setting and this also needed further assessment (4). It is, therefore, gratifying to see that some of the other

participants of the workshop have followed up this suggestion and reevaluated previously collected data sets to determine whether physician and patient assessments of severity of reflux symptoms agreed with each other (5).

The authors conducted a *post hoc* analysis of four studies involving 2,674 GERD patients treated with proton pump inhibitors. In each study, GERD symptoms were assessed by a clinician using a four-point scale (none, mild, moderate, and severe). At the same visit, patients completed a validated gastrointestinal symptoms rating scale (GSRS) questionnaire. The GSRS assesses symptoms on a seven-point Likert scale so this was collapsed to a four-point scale (*e.g.*, none, mild/minor, and moderate, moderately severe/severe/very severe) to compare with physician's assessment. Agreement between patient and clinician as to the severity of GERD symptoms was modest prior to therapy. For example, the clinician only agreed with the patient on the severity of heartburn in 48–52% of cases (5). In three out of four studies, the clinician underestimated the severity of symptoms compared to the patient. Correlation was better after therapy especially if the patient reported no reflux symptoms where agreement occurred in 73–83% of cases (5).

This is an important piece of work because it provides evidence for the increasing move toward patient-oriented outcomes in clinical trials. For a long time, the primary outcome measure for many clinical trials has been the physicians' assessment of improvement. Patient-centeredness as well as best practice in psychometrics dictate that this assessment should be made by a properly validated symptom questionnaire (6). This study should be of interest to all designing and interpreting trials in GERD, in that it bridges the gap between the two methods of assessment, examining the degree of agreement between the two methods. The principal weakness in this type of analysis is that it was not planned at

inception, and is therefore a *post hoc* analysis of a convenience sample of trial data available to the authors. However, the trials are large and well conducted, and a table summarizing the main trial procedures relevant to this analysis (blinding of outcome assessments, types of assessment conducted by patients and doctors, timing and interventions) has been provided. It seems reasonable to assume that the findings are likely to be generalizable to other trials and symptom outcome measures.

For those interested in pooling results of clinical trials in meta analyses, the most important finding is that agreement on absence *versus* presence of symptoms is most in agreement with clinical impression. This is in concordance with other data presented at the Marrakech workshop that showed that total absence of symptoms correlated best with endoscopic healing of esophagitis (2) and formed the basis of the workshop's recommendation that total absence of symptoms should be the reference standard for clinical trial outcomes. This is consistent with findings in other gastrointestinal disorders such as nonulcer dyspepsia (7). The researcher is presented with a number of choices for outcome measures to assess patient's response to treatment in a trial, which have been reviewed recently (8). McColl has also reviewed previously the most desirable characteristics and appropriate validation of symptom outcome measures in general (6). A validated measure is likely to consist of a five- or seven-point scale with multiple questions that has undergone a full psychometric evaluation including examination of test-retest reliability, validity, and responsiveness in a representative population.

This paper is not going to alter clinical practice but has the potential to change how we conduct clinical research in GERD. Future trials should use self-completed questionnaires to assess GERD symptoms rather than relying on physician assessment. For the patient who is experiencing the symptoms, this seems almost self-evident. It does not appear obvious to researchers; however, as a systematic review of 108 GERD trials (2) found that only 53 included a patient completed symptom evaluation and in most of these it was unclear whether that or the physician assessment was used as the primary outcome. Post-Marrakech, we are at least able to agree how we should measure out-

comes in GERD clinical trials. The challenge is to further develop well-validated self-completed reflux symptom questionnaires (8–10).

REFERENCES

1. Dent J, Armstrong D, Delaney B, et al. Symptom evaluation in reflux disease: Workshop background, processes, terminology, recommendations, and discussion outputs. *Gut* 2004;53(Suppl 4):iv1–24.
2. Sharma N, Donnellan C, Preston C, et al. A systematic review of symptomatic outcomes used in oesophagitis drug therapy trials. *Gut* 2004;53(Suppl 4):iv58–65.
3. Moayyedi P, Duffy J, Delaney B. New approaches to enhance the accuracy of the diagnosis of reflux disease. *Gut* 2004;53(Suppl 4):iv55–7.
4. McColl E. Best practice in symptom assessment: A review. *Gut* 2004;53(Suppl 4):iv49–54.
5. McColl E, Junghard O, Wiklund I, et al. Assessing symptoms of gastroesophageal reflux disease: How well do clinicians' assessments agree with those of their patients? *Am J Gastroenterol* 2005;100:11–18.
6. McColl E, Jacoby A, Thomas L, et al. Design and use of questionnaires: A review of best practice applicable to surveys of health service staff and patients. *Health Technol Assess (Winchester, England)* 2001;5(31):1–256.
7. Moayyedi P, Delaney BC, Vakil N, et al. The efficacy of proton pump inhibitors in non-ulcer dyspepsia: A systematic review and economic analysis. *Gastroenterology* 2004;127:1329–37.
8. Fraser A, Delaney B, Moayyedi P. Symptom-based outcome measures for dyspepsia and GERD trials: A systematic review. *Am J Gastroenterol* 2004 in press.
9. Gralnek IM. Diagnosis and management of gastroesophageal reflux disease in the primary care setting: Can health-related quality of life play a role? *Am J Gastroenterol* 2001;96(8 Suppl):S54–6.
10. Shaw MJ, Talley NJ, Beebe TJ, et al. Initial validation of a diagnostic questionnaire for gastroesophageal reflux disease. *Am J Gastroenterol* 2001;96:52–7.

Reprint requests and correspondence: Paul Moayyedi, BSc, MB ChB, PhD, MPH, FRCP, Professor of Gastroenterology, Department of Medicine, Division of Gastroenterology, McMaster University Medical Center, 1200 Main St., West, HSC 3N51D, Hamilton, Ontario, Canada L8N 3Z5. e-mail: moayyep@mcmaster.ca
 Received October 31, 2004; accepted November 2, 2004.
