DISEASES OF THE ESOPHAGUS

Questionnaire

Validation of the Mayo Dysphagia Questionnaire

A. B. M. Grudell,[1] J. A. Alexander,[1] F. B. Enders,[2] R. Pacifico,[3] M. Fredericksen,[1] J. L. Wise,[1] G. R. Locke III,[1,4] A. Arora,[1] T. Zais,[1] N. J. Talley,[1,4] Y. Romero[1,4,5]

Divisions of [1]Gastroenterology and Hepatology and [2]Biostatistics, Mayo Clinic, Rochester, Minnesota, [3]Division of Gastroenterology, University of Michigan, Ann Arbor, Michigan, [4]GI Outcomes Research Unit and [5]Department of Otolaryngology, Mayo Clinic, Rochester, Minnesota, USA

SUMMARY. **While multiple instruments characterize upper gastrointestinal symptoms, a validated instrument devoted to the measurement of a spectrum of esophageal dysphagia attributes is not available. Therefore, we constructed and validated the Mayo Dysphagia Questionnaire (MDQ). The 27 items of the MDQ underwent content validity, feasibility, concurrent validity, reproducibility, internal consistency, and construct validity testing. To assess content validity, five esophageal subspecialty gastroenterologists reviewed the items to ensure inclusion of pertinent domains. Feasibility testing was done with eight outpatients who refined problematic items. To assess concurrent validity, 70 patient responses on the MDQ were compared to responses gathered in a structured patient-physician interview. A separate group of 70 outpatients completed the MDQ twice to assess the reproducibility of each item. A total of 148 patients participated in the validation process (78 [53%] men; mean age 62). On average, the MDQ took 6 minutes to complete. A single item (odynophagia) tested poorly with a kappa value of < 0.4. Otherwise, the majority of concurrent validity kappa values were in the good to excellent range with a mean of 0.63 (95% CI 0.22–0.89). The majority of reproducibility kappa values were also in the good to excellent range with a median kappa value of 0.76 (interquartile range: 0.67–0.81). Cronbach's alpha values were excellent in the range of 0.86–0.88. Spearman rank correlation coefficients to assess construct validity were also excellent in the range of 0.87–0.98. Thus, the MDQ is a concise instrument that demonstrates overall excellent concurrent validity, reproducibility, internal consistency, and construct validity for the features of esophageal dysphagia.**

KEY WORDS: **dysphagia, questionnaire, swallowing difficulty.**

INTRODUCTION

While multiple instruments characterize a variety of upper gastrointestinal symptoms,[1–11] a validated instrument devoted to the measurement of a spectrum of esophageal dysphagia attributes is not available. An instrument designed to diagnose and gauge dysphagia severity would improve uniformity in the patients entered into clinical research trials and provide a consistent assessment tool to measure efficacy end-points.

Dysphagia items have been included in the Gastroesophageal Reflux Questionnaire,[4] the Gastrointestinal Quality of Life Index,[3] the M.D. Anderson Dysphagia Inventory,[6] the Supraesophageal Reflux Questionnaire,[9] the SWAL-QOL,[10,11] and the modified Achalasia Symptom Score.[12] However, none of these instruments are devoted to assessing many different aspects of esophageal dysphagia in a succinct manner. While the Mayo Dysphagia Questionnaire (MDQ) shares the majority of its dysphagia items with other Mayo Clinic-developed instruments,[4,9] the 96-item Gastroesophageal Reflux Questionnaire[4] includes sections on irritable bowel syndrome, chest pain, and bowel habits, while the 29-item Supraesophageal Reflux Questionnaire[9] focuses on throat clearing, cough, globus sensation, sore throat and hoarseness.

We aimed to construct and validate a simple instrument to measure dysphagia using items from

Address correspondence to: Yvonne Romero, MD, Division of Gastroenterology and Hepatology, Gonda 9A, Mayo Clinic, Rochester, Rochester, MN 55905. Email: romero.yvonne@mayo.edu

previously validated questionnaires, encompassing a spectrum of domains including: dysphagia onset, frequency and severity; dysphagia for a variety of foodstuffs; previous episodes of food impaction and odynophagia.

METHODS

The MDQ is a 27-item instrument comprised of 25 items from previously validated Mayo Clinic instruments,[4,8,9,11] plus one item from the Marks et al.,[1] Cox et al. modified dysphagia scale,[2] and one new item. Sixteen items use a dichotomous format, nine items use a Likert scale, one item offers multiple non-hierarchical options and the final item is a scale. Three symptom domains were included in the body of the questionnaire, with three items per domain, to detail symptom duration, frequency and severity. The modified dysphagia scale[2] contains six options regarding the consistency of foodstuff that reliably causes dysphagia. One novel item inquires as to previous episodes of impaction. The MDQ contains items independent of the dysphagia domain. These items address heartburn, acid regurgitation, prior esophageal procedures (dilation, fundoplication and esophagectomy), seasonal allergies, asthma, and current use of acid suppression medications, as these factors can impact dysphagia and alter the overall interpretation of the dysphagia items. Due to its stem-and-leaf format, a patient can answer as few as 13 or as many as 27 items. The study was approved by the Mayo Clinic Foundation Institutional Review Board on October 14, 2004.

Content validity and feasibility methods

Five esophageal subspecialty gastroenterologists (AA, GRL, JAA, JLW, YR) reviewed the MDQ to ensure all pertinent domains were included in the questionnaire as an assessment of content validity. Although most items had been previously validated, the MDQ was piloted with eight outpatients to gather their input on the clarity of each item. Problematic items and the general format underwent revision until there was agreement that the questionnaire was intelligible and easy to read. The final version of the MDQ was subjected to concurrent validity and reproducibility testing.

Concurrent validity methods

Concurrent validity is an estimate of how well a physician's interview (the standard diagnostic test) corresponds to answers marked on a self-administered questionnaire. A total of 70 unselected outpatients were recruited from the Esophagus Clinic or the Upper Endoscopy Suite (Mayo Clinic, Rochester Minnesota, MN, US) in November 2004. All patients in these areas were invited to participate in an effort to limit selection bias. Subjects completed the MDQ and then underwent a structured interview by an investigator physician (ABMG) who was blind to their written responses. After the interview, the physician completed a MDQ based on the responses given during the interview. The investigator (ABMG) was a second-year internal medicine resident at the time of the interviews who was trained to conduct structured interviews. The generalizeability of the study is strengthened by having a nonesophageal expert interview the patient.

Reproducibility methods

In order to assess reliability, a separate group of 150 unselected outpatients were invited to complete an MDQ twice and were mailed study materials 2 weeks prior to their scheduled esophagogastroduodenoscopy or Esophagus Clinic appointment. The first 70 patients to return the questionnaire and consent form, and complete a second MDQ, were included in the analysis.

Statistical methods

The concurrent validity and reproducibility of each item were assessed using the simple kappa statistic (κ), which can range from -1 to $+1$ with κ-value of 0 indicating that the observed agreement is no different from that which would be found by chance alone. In general, κ-values greater than 0.8 are considered excellent, those greater than 0.6 are good, and those greater than 0.4 are fair. The internal consistency of the items in the three main symptom domains was assessed using the Cronbach's alpha (α) statistic. Skip patterns within the questionnaire were coded as separate responses to allow inclusion of all questions in the dysphagia domain when calculating the Cronbach's α. In general, Cronbach's α values greater than 0.7 are good and those between 0.8 and 0.90 are excellent. Cronbach's α values greater than 0.9 suggest that too many items may be included for a symptom domain. Construct validity comparing the dysphagia items with the food impaction item was assessed with the Spearman rank correlation coefficient. Spearman rank correlation coefficients range from 0 to 1.0 with those values closer to 1.0 being excellent. All statistical analysis was completed with the SAS System software (SAS Institute, Inc, Cary, NC, USA).

RESULTS

On average, the MDQ took 6 min (range 5–9 min) to complete. One hundred and forty-eight patients

Table 1 Baseline patient characteristics

Type of validity testing	Feasibility N = 8	Concurrent validity N = 70	Reproducibility N = 70
Age in years (SD)	67.3 (14.7)	60.6 (13.2)	57.7 (14.6)
Male : Female N (%)	5 (62) : 3 (38)	38 (54) : 32 (46)	35 (50) : 35 (50)
Dysphagia N (%)	n/a	36 (51)	36 (51)

SD = standard deviation, n/a = not applicable.

Table 2 Concurrent validity and reproducibility of the mayo dysphagia questionnaire

Mayo Dysphagia Questionnaire Item	Concurrent validity N = 70		Reproducibility N = 70	
	κ	95% CI	κ	95% CI
Dysphagia, presence (y/n)	0.97	0.92–1.00	0.86	0.74–0.98
Onset	0.96	0.93–1.00	0.84	0.73–0.96
Frequency	0.79	0.70–0.87	0.82	0.71–0.93
Severity	0.72	0.61–0.83	0.72	0.59–0.85
Heartburn, presence (y/n)	0.53	0.27–0.78	0.68	0.46–0.90
Onset	0.61	0.44–0.79	0.58	0.40–0.76
Frequency	0.53	0.37–0.70	0.67	0.52–0.81
Severity	0.42	0.26–0.58	0.56	0.40–0.71
Acid regurgitation, presence (y/n)	0.79	0.62–0.97	0.95	0.85–1.00
Onset	0.68	0.52–0.83	0.79	0.66–0.91
Frequency	0.63	0.49–0.78	0.76	0.65–0.87
Severity	0.49	0.34–0.64	0.68	0.55–0.81
Dysphagia in regard to particular foodstuffs or consistencies (y/n)				
Solids only	0.57	0.31–0.84	0.76	0.59–0.93
Solids and/or liquids	0.78	0.66–0.90	0.76	0.63–0.88
Liquids only	0.66	0.04–1.00	0.49	−0.11–1.00
Water	0.43	0.16–0.70	0.81	0.61–1.00
Grits	0.75	0.51–0.98	0.51	0.18–0.85
Banana	0.51	0.25–0.76	0.77	0.52–1.00
Apple	0.70	0.46–0.94	0.71	0.48–0.94
Bread	0.64	0.33–0.96	0.81	0.62–1.00
Meat	0.72	0.36–1.00	0.77	0.53–1.00
Food impaction (y/n)	0.60	0.35–0.85	0.78	0.60–0.96
Odynophagia (y/n)	0.30	0.05–0.55	0.26	−0.01–0.53

For each item in the questionnaire, the kappa value (κ) and corresponding 95% confidence intervals are reported.

participated in the validation process (Table 1). Overall, participants had a mean age of 62 years (SD = 14) with a male : female ratio of 1 : 1. Concurrent validity and reproducibility testing results are presented in Table 2. The Cronbach's α was 0.86 for patients and 0.88 for physicians. The Spearman correlation coefficients were 0.87, 0.89 and 0.98 when the food impaction item was compared to dysphagia severity, frequency, and onset, respectively.

Concurrent validity results

Seventy of 75 patients (93%) invited to participate in the concurrent validity process elected to do so. Of these 70, 36 (51%) had dysphagia as a concern for evaluation. The majority of concurrent validity κ-values for the MDQ were in the good to excellent range. The odynophagia item was the only item that did not test well with a κ-value of 0.30 (95% CI 0.05–0.55). The mean concurrent validity κ of

the remaining items was 0.63 (95% CI 0.22–0.89), with values ranging from 0.42 to 0.97. There was excellent concordance between the physician interview and the MDQ for the presence of dysphagia and its onset; with good concordance for the frequency and severity of dysphagia, previous episode of food impaction, and dysphagia for foodstuffs with differing consistencies. Since it is common clinical practice to group foodstuffs into three main categories of bolus type, we analyzed the concurrent validity of: (i) fibrous solids (meat, bread and apple); (ii) soft solids (oatmeal, grits and banana); and (iii) liquids/water, as collapsed categories with the following κ-values: 0.57 (95% CI 0.31–0.84); 0.78 (95% CI 0.66–0.90); and 0.66 (95% CI 0.04–1.00), respectively.

Reproducibility results

Of the 70 patients participating in the reproducibility phase of this study, 36 (51%) had dysphagia as a

concern for evaluation. The median (interquartile range) time between completion of the first and second questionnaires was 6 days (4–10). Excluding the odynophagia item, which demonstrated marginal performance, the reproducibility κ-values for the rest of the MDQ items ranged from 0.51 to 0.95. The MDQ was highly reproducible for the presence of dysphagia, dysphagia onset and frequency, and dysphagia with water and bread. Items concerning dysphagia severity, any previous episode of impaction, and specifics regarding the type of foodstuff associated with dysphagia, particularly banana, apple and meat demonstrated good reproducibility, while the grits item demonstrated fair reproducibility. The κ-values for the three main bolus types were 0.76 (95% CI 0.59, 0.93) for fibrous solids, 0.76 (95% CI 0.63–0.88) for soft solids and 0.49 (95% CI –0.11–1.00) for liquids/water.

DISCUSSION

The MDQ is a concise instrument that demonstrates excellent internal consistency, construct validity, concurrent validity and reproducibility for the features of esophageal dysphagia. Although a number of validated instruments have included at least one item concerning dysphagia,[3,4,6,9–11,13] a broad yet succinct instrument is lacking. No currently available instrument quantitates a number of key dysphagia elements (including when dysphagia was first noted, its frequency and severity, and foodstuffs that reliably pose a challenge [liquids, solids or both]). Unique to the MDQ is a food impaction item, which may be helpful for research in eosinophilic esophagitis and lichen planus. That said, the MDQ was validated to broadly assess dysphagia, not dysphagia due to a specific disease process. Based upon the validation data, we conclude that the MDQ, albeit brief, is the most comprehensive instrument to evaluate esophageal dysphagia currently available.

The MDQ has limitations. Since previously validated items were used in the questionnaire, we could not alter the wording of the items or create a global time window for the entire instrument. A second weakness is that the reproducibility estimate may be affected by recall bias since 10% of subjects completed the second questionnaire 2 days after the first. Because this time window has been used in other validation studies, this should be of minimal significance.[4,9,13]

Overall, the κ-values of the MDQ were high and well acceptable for both reproducibility and concurrent validity. Nonetheless, the range between the highest and lowest values was broad, likely due to the sample size of the study. Given that the perception of dysphagia is a very personal experience that is highly dependent upon the type of material being swallowed and the chewing behavior of the subject, and given that the κ ranges of many items were similar to, if not better than, those demonstrated by other validated instruments, we anticipate the MDQ will allow researchers and clinicians to create a uniform language for characterizing patients with esophageal dysphagia.[1–9,13]

In conclusion, the MDQ is a concise yet comprehensive instrument that demonstrates overall excellent concurrent validity, reproducibility, and internal consistency for the diagnosis and measurement of esophageal dysphagia. Our hope is that its use will improve uniformity in clinical research trials; functioning as a consistent assessment tool to reliably measure efficacy end-points.

References

1 Marks R D, Richter J E. Peptic strictures of the esophagus. Am J Gastroenterol 1993; 88: 1160–73.
2 Cox J G C, Winter R K, Maslin S C et al. Balloon or bougie for dilatation of benign oesophageal stricture? An interim report of a randomized controlled trial. Gut 1988; 29: 1741–7.
3 Eypasch E, Williams J I, Wood-Dauphinee S et al. Gastrointestinal quality of life index: development, validation and application of a new instrument. Br J Surg 1995; 82: 216–22.
4 Locke G R, Talley N J, Weaver A L, Zinsmeister A R. A new questionnaire for gastroesophageal reflux disease. Mayo Clin Proc 1994; 69: 539–47.
5 Svedlund J, Sjodin I, Dotevall G. GSRS-A clinical rating scale for gastrointestinal symptoms in patients with irritable bowel syndrome and peptic ulcer disease. Dig Dis Sci 1998; 33: 129–34.
6 Chen A Y, Frankowski R, Bishop-Leone J et al. The development and validation of a dysphagia-specific quality-of-life questionnaire for patients with head and neck cancer. Arch Otolaryngol Head Neck Surg 2001; 127: 870–6.
7 Wiklund I K, Junghard O, Grace E et al. Quality of life in reflux and dyspepsia patients. Psychometric documentation of a new disease-specific questionnaire (QOLRAD). Eur J Surg 1998; 583 (Suppl.): 41–9.
8 Dhillon S, Romero Y, Harmsen W S, Beri G, Schleck C, Zinsmeister A R. Concurrent validity and reproducibility characteristics of the reflux symptoms questionnaire and the reflux symptoms questionnaire 1-W. Am J Gastroenterol 2002; 97 (Suppl. 9): S243.
9 Dauer E, Thompson D, Zinsmeister A R et al. Supraesophageal reflux: validation of a symptom questionnaire. Otolaryngol Head Neck Cancer 2006; 134: 73–80.
10 McHorney C A, Bricker D E, Robbins J, Kramer A E, Rosenbek J C, Chignell K A. The SWAL-QOL outcomes tool for oropharyngeal dysphagia in adults. II. Item reduction and preliminary scaling. Dysphagia 2000; 15: 122–33.
11 McHorney C A, Robbins J, Lomax K et al. The SWAL-QOL and SWAL-CARE outcome tool for oropharyngeal dysphagia in adults. III. Documentation of reliability and validity. Dysphagia 2002; 17: 97–114.
12 Eckardt V F, Aignherr C, Bernhard G. Predictors of outcome in patients with achalasia treated by pneumatic dilation. Gastroenterology 1992; 103: 1732–8.
13 Romero Y, Thistle J L, Longsreth G F et al. A questionnaire for the assessment of biliary symptoms. Am J Gastro 2003; 98: 1042–51.