The Infant Gastroesophageal Reflux Questionnaire Revised: Development and Validation as an Evaluative Instrument

LEAH KLEINMAN,* MARGARET ROTHMAN,† RICHARD STRAUSS,† SUSAN R. ORENSTEIN,§ SUZANNE NELSON,¶ YVAN VANDENPLAS,‖ SALVATORE CUCCHIARA,# and DENNIS A. REVICKI**

*Center for Health Outcomes Research, United BioSource Corporation, Seattle, Washington; †Johnson & Johnson, Raritan, New Jersey; §Pediatric Gastroenterology, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania; ¶Division of Gastroenterology, Hepatology, and Nutrition, Children's Memorial Hospital, Chicago, Illinois; ‖AZ-VUB-Kinderziekhuis, Brussels, Belgium; #Department of Pediatric Gastroenterology, University of Rome, Rome, Italy; and **Center for Health Outcomes Research, United BioSource Corporation, Bethesda, Maryland

Background & Aims: **Gastroesophageal reflux disease (GERD) is frequently experienced by infants, and disease-specific measures are needed to evaluate treatment benefits. We revised the Infant Gastroesophageal Reflux Questionnaire (I-GERQ) on the basis of information from parents of infants with GERD and physicians and subjected it to a psychometric evaluation.** ***Methods:*** **A 3-week, multi-country observational study of 185 caregivers of infants younger than 18 months with GERD and 93 caregivers of control infants was conducted. Caregivers completed the I-GERQ-R weekly and recorded symptoms in a Daily Diary. Caregivers and physicians rated global disease severity and change in overall GERD symptoms.** ***Results:*** **Slightly more than half of infants were male with a mean age of 6.7 months, and most infants had been diagnosed with GERD for a little more than 2 months (mean, 66.7 days). Internal consistency reliability for the I-GERQ-R ranged from 0.86 to 0.87, and test-retest reliability was 0.85. Construct validity was demonstrated by significant differences between cases and controls on all item scores (all $P < .01$) and the total score ($P < .0001$), correlations with relevant Daily Diary symptoms, and both physician-rated ($P < .05$) and caregiver-rated disease severity ($P < .05$). Mean baseline to 3-week I-GERQ-R change scores for those infants whose caregivers reported improvement was −5.7 compared with −0.3 for those whose caregivers reported worse/same ($P < .001$). Physician ratings of change resulted in similar findings, with mean changes of −5.7 for those rated improved and −0.1 for those rated as worse/same ($P < .0001$).** ***Conclusion:*** **This study demonstrated the I-GERQ-R is a reliable, valid, and clinically responsive measure of infant GERD symptoms.**

Gastroesophageal reflux disease (GERD) in infants is among the most common causes for physician consultation worldwide. The lack of a reliable and valid noninvasive instrument to evaluate progression and remission of disease over time and to evaluate treatments for infant GERD in clinical trials, including international trials, has restricted clinical management and practice guidance. Symptoms in infants with GERD are considerably different from those in older children and adults,[1] mandating the development of infant-specific instruments for such purposes.

The pervasiveness of the main symptoms of GERD, regurgitation and crying, in normal infants makes challenging the noninvasive distinction between physiologic reflux and GERD. Recurrent regurgitation occurs in two thirds of 4-month-old infants.[2] Normal infants cry up to 2 hours daily.[3] Nonetheless, characterizing and quantifying the symptoms or symptom clusters can differentiate infants with GERD from normal infants.[4]

Both the North American Society for Pediatric Gastroenterology, Hepatology and Nutrition (NASPGHAN) and the European Society of Pediatric Gastroenterology, Hepatology and Nutrition (ESPGHAN) recommend using clinical symptoms for diagnosing and evaluating treatment of uncomplicated GERD in infants.[5,6] In the United States, a number of recent regulations now encourage clinical trials of pharmacotherapies in infants and children,[7–9] heightening the need for psychometrically sound evaluative measures. A psychometrically sound method of assessing infant GERD symptoms is required.

Symptom questionnaires have been developed to assess GERD in adults[10-15]; less attention has been paid to symptom assessment in infants and children. Before

Abbreviations used in this paper: ESPGHAN, European Society of Pediatric Gastroenterology, Hepatology and Nutrition; GERD, gastroesophageal reflux disease; ICC, intraclass correlation coefficient; I-GERQ, Infant Gastroesophageal Reflux Questionnaire; I-GERQ-R, Infant Gastroesophageal Reflux Questionnaire Revised; NASPGHAN, North American Society for Pediatric Gastroenterology, Hepatology and Nutrition; OTE, Overall Treatment Effect Scale; ROC, receiver operating curve; SD, standard deviation.

1542-3565/06/$32.00
doi:10.1016/j.cgh.2006.02.016

2005, only the Infant Gastroesophageal Reflux Questionnaire (I-GERQ)[4,16] had been developed and validated for the assessment of GERD-related symptoms in infants.

The I-GERQ was designed and validated as an instrument to obtain baseline clinical information and as a diagnostic questionnaire. It was validated for diagnosis of GERD in children ages 1–14 months by using abnormal pH probe studies and/or abnormal esophageal biopsies as gold standards. However, it was neither designed nor validated as an instrument to evaluate response to therapeutic interventions, and it is fairly long. Appendix A (see supplemental material online at www.cghjournal.org) indicates the items found to be the best discriminators within the 138-item I-GERQ. Application as an evaluative instrument necessitated several modifications: establishing a recall period, increasing and refining response options, eliminating redundant and ambiguous questions, and establishing and testing a scoring algorithm to assess symptomatic change over time. Use of the I-GERQ in countries other than the United States dictated translation into languages other than US English. This study's objective was therefore to revise the I-GERQ, according to best practices of questionnaire development, as an evaluative instrument and to translate the revised I-GERQ (I-GERQ-R) into several languages. Subsequently a prospective multinational observational study evaluated the new instrument's reliability, validity, and responsiveness. Cut scores for distinguishing between infants with and without GERD were evaluated for diagnostic sensitivity and specificity.

Methods

Content Validation

Validation of instruments for multi-country trials requires that the items have the same meaning across cultures.[14,15] The I-GERQ was thus revised according to established translation and validation guidelines, on the basis of input from focus groups with parents of infants with GERD held in France, Germany, Poland, and the United States. This input guided revision of individual items and response options. An international panel of physicians with expertise in infant GERD also reviewed and refined the questionnaire further.

Finally, the I-GERQ-R (revised) was translated into multiple European languages (Dutch, Finnish, French, German, Italian, Polish, Portuguese, Spanish, and UK English) by using standard translation methodology.[17-19] For cultural validation, a version was pilot-tested in the appropriate countries with a small sample of infant GERD caregivers who completed the questionnaire and then responded to questions about item comprehensibility, relevance, and acceptability. Final revisions were then made. For example, UK parents preferred "regurgitation" to "spit-up," so the UK version reflects this. Pilot testing was independent of the later validation study, described below.

Design and Procedures

The multinational observational study was conducted November 2002–February 2003 at 16 centers in 7 countries: Belgium (3 sites), France (3 sites), Italy (2 sites), The Netherlands (1 site), Poland (3 sites), UK (1 site), and the United States (3 sites). The majority of centers were pediatric gastroenterology clinics. Healthy infant controls were recruited from affiliated general pediatric clinics or were siblings of pediatric gastroenterology outpatients. This study was approved by all appropriate Ethical Review Committees or Institutional Review Boards.

The study population consisted of primary caregivers of infants with (cases) or without (healthy controls) a diagnosis of GERD. For inclusion, the infants were required to be <18 months of age at baseline and to have (cases) or not have (controls) a current diagnosis of GERD made by the usual practice at their clinical center. Healthy control infants were required to not have any concomitant illness that could produce symptoms like GERD and to not have any symptoms such as spitting up or fussiness to a degree that they could be considered signs of GERD. Caregivers were required to live in the same household as the infant, to be able to understand and read the appropriate language, and to provide written informed consent. Caregivers were ineligible if their infant had a condition that would interfere with the caregiver's ability to complete study questionnaires accurately, if the infant was participating in a clinical trial with investigational or approved medications for gastrointestinal complaints, or if the caregiver had a condition that, in the opinion of the investigator, would interfere with the requirements of the study.

To assess responsiveness of the new questionnaire, the study extended for 3 weeks between a baseline visit and a follow-up visit. During the baseline visit, the caregiver completed the I-GERQ-R and was instructed in completing a GERD Daily Symptom Diary. A physician assessed the infant and prescribed an active intervention including positional changes, feeding changes, and/or medication. The caregiver was asked to complete for 21 days a Daily Symptom Diary daily and a weekly I-GERQ-R (4 times total, including the baseline). At the follow-up visit, the same physician reassessed the infant. Half of the healthy controls' caregivers participated in the full 3-week study; the remainder participated in only the baseline assessment.

Study Measures

Infant Gastroesophageal Reflux Questionnaire Revised. The I-GERQ-R used in the study was a 14-item caregiver-completed symptom assessment scale; Appendix B (see supplemental material online at www.cghjournal.org) shows the final 12-item I-GERQ-R. All items were based on a 1-week recall period. Response choices ranged from 2–5 categories; higher scores indicated greater symptom burden.