Title

Development of an Online Library of Patient-Reported Outcome Measures in Gastroenterology: The GI-PRO Database

Journal

The American Journal of Gastroenterology, 109(2)

ISSN

0002-9270

Authors

Khanna, Puja
Agarwal, Nikhil
Khanna, Dinesh
et al.

Publication Date

2014-02-01

DOI

10.1038/ajg.2013.401

Peer reviewed

Published in final edited form as:
Am J Gastroenterol. 2014 February ; 109(2): 234–248. doi:10.1038/ajg.2013.401.

Development of an Online Library of Patient-Reported Outcome Measures in Gastroenterology: The GI-PRO Database

Puja Khanna, MD, MPH[1], **Nikhil Agarwal, MD**[2,3], **Dinesh Khanna, MD, MSc**[1], **Ron D. Hays, PhD**[4], **Lin Chang, MD**[5,6], **Roger Bolus, PhD**[5,7], **Gil Melmed, MD, MSHS**[3], **Cynthia B. Whitman, MPH**[7], **Robert M. Kaplan, PhD**[4], **Rikke Ogawa, MLIS**[8], **Bradley Snyder, MS**[7], and **Brennan M.R. Spiegel, MD, MSHS**[2,4,5,7]

[1]Division of Rheumatology, University of Michigan, Ann Arbor, Michigan, USA

[2]Department of Gastroenterology, VA Greater Los Angeles Healthcare System, Los Angeles, California, USA

[3]Department of Gastroenterology, Cedars-Sinai Medical Center, Los Angeles, California, USA

[4]Division of General Internal Medicine and Health Services Research, David Geffen School of Medicine at UCLA, and Department of Health Services, UCLA School of Public Health, Los Angeles, California, USA

[5]Division of Digestive Diseases, David Geffen School of Medicine at UCLA, Los Angeles, California, USA

[6]Center for Neurobiology of Stress, David Geffen School of Medicine at UCLA, Los Angeles, California, USA

[7]UCLA/VA Center for Outcomes Research and Education, Los Angeles, California, USA

[8]Biomedical Library of the Health Sciences, University of California at Los Angeles, Los Angeles, California, USA

Abstract

OBJECTIVES—Because gastrointestinal (GI) illnesses can cause physical, emotional, and social distress, patient-reported outcomes (PROs) are used to guide clinical decision making, conduct

Correspondence: Brennan M.R. Spiegel, MD, MSHS, Department of Gastroenterology, VA Greater Los Angeles Healthcare System, 11301 Wilshire Boulevard, Building 115, Room 215, Los Angeles, California 90073, USA. bspiegel@mednet.ucla.edu.

CONFLICT OF INTEREST

Guarantor of the article: Brennan M.R. Spiegal, MD, MSHS.

Specific author contributions: P. Khanna and N. Agarwal designed the study, reviewed the citations for systematic review, assigned quality scores for the selected manuscripts, and wrote the first draft. D. Khanna and B. Spiegel are Pis of the NIH PROMIS grant. They designed the study, reviewed the citations for systematic review, assigned quality scores for the selected manuscripts, and revised the first draft. R.D. Hays designed the study, and developed and assigned quality scores for the selected manuscripts. L. Chang, R. Bolus, and G. Melmed designed the study and assigned quality scores for the selected manuscripts; CB. Whitman and R.M. Kaplan assigned quality scores for the selected manuscripts; R. Ogawa conducted the systematic review; B. Snyder designed the PRO website. All authors revised the manuscript and approved the final version.

DISCLAIMER

The opinions and assertions contained herein are the sole views of the authors and are not to be construed as official or as reflecting the views of the Department of Veteran Affairs.

research, and seek drug approval. It is important to develop a mechanism for identifying, categorizing, and evaluating the over 100 GI PROs that exist. Here we describe a new, National Institutes of Health (NIH)-supported, online PRO clearinghouse—the GI-PRO database.

METHODS—Using a protocol developed by the NIH Patient-Reported Outcome Measurement Information System (PROMIS®), we performed a systematic review to identify English-language GI PROs. We abstracted PRO items and developed an online searchable item database. We categorized symptoms into content "bins" to evaluate a framework for GI symptom reporting. Finally, we assigned a score for the methodological quality of each PRO represented in the published literature (0–20 range; higher indicates better).

RESULTS—We reviewed 15,697 titles ($\kappa > 0.6$ for title and abstract selection), from which we identified 126 PROs. Review of the PROs revealed eight GI symptom "bins": (i) abdominal pain, (ii) bloat/gas, (iii) diarrhea, (iv) constipation, (v) bowel incontinence/soilage, (vi) heartburn/reflux, (vii) swallowing, and (viii) nausea/vomiting. In addition to these symptoms, the PROs covered four psychosocial domains: (i) behaviors, (ii) cognitions, (iii) emotions, and (iv) psychosocial impact. The quality scores were generally low (mean 8.88±4.19; 0 (min)−20 (max)). In addition, 51% did not include patient input in developing the PRO, and 41% provided no information on score interpretation.

CONCLUSIONS—GI PROs cover a wide range of biopsychosocial symptoms. Although plentiful, GI PROs are limited by low methodological quality. Our online PRO library (www.researchcore.org/gipro/) can help in selecting PROs for clinical and research purposes.

INTRODUCTION

Patients typically seek healthcare because they experience symptoms. Healthcare providers must elicit, measure, and interpret patient symptoms as part of the clinical evaluation. Patient-Reported Outcomes (PROs) capture the patients' illness experience in a structured format (1) and may help bridge the gap between patients and providers. Health-related quality-of-life (HRQOL) measures capture health directly reported by the patient (e.g., physical, emotional, or social symptoms) and can help to direct care and improve clinical outcomes. When clinicians systematically collect HRQOL data in the right place at the right time, it can effectively aid in detection and management of conditions (2,3), improve satisfaction with care (4), and enhance the patient-provider relationship (4–8).

In addition to their use in clinical practice, PROs also play an important role in clinical trials and other research endeavors. For example, HRQOL measures have gained traction as an outcome in clinical research, including clinical trials. HRQOL measures can document patient improvement or decrement over time, and help estimate the benefits of clinical interventions. In addition, the United States Food and Drug Administration (FDA) now considers the patient report as an important aspect of drug approval, and has developed guidance for use of PROs in clinical trials (1). The National Institute of Health has also supported a major PRO initiative, called the Patient-Reported Outcome Measurement Information System (PROMIS®; www.nihpromis.org), designed to develop and evaluate several ways of assessing HRQOL domains (9,10). The vision of the newly created Patient-Centered Outcomes Research Institute (PCORI: http://www.pcori.org/) is to provide patients

and the public with “the information they need to make decisions that reflect their desired health outcomes.” Finally, the rising prominence of the Chronic Care Model, which emphasizes the centrality of the provider-patient relationship in clinical decision making (2,3), places the patient report front and center. In short, there is a confluence of scientific, regulatory, and political factors that amplify the importance of PRO research.

Gastrointestinal (GI) illnesses are associated with physical, mental, and social distress (11). For this reason, patients, providers, investigators, and regulators are interested in using PROs to guide clinical decision making, conduct clinical research, and achieve drug approval in GI. Over the past two decades, investigators have developed PROs that measure a range of GI symptoms, including physical, emotional, and social features of digestive disorders. It is important for GI providers and researchers to be aware of existing PROs, and to have a mechanism for easily identifying and assessing the quality of these instruments.

The purpose of this article is to help guide clinicians and investigators by developing an online library of English-language GI PROs published to date. Using a systematic protocol developed by the National Institutes of Health (NIH) PROMIS network (12,13), we searched the literature to identify GI PROs. We categorized the instruments in a searchable PRO clearinghouse, called the GI-PRO database. This paper describes how we developed this resource, and explains how interested stakeholders can use this database to assist with selecting PROs for both clinical and research purposes.

METHODS

Study overview

We performed a structured search to identify English-language PROs across all luminal diseases and other illnesses that directly affect the GI tract (e.g., systemic sclerosis). We developed a search strategy that targets studies describing the development and evaluation of English-language PROs that measure GI symptoms, including physical, emotional, cognitive, and social symptoms attributable to disorders affecting the GI tract. We then abstracted individual items from each PRO, and developed a comprehensive item library that is searchable using an online relational database. We developed “bins” within which to categorize physical symptoms, and used this to assess a framework for GI physical symptom reporting, similar to one developed previously for irritable bowel syndrome (IBS) (14), and a process supported by the NIH PROMIS network (13). Finally, we categorized each PRO, and assigned a score for the methodological quality of each instrument. In the sections that follow, we describe our search strategy, abstraction methods, quality scoring, and development of the GI-PRO database.

Literature search

We performed a systematic review of PubMed, the Cochrane database, and the Patient-Reported Outcomes and Quality of Life Instruments Database (PROQOLID) to identify English-language publications from 1946 through January 2012. We developed our search strategy using PROMIS criteria in concert with an expert librarian (R.O.), and mirrored the approach employed by the University of Pittsburgh PROMIS investigators (12). We next

developed a set of keywords related to GI disorders using a combination of Medical Subject Headings (MeSH) and text words (15). After developing the search terms, we applied both the Pittsburgh strategy (12) and a previously validated PRO filter developed by Terwee *et al* (16), and compared retrieval sets. We added limits to the search strategy to find English-language articles discussing human subjects.

Before finalizing our search strategy, we conducted pilot tests to troubleshoot any shortcomings in the filter. In test runs, we discovered two complications with the initial search strategy: the Pittsburgh filter did not include several known test articles relevant to our target population, and the Terwee filter included a large number of citations whose topics were beyond the scope of the project. In order to fine-tune the filters for our purposes, we developed a test set of 42 known PROs indexed in PubMed that we determined must be included in the final search. We employed this test set to update the search strategy that we itera-tively modified to optimize sensitivity. To combine, modify, and create our final search strategy, we reviewed the search concepts in a group session with the librarian to identify salient text words and MeSH terms from both the Pittsburgh and Terwee filters. To guide triage decisions about retaining search terms, we focused on concept relevance and potential for error (e.g., multiple meanings of a term causing irrelevant retrieval). We identified variants of terms, including British spellings and alternate plurals for phrase searches, and included these permutations as part of the revised filter. As we reviewed terms for inclusion or exclusion, we iteratively verified the filter against the set of known articles to identify potential for error. Of the 42 known articles, we retrieved all but one using the final search strategy. The missing article was not in English and was therefore ineligible, as we focused on English-language articles. Table 1 provides the final search strategy.

Study selection

Four authors (P.K., N.A., B.M.R.S., and D.K.) reviewed the citations generated from the search strategy. We divided the review into three stages: titles, abstracts, and manuscripts. Pairs of reviewers assessed each generated title for relevancy, and rejected titles that fulfilled one of the following explicit exclusion criteria: (i) not written in English, (ii) did not concern a clinical question relevant to human subjects, (iii) did not pertain to a luminal disorder, or illness with GI manifestations; (iv) did not pertain to PROs or diseases severity/activity indices; (v) referred only to objective disease markers, including radiographic, histologic, biochemical, endoscopic, or stool markers; (vi) referred to a translation of existing instruments into other languages; and/or (vii) was a self-described editorial, review article, letter, case reports, or opinion piece.

We included PROs that exclusively included items measuring physical, emotional, or social symptoms as reported by the patient. Instruments that measure disease activity via biomarkers (i.e., lab studies, endoscopic findings, clinical findings, and so on) were excluded. For example, the Crohn's Disease Activity Index is a severity scale that includes biomarker data and physical examination findings, and was therefore excluded. This does not mean the Crohn's Disease Activity Index is unimportant or that it includes some PRO information—but simply indicates that it is not exclusively patient-reported information.

We next sought to identify and resolve discordant assessments between reviewers. This occurred through discussion between the two raters and, if there was uncertainty, included oversight by a third-party arbiter who reviewed and discussed the discordant titles. We remained conservative by accepting a title if there was uncertainty about how best to deliberate.

The reviewers then assessed the relevancy of all abstracts corresponding with the remaining titles, and excluded abstracts for the following reasons: (i) fulfilled one or more of the title exclusion criteria, and (ii) did not provide original data (or was solely a review article). For the last cycle, the reviewers evaluated the relevancy of all manuscripts corresponding with the remaining abstracts, and included manuscripts if they featured PRO items pertaining to GI symptoms and their dimensions. We then performed manual reviews of the reference lists from key review articles to identify additional studies missed by the computer-assisted searches.

Interrater reliability was measured using the κ-statistic. The κ refers to interrater agreement between two independent raters evaluating titles and abstracts from the literature search, assigning either “accept” or “decline” to each title or abstract. We calculated a standard unweighted κ based on the resulting 2×2 table, using the standard formula = (observed agreement - expected agreement)/ (1 –expected agreement) (17).

Item “binning”

We shared the final list of instruments with three GI experts with experience in PRO development (William Chey (University of Michigan), Douglas Drossman (University of North Carolina), and Jan Irvine (University of Toronto)). We asked the experts to comment on the comprehensiveness of the search and identify potentially missing instruments. We then developed a framework of “bins” and “subbins” to categorize GI symptoms contained within PRO measures. “Binning” is a structured process, described and endorsed by the NIH PROMIS network, to position individual PRO items within posited domains (13). Using a top-down approach, we first defined a set of predefined bins as recommended by the PROMIS domain protocol documentation (13). We remained flexible to add new bins for items that did not easily fit into a predefined bin. We established key words for each bin, and assigned each PRO item a bin label. This resulted in a conceptual framework of bins for categorizing GI physical symptoms, along with a list of individual items within symptom bin. We then summarized this information and made it available in our online GI-PROS database, as described below.

Quality scoring of PROs

Following collection of existing PROs, development of bins, and categorization of PRO items within bins, we assigned a quality score to each PRO instrument. We employed a checklist of PRO methodological “best practices” based on previously developed quality scores, including the Evaluating the Measurement of Patient-Reported Outcomes (EMPRO) (18) and the COnsensus-based Standards for the selection of health status Measurement INstruments (COSMIN).

EMPRO has 39 items that assess 8 attributes: conceptual and measurement model, reliability, validity, responsiveness, interpretability burden, alternate modes of administration, and cross-cultural and linguistic adaptations of measures (18). COSMIN includes over 100 items that assess 12 PRO attributes: internal consistency, reliability, measurement error, content validity, structural validity, hypothesis testing, cross-cultural validity, criterion validity, responsiveness, interpretability, generalizability of results, and item response theory (19–22). Both checklists are limited by their very high response burden, and the COSMIN is further limited by a low interrater reliability (19).

In light of these shortcomings, we modified the instruments to develop a brief checklist of methodological best practices that is tenable for use in evaluating a large number of instruments. The resulting checklist consists of eight items used to evaluate key properties of PRO instruments: (i) item writing process, (ii) content validity, (iii) reliability of measures, (iv) reliability for multiple English-language subgroups, (v) construct validity, (vi) interpretability, (vii) response burden, and (viii) mode of administration. Six of the items have a polytomous rating scale and two have a dichotomous response scale (Table 2). Detailed definitions of the checklist items are available on the website.

Development of online database

We developed a publicly available online PRO library, called the GI-PRO database, for users to identify and categorize all available English-language GI PROs identified by our search strategy. We programmed the relational database with the following forward and backward search functions:

- *PRO-level search:* users can enter the name of a known PRO to access information about the instrument, including the name of its developers, year of publication, supporting manuscript information, domain coverage, number of items, target conditions and populations, and quality checklist scores (forward search).
- *Symptom-level search:* users can enter a symptom of interest (e.g., pain, bloating, diarrhea, incontinence, and so on) to access a list of all questionnaires with items that pertain to the target symptom (backward search). This allows the user to view a list of every permutation of the symptom published to date, as identified by our search.

A list of searchable key terms is available in Appendix 1

RESULTS

Search results

The search strategy identified 15,697 titles, of which 183 met our final inclusion criteria ($\kappa >$ 0.6 for title and abstract selection). There were 121 PRO instruments, comprising 2,372 items, described in the included studies. The expert panel identified 5 additional instruments that were not captured in our systematic review, for a total of 126 instruments. The majority of PROs had one publication and we found that there was a mean of 1.25 publications per PRO (s.d. = 0.72; range 1.00–5.00; median 1.00). A higher number of publications per PRO were associated with higher quality as subsequent publications assessed quantitative aspects

such as longitudinal assessments and minimally important differences. The PROs cover a range of conditions, including achalasia (23), celiac sprue (24,25), dyspepsia (26–43), eosinophilic esophagitis (44), fecal incontinence (45–60), functional GI disorders (25,41,49,53,60–86), gastroesophageal reflux disease (30,38–40,43,87–110), GI malignancies (48,77,111–115), postgastrectomy (113,116), ileal conduit diversion (117), ileostomy (118), inflammatory bowel disease (119–127), pregnancy-related GI symptoms (128–130), systemic sclerosis (131–133), and radiation enteritis (57), among others. In all, 15 PROs apply to the pediatric population (44,45,60,77,80,97,100,102,115,123,124,134–137), and 6 apply specifically to women (56,128–130,138,139).

Item binning

The conceptual framework in Figure 1 represents the major GI symptom bins identified by our literature search. The framework posits that GI symptoms are represented by eight content bins: (i) abdominal pain, (ii) bloat/gas, (iii) diarrhea, (iv) constipation, (v) bowel incontinence/soilage, (vi) heartburn/reflux, (vii) swallowing, and (viii) nausea/vomiting. We discuss each bin further in the following paragraphs.

Abdominal pain—Of the included PROs, 12 included symptoms of abdominal pain (41,44,49,61,63,66,114,131,134,140–142). The PROs reveal that abdominal pain is multifaceted; PRO items cover many dimensions of abdominal pain, including intensity, frequency, bothersomeness, location, and pain interference.

Gas/bloat—Of the included PROs, 15 included symptoms related to gas, flatus, and bloating (30,49,56,62,63,66,71,76, 82,84,88,99,114,140,143). The PRO items can be divided into patient descriptions of the *look* vs. *feel* of bloating, as we have also discovered in previous qualitative work (14). Many PRO items refer to “flatulence” as a related but separate symptom that indicates passing gas (in contrast to perceived gas retention with subsequent visible bloating). In addition, many PRO items refer to “gurgling,” or “rumbling” of contents inside the belly. We currently group these latter symptoms within the gas/bloat/ flatulence domain.

Defecatory symptoms—The literature search identified 2 major defecatory symptom bins: diarrhea and constipation. There were 24 PROs with items pertaining to diarrhea (44,49,51,56,57,59, 61,63,66,70,71,113,114,116,117,119,120,131,134,140,144–147); the items included bowel urgency (i.e., having to rush to the bathroom), increased stool frequency, and loose stool form. Seventeen PROs included items pertaining to constipation (53,58,60,61,66,69,76–86), including incomplete evacuation, straining, infrequent stools, and hard or lumpy stools.

Bowel incontinence/soilage—Previous research, focused in IBS, revealed that incontinence is a separate bin from urgency (14). Our search yielded 16 PROs with items pertaining to bowel incontinence (45–60). This bin includes a spectrum of symptoms ranging from leakage, underpant soilage, and outright incontinence. It is noteworthy that these PRO items were often distinct from urgency or diarrhea, again suggesting a stand-alone bin for incontinence/soilage.

Nausea/vomiting—Our search revealed 11 PROs with items pertaining to nausea and vomiting (44,66,110,115,128–130,136,137, 140,148,149). These items capture a range of increasingly severe foregut symptoms that begins with feeling sick to the stomach and ends with vomiting up stomach contents. Intermediate symptoms include low appetite, feeling sick to the stomach, dry heaving, and queasiness in the belly.

Gastroesophageal reflux (GER)—We identified 29 PROs with items pertaining to GER (30,38–40,43,87–109). Patients with GER experience a wide range of foregut symptoms. Our review of extant items, coupled with evaluation of focus group results, identified four subdomains of GER symptoms, including: (i) liquid and food sensations (reflux, regurgitation, choking, bad-tasting liquids), (ii) painful sensations (heartburn, chest pain, throat burn), (iii) belching gas/hiccups, and (iv) head and neck sensations.

Disrupted swallowing—We identified 13 PROs with items pertain -ingto disrupted swallowing, including dysphagia and odynophagia symptoms (23,30,44,58,88,90,106,108,112–114,143,144). Patients with disrupted swallowing describe a range of symptoms ranging from transient food “sticking” to complete inability to swallow solids or liquids. These symptoms typically progress along a clinical spectrum, as captured in the extant items in the literature.

Psychological symptoms, social symptoms, and health behaviors

In addition to physical symptoms, published GI PROs cover a range of psychosocial symptoms attributable to GI disorders. Table 3 provides a list of psychosocial domains within the PROs identified by our search. The table highlights the breadth and depth of psychosocial illness experiences of GI patients. In addition, many PROs address health behaviors attributable to GI diseases. These can be classified into *avoidance and restrictive behaviors* (e.g., avoiding social events, travel, or culprit foods), *proactive and preventative behaviors* (e.g., wearing loose fitting clothes, staying near bathrooms, exercising to avoid GI symptoms), and *reactive behaviors* (e.g., taking medicines, running to the bathroom, separating from others during symptoms, applying heating pads).

Quality scoring

Table 4 provides the results of quality scoring. The mean score was 8.88 (s.d. = 4.19) out of 20 (higher indicates better). Only 59% of instruments provided information on how to interpret scores, and 51% did not include any patient input in developing the PRO instrument. Only 46% of GI PROs were based on cognitive debriefings of the items. The results of the full checklist are presented in Table 4. Table 5 lists the highest scoring instruments by disease category.

Online PRO library

The online library is accessible at http://www.researchcore.org/gipro/ and on the homepage of www.ResearchCORE.org. Figure 2 provides a screenshot of the main search page of the GI-PRO database and demonstrates the results of search. Users can employ the database to conduct forward and backward searches of PRO instruments, domains, and individual items.

DISCUSSION

Healthcare providers measure and interpret the patient report in order to direct clinical decision making. The literature suggests that HRQOL measurement can aid in detection and management of conditions (2,3), improve patient satisfaction (4), and enhance the patient-provider relationship at the center of chronic disease care (4–8). For this reason, it is important to have a systematic accounting of the available PROs. This is especially true in GI, where most illnesses announce themselves through patient-reported symptoms rather than abnormal biomarkers. We have developed an online clearinghouse of PROs that pertain to luminal digestive diseases—the GI-PRO database—that is now available to help clinicians, researchers, clinical trialists, and educators identify available GI PROs, evaluate their characteristics, and assess their level of methodological quality.

The database provides a novel way to identify GI PROs. The PRO repository may be useful for clinical studies and clinical practice. For example, researchers interested in finding an HRQOL measure for a study can readily access disease-targeted measures from the website. In addition, clinicians seeking PROs or individual questions to evaluate a patient's symptoms or biopsychosocial illness context can obtain guidance from PROs listed in the registry. For example, clinicians may seek guidance for how best to measure the HRQOL impact of IBS, and upon reviewing the PRO offerings may enhance their approach to illness assessment in a patient with chronic IBS symptoms. This repository may also serve as a resource to learn more about the breadth and depth of symptoms for common GI illnesses, as the online library encompasses the full spectrum of common GI diseases.

Our study has five key findings. First, using an NIH-endorsed strategy of systematic review, we identified over 100 PROs (to date) that pertain to luminal GI illnesses. In light of this extensive collection of PROs, interested stakeholders may benefit from our GI-PROs database to help them navigate the PRO terrain in GI.

Second, based on a structured review of the available PROs, we developed a framework for categorizing GI physical symptoms (Figure 1). According to this framework, almost every GI symptom can be "binned" into one of the eight categories. This parsimonious model indicates that the alimentary tract is surprisingly efficient in its symptom expression; the symptom dictionary is relatively narrow. Whereas the variety of underlying GI disorders is expansive, their symptom expressions funnel into a narrow taxonomy of defined presentations.

Our research group used this framework to develop a GI item bank for the PROMIS consortium—an NIH Roadmap Initiative with the goal of building, evaluating, and disseminating a toolbox of publicly available PRO item banks across the human illness experience. The PROMIS item banks are different from traditional paper-and-pencil questionnaires, because they can be electronically administered using computerized adaptive testing based on item parameters estimated using item response theory (150,151). This yields highly efficient and short questionnaires that can be implemented in busy clinical systems while preserving reliability and validity. Each PROMIS item bank is based on an underlying framework derived from several sources, including literature searches. With this

background, we used the framework in Figure 1 to serve as the backbone for our GI item bankfor PROMIS (www.nihpromis.org). Based on this symptom framework, the PROMIS GI bank provides a multidimensional tool for measuring symptoms across the full breadth and depth of GI conditions.

Third, we found that the existing GI PROs measure a wide range of psychosocial symptoms. This emphasizes that GI illnesses not only generate physical symptoms, but also impact emotional, cognitive, and social functioning. Table 3 reveals the full range of psychosocial subscales contained within the existing GI PROs. It is now possible, using the GI-PROs database, to efficiently search these terms to identify PROs that measure each concept. In addition, the GI-PROs database allows users to view the terminology employed by PROs to capture psychosocial symptoms in GI.

Fourth, we found that many PROs measure health behaviors related to GI illnesses. These behaviors can be divided into *avoidance and restrictive behaviors* (e.g., avoiding social events), *proactive and preventative behaviors* (e.g., wearing loose fitting clothes), and *reactive behaviors* (e.g., taking medicines). This points out the continuum of patient reporting: physical symptoms may have an affective consequence (e.g., bother) that leads to health behaviors that ultimately affect overall HRQOL. The existing GI PROs sample from across this spectrum.

Finally, we found that the overall methodological quality of GI PROs is typically not high (with notable exceptions). For example, only 46% of PROs relied on patients to develop the content of the questionnaire. This is problematic, because patients are the gold standard for content development, and are, in fact the “P” in PRO. As mandated by the FDA (152) and the NIH (13) protocols for PRO development, future GI PROs must rely on patients, first and foremost, for item development. In addition, patients should carefully review the resulting items before accepting them as final. The usual approach is to conduct cognitive debriefing interviews that allow patients to evaluate items and their response scales. Cognitive interviewing is a powerful tool for gaining a better understanding of the underlying or covert process involved in responding to survey items through the use of verbal probing techniques (153,154). Only 52% of GI PROs in this review involved patients in cognitive debriefing of candidate items. Another shortcoming is that virtually none of the PRO studies provided data across important patient subgroups. For example, the performance of PROs across gender, age, or race remains largely unknown. The PROs are further limited by rarely providing data on how to interpret the scores generated by the instruments. Without knowing the minimally clinically important difference on a scale, it is difficult to establish the meaning of PRO improvements or decrements over time. Finally, only 29% of studies provided data on the response burden of the PRO to patients. As PROs move into everyday care (152,155), it will be increasingly important to develop instruments that are efficient and easily administered in busy clinical environments.

Our database has several strengths. First, it is in the public domain—funded by the NIH and available to all interested stakeholders, including clinicians, researchers, and patients themselves. We relied on a rigorous systematic review using an NIH-endorsed search strategy and supplemented our review by reaching out to experts in the field. In addition to

providing descriptive information about existing PROs, we provide quantitative information about the methodological quality of these instruments using a checklist. We hope that after reviewing the checklist, researchers will have a better idea of what characteristics are needed to develop a high-quality PRO. Although this exercise was not specifically intended to "grade" PROs, but rather document their fundamental characteristics, researchers will be able to see which published instruments have the elements on our checklist. This will help guide development of PROs from researchers as it will provide a list of instruments with sound methodology. Finally, we have programmed the database to be highly functional, providing users with a unique interface to efficiently search the existing GI PROs. In time, we intend to expand the GI-PROs database as new instruments are published—an ongoing process. In addition, we hope this model will serve as an impetus for other subspecialists to evaluate their own PRO literature and consolidate the findings using a similar database structure.

Our database also has important limitations. In particular, quality scoring is an imprecise science; subjectivity can easily undermine a seemingly rigorous scoring system. This is well known to the developers of previous PRO quality assessment scores (19), and is not lost on us. We anticipate that many of our quality scores are debatable and are dynamic (as publications continue to improve quality scoring), and we remain modest in our assessments to date. If PRO authors believe our assessments are incorrect, then we will remain open and flexible to consider updates to our existing scores in an effort to improve this shared, publicly available resource (to express any concerns about the accuracy of the database, please go to http://www.researchcore.org/gipro). Indeed, one benefit of an online, dynamic database is the ability to update its information at any time. Furthermore, as more data are collected for a PRO, quality scores can be updated to keep pace with incremental information. Another limitation is that we limited our database to English-language PROs. Although most PROs are in English, there are undoubtedly many outstanding instruments in other languages. With more time and resources, the GI-PROS database could be expanded to include the non-English literature. Our original abstractions did not collect data on translations or trans-latability as well; future work will aim to include these important attributes of the included PROs. In addition, despite our extensive search, it remains likely that we have missed some PROs, and hence we acknowledge that this library is a work in progress and that some PRO instruments, especially those related to luminal GI disorders other than IBS, may have been missed. We welcome notification of an oversight, and remain open to update the database to incorporate all eligible PROs. Finally, we developed a streamlined quality checklist based on previously developed quality scores such as EMPRO (18) and COSMIN checklists (19–22). However, these checklists have over 140 items that assess every possible attributes of PRO measures; as such, they are highly cumbersome and time consuming. To populate our database, we sought to simplify the process to avoid the high response burden and low interrater reliability of existing scoring methods. As a result, we did not capture all aspects of methodological quality, but instead focused on particular areas of importance.

In summary, we have developed a publicly available online library of PRO measures that itemizes physical, emotional, and social symptoms pertinent for patients with GI distress.

The GI-PROS database is searchable at multiple levels, including PROs, domains, individual items, and methodological quality.

Acknowledgments

Financial support: NIH/NIAMS U01 AR057936A, the National Institutes of Health through the NIH Roadmap for Medical Research Grant (AR052177). Puja Khanna was supported by ACR Research and Education Foundation Career Development Bridge Funding Award 2011–12. Dinesh Khanna was also supported by NIAMS K24AR063120-02. Ron Hays was supported by NIH/NIA grants P30-AG028748 and P30-AG021684, and NCMHD grant 2P20MD000182. Brennan Spiegel has received grant support from Ironwood, Movetis, and Shire Pharmaceuticals. Lin Chang has received grant support from Prometheus, Rose Pharma, and Takeda.

Potential competing interests: Brennan Spiegel has served as an advisor for AstraZeneca, Movetis, Prometheus, Ironwood, Salix, and Shire; Gil Melmed has served as a consultant for Centocor, Amgen, Setpoint, Celgene, as advisor for UCB, and is on the speaker's bureau for Prometheus; Lin Chang has served as a consultant to Albireo, Movetis, Ironwood, Prometheus, Rose Pharma, Salix, and Takeda North America; Ron D. Hays has served as a consultant for Amgen, Allergan, Pfizer, and the Critical Path Institute.

REFERENCES

1. von Gruenigen VE, Huang HQ, Gil KM, et al. Assessment of factors that contribute to decreased quality of life in Gynecologic Oncology Group ovarian cancer trials. Cancer. 2009; 115:4857–4864. [PubMed: 19691091]
2. Wagner EH, Austin BT, Von KorffM. Organizing care for patients with chronic illness. Milbank Q. 1996; 74:511–544. [PubMed: 8941260]
3. Wagner EH, Austin BT, Davis C, et al. Improving chronic illness care: translating evidence into action. Health Aff (Millwood). 2001; 20:64–78. [PubMed: 11816692]
4. Marshall S, Haywood K, Fitzpatrick R. Impact of patient-reported outcome measures on routine practice: a structured review. J Eval Clin Pract. 2006; 12:559–568. [PubMed: 16987118]
5. Dobscha SK, Gerrity MS, Ward ME. Effectiveness of an intervention to improve primary care provider recognition of depression. Eff Clin Pract. 2001; 4:163–171. [PubMed: 11525103]
6. Taenzer P, Bultz BD, Carlson LE, et al. Impact of computerized quality of life screening on physician behaviour and patient satisfaction in lung cancer outpatients. Psychooncology. 2000; 9:203–213. [PubMed: 10871716]
7. Velikova G, Booth L, Smith AB, et al. Measuring quality of life in routine oncology practice improves communication and patient well-being: a randomized controlled trial. J Clin Oncol. 2004; 22:714–724. [PubMed: 14966096]
8. Detmar SB, Muller MJ, Schornagel JH, et al. Health-related quality-of-life assessments and patient-physician communication: a randomized controlled trial. JAMA. 2002; 288:3027–3034. [PubMed: 12479768]
9. Reeve BB, Hays RD, Bjorner JB, et al. Psychometric evaluation and calibration of health-related quality of life item banks: plans for the Patient-Reported Outcomes Measurement Information System (PROMIS). Med Care. 2007; 45:S22–S31. [PubMed: 17443115]
10. Cella D, Yount S, Rothrock N, et al. The Patient-Reported Outcomes Measurement Information System (PROMIS): progress of an NIH Roadmap cooperative group during its first two years. Med Care. 2007; 45:S3–S11. [PubMed: 17443116]
11. Spiegel B, Khanna D, Bolus R, et al. Understanding gastrointestinal distress: a framework for clinical practice. Am J Gastroenterol. 2011; 106:380–385. [PubMed: 21378758]
12. Klem, M.; Saghafi, E.; Abromitis, R., et al. Creation of exhaustive item banks: the role of the health sciences librarian. PROMIS Conference: Building Tomorrow's Patient-Reported Outcome Measure; Gaithersburg, MD: 2006.
13. DeWalt DA, Rothrock N, Yount S, et al. Evaluation of item candidates: the PROMIS qualitative item review. Med Care. 2007; 45:S12–S21. [PubMed: 17443114]

14. Spiegel BM, Bolus R, Agarwal N, et al. Measuring symptoms in the irritable bowel syndrome: development of a framework for clinical trials. Aliment Pharmacol Ther. 2010; 32:1275–1291. [PubMed: 20955447]
15. Cochrane Handbook for Systematic Reviews of Interventions, Version 5.0.2. [Accessed 30 March 2010] 2009. at http://www.cochrane-handbook.org/
16. Terwee CB, Jansma EP, Riphagen II, et al. Development of a methodological Pub Med search filter for finding studies on measurement properties of measurement instruments. Qual Life Res. 2009; 18:1115–1123. [PubMed: 19711195]
17. Agresti, A. Categorical Data Analysis. 2nd edn. Hoboken, NJ: John Wiley and Sons; 2002.
18. Valderas JM, Ferrer M, Mendivil J, et al. Development of EMPRO: a tool for the standardized assessment of patient-reported outcome measures. Value Health. 2008; 11:700–708. [PubMed: 18194398]
19. Mokkink LB, Terwee CB, Gibbons E, et al. Inter-rater agreement and reliability of the COSMIN (COnsensus-based Standards for the selection of health status Measurement Instruments) checklist. BMC Med Res Methodol. 2010; 10:82. [PubMed: 20860789]
20. Mokkink LB, Terwee CB, Knol DL, et al. The COSMIN checklist for evaluating the methodological quality of studies on measurement properties: a clarification of its content. BMC Med Res Methodol. 2010; 10:22. [PubMed: 20298572]
21. Mokkink LB, Terwee CB, Patrick DL, et al. The COSMIN study reached international consensus on taxonomy terminology and definitions of measurement properties for health-related patient-reported outcomes. J Clin Epidemiol. 2010; 63:737–745. [PubMed: 20494804]
22. Mokkink LB, Terwee CB, Patrick DL, et al. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. Qual Life Res. 2010; 19:539–549. [PubMed: 20169472]
23. Urbach DR, Tomlinson GA, Harnish JL, et al. A measure of disease-specific health-related quality of life for achalasia. Am J Gastroenterol. 2005; 100:1668–1676. [PubMed: 16144120]
24. Cranney A, Zarkadas M, Graham ID, et al. The Canadian celiac health survey—the Ottawa chapter pilot. BMC Gastroenterol. 2003; 3:8. [PubMed: 12740024]
25. Hauser W, Gold J, Stein J, et al. Health-related quality of life in adult coeliac disease in Germany: results of a national survey. Eur J Gastroenterol Hepatol. 2006; 18:747–754. [PubMed: 16772832]
26. De Luca L, Zagari RM, Pozzato P, et al. Measuring dyspepsia: a new severity index validated in Bologna. Dig Liver Dis. 2004; 36:806–810. [PubMed: 15646426]
27. Leidy NK, Farup C, Rentz AM, et al. Patient-based assessment in dyspepsia: development and validation of Dyspepsia Symptom Severity Index (DSSI). Dig Dis Sci. 2000; 45:1172–1179. [PubMed: 10877234]
28. Madsen LG, Hansen JM, Gronvold M, et al. The validity of a symptom diary in ratings of dyspepsia measured against a detailed interview: do patients and clinicians agree in their assessment of symptoms? Aliment Pharmacol Ther. 2007; 26:905–912. [PubMed: 17767475]
29. Mujakovic S, de Wit NJ, van Marrewijk CJ, et al. Psychopathology is associated with dyspeptic symptom severity in primary care patients with a new episode of dyspepsia. Aliment Pharmacol Ther. 2009; 29:580–588. [PubMed: 19053982]
30. Rothman M, Farup C, Stewart W, et al. Symptoms associated with gastroesophageal reflux disease: development of a questionnaire for use in clinical trials. Dig Dis Sci. 2001; 46:1540–1549. [PubMed: 11478508]
31. Kulich KR, Malfertheiner P, Madisch A, et al. Psychometric validation of the German translation of the Gastrointestinal Symptom Rating Scale (GSRS) and Quality of Life in Reflux and Dyspepsia (QOLRAD) questionnaire in patients with reflux disease. Health Qual Life Outcomes. 2003; 1:62. [PubMed: 14613560]
32. Adam B, Liebregts T, Saadat-Gilani K, et al. Validation of the gastrointestinal symptom score for the assessment of symptoms in patients with functional dyspepsia. Aliment Pharmacol Ther. 2005; 22:357–363. [PubMed: 16098003]

33. Revicki DA, Rente AM, Dubois D, et al. Gastroparesis Cardinal Symptom Index (GCSI): development and validation of a patient reported assessment of severity of gastroparesis symptoms. Qual Life Res. 2004; 13:833–844. [PubMed: 15129893]
34. Hu WH, Lam KF, Wong YH, et al. The Hong Kong index of dyspepsia: a validated symptom severity questionnaire for patients with dyspepsia. 1 Gastroenterol Hepatol. 2002; 17:545–551.
35. Talley NJ, Haque M, Wyeth JW, et al. Development of a new dyspepsia impact scale: the Nepean Dyspepsia Index. Aliment Pharmacol Ther. 1999; 13:225–235. [PubMed: 10102954]
36. Sander GB, Mazzoleni LE, Francesconi CF, et al. Development and validation of a cross-cultural questionnaire to evaluate nonulcer dyspepsia: the Porto Alegre Dyspeptic Symptoms Questionnaire (PADYQ). Dig Dis Sci. 2004; 49:1822–1829. [PubMed: 15628711]
37. de la Loge C, Trudeau E, Marquis P, et al. Cross-cultural development and validation of a patient self-administered questionnaire to assess quality of life in upper gastrointestinal disorders: the PAGI-QOL. Qual Life Res. 2004; 13:1751–1762. [PubMed: 15651545]
38. Revicki DA, Rente AM, Tack J, et al. Responsiveness and interpretation of a symptom severity index specific to upper gastrointestinal disorders. Clin Gastroenterol Hepatol. 2004; 2:769–777. [PubMed: 15354277]
39. Armstrong D, Veldhuyzen SJ, Chung SA, et al. Validation of a short questionnaire in English and French for use in patients with persistent upper gastrointestinal symptoms despite proton pump inhibitor therapy: the PASS (Proton pump inhibitor Acid Suppression Symptom) test. Can J Gastroenterol. 2005; 19:350–358. [PubMed: 15997268]
40. Wiklund IK, 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 Suppl. 1998; 583:41–49. [PubMed: 10027672]
41. Mouzas IA, Fragkiadakis N, Moschandreas J, et al. Validation and results of a questionnaire for functional bowel disease in out-patients. BMC Public Health. 2002; 2:8. [PubMed: 12022923]
42. Arslan G, Lind R, Olafsson S, et al. Quality of life in patients with subjective food hypersensitivity: applicability of the 10-item short form of the Nepean Dyspepsia Index. Dig Dis Sci. 2004; 49:680–687. [PubMed: 15185878]
43. Fraser A, Delaney BC, Ford AC, et al. The Short-Form Leeds Dyspepsia Questionnaire validation study. Aliment Pharmacol Ther. 2007; 25:477–486. [PubMed: 17270004]
44. Flood EM, Beusterien KM, Amonkar MM, et al. Patient and caregiver perspective on pediatric eosinophilic esophagitis and newly developed symptom questionnaires*. Curr Med Res Opin. 2008; 24:3369–3381. [PubMed: 19032119]
45. Brandt ML, Daigneau C, Graviss EA, et al. Validation of the Baylor Continence Scale in children with anorectal malformations. J Pediatr Surg. 2007; 42:1015–1021. discussion 21. [PubMed: 17560212]
46. de Cassia Domansky R, de Gouveia Santos VL. Cross-cultural adaptation and validation of the Portuguese version of the Bowel Function in the Community instrument. J Wound Ostomy Continence Nurs. 2007; 34:671–677. [PubMed: 18030108]
47. Krysa J, Lyons M, Williams AB. A simple quality of life questionnaire for patients with faecal incontinence. Int J Colorectal Dis. 2009; 24:1213–1217. [PubMed: 19575210]
48. Temple LK, Bacik J, Savatta SG, et al. The development of a validated instrument to evaluate bowel function after sphincter-preserving surgery for rectal cancer. Dis Colon Rectum. 2005; 48:1353–1365. [PubMed: 15868235]
49. Bakx R, Sprangers MA, Oort FJ, et al. Development and validation of a colorectal functional outcome questionnaire. Int J Colorectal Dis. 2005; 20:126–136. [PubMed: 15449078]
50. Macmillan AK, Merrie AE, Marshall RJ, et al. Design and validation of a comprehensive fecal incontinence questionnaire. Dis Colon Rectum. 2008; 51:1502–1522. [PubMed: 18626716]
51. Due U, Ottesen M. The Danish anal sphincter rupture questionnaire: validity and reliability. Acta Obstet Gynecol Scand. 2009; 88:36–42. [PubMed: 19023680]
52. Lukacz ES, Lawrence JM, Buckwalter JG, et al. Epidemiology of prolapse and incontinence questionnaire: validation of a new epidemiologic survey. Int Urogynecol J Pelvic Floor Dysfunct. 2005; 16:272–284. [PubMed: 15856132]

53. Bharucha AE, Locke GR III, Seide BM, et al. A new questionnaire for constipation and faecal incontinence. Aliment Pharmacol Ther. 2004; 20:355–364. [PubMed: 15274673]
54. Reilly WT, Talley NJ, Pemberton JH, et al. Validation of a questionnaire to assess fecal incontinence and associated risk factors: fecal incontinence questionnaire. Dis Colon Rectum. 2000; 43:146–153. discussion 53-4. [PubMed: 10696886]
55. Rockwood TH, Church JM, Fleshman JW, et al. Fecal Incontinence Quality of Life Scale: quality of life instrument for patients with fecal incontinence. Dis Colon Rectum. 2000; 43:9–16. discussion −7. [PubMed: 10813117]
56. Bug GJ, Kiff ES, Hosker G. A new condition-specific health-related quality of life questionnaire for the assessment of women with anal incontinence. BJOG. 2001; 108:1057–1067. [PubMed: 11702838]
57. Abayomi J, Kirwan J, Hackett A. The prevalence of chronic radiation enteritis following radiotherapy for cervical or endometrial cancer and its impact on quality of life. Eur J Oncol Nurs. 2009; 13:262–27. [PubMed: 19640788]
58. Khanna D, Hays RD, Maranian P, et al. Reliability and validity of the University of California, Los Angeles Scleroderma Clinical Trial Consortium Gastrointestinal Tract Instrument. Arthritis Rheum. 2009; 61:1257–1263. [PubMed: 19714600]
59. Bols EM, Hendriks EJ, Deutekom M, et al. Inconclusive psychometric properties of the Vaizey score in fecally incontinent patients: a prospective cohort study. Neurourol Urodyn. 2010; 29:370–377. [PubMed: 19634170]
60. Nanigian DK, Nguyen T, Tanaka ST, et al. Development and validation of the fecal incontinence and constipation quality of life measure in children with spina bifida. J Urol. 2008; 180:1770–1773. discussion 3. [PubMed: 18721959]
61. Roalfe AK, Roberts LM, Wilson S. Evaluation of the Birmingham IBS symptom questionnaire. BMC Gastroenterol. 2008; 8:30. [PubMed: 18651941]
62. Wiklund IK, Fullerton S, Hawkey CJ, et al. An irritable bowel syndrome-specific symptom questionnaire: development and validation. Scand J Gastroenterol. 2003; 38:947–954. [PubMed: 14531531]
63. van der Veek PP, Van Rood YR, Masclee AA. Symptom severity but not psychopathology predicts visceral hypersensitivity in irritable bowel syndrome. Clin Gastroenterol Hepatol. 2008; 6:321–328. [PubMed: 18258487]
64. Schmulson M, Ortiz O, Mejia-Arangure JM, et al. Further validation of the IBS-QOL: female Mexican IBS patients have poorer quality of life than females from North Carolina. Dig Dis Sci. 2007; 52:2950–2955. [PubMed: 17415635]
65. Jones MP, Keefer L, Bratten J, et al. Development and initial validation of a measure of perceived stigma in irritable bowel syndrome. Psychol Health Med. 2009; 14:367–374. [PubMed: 19444714]
66. Bengtsson M, Ohlsson B, Ulander K. Development and psychometric testing of the Visual Analogue Scale for Irritable Bowel Syndrome (VAS-IBS). BMC Gastroenterol. 2007; 7:16. [PubMed: 17475020]
67. Labus JS, Bolus R, Chang L, et al. The Visceral Sensitivity Index: development and validation of a gastrointestinal symptom-specific anxiety scale. Aliment Pharmacol Ther. 2004; 20:89–97. [PubMed: 15225175]
68. Reilly MC, Bracco A, Ricci JF, et al. The validity and accuracy of the Work Productivity and Activity Impairment questionnaire—irritable bowel syndrome version (WPALIBS). Aliment Pharmacol Ther. 2004; 20:459–467. [PubMed: 15298641]
69. Luther SL, Nelson AL, Harrow JJ, et al. A comparison of patient outcomes and quality of life in persons with neurogenic bowel: standard bowel care program vs colostomy. J Spinal Cord Med. 2005; 28:387–393. [PubMed: 16869085]
70. 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. 1988; 33:129–134. [PubMed: 3123181]
71. Wong E, Guyatt GH, Cook DJ, et al. Development of a questionnaire to measure quality of life in patients with irritable bowel syndrome. Eur J Surg Suppl. 1998; 583:50–56. [PubMed: 10027673]

72. Groll D, Vanner SJ, Depew WT, et al. The IBS-36: a new quality of life measure for irritable bowel syndrome. Am J Gastroenterol. 2002; 97:962–971. [PubMed: 12003433]

73. Patrick DL, Drossman DA, Frederick IO, et al. Quality of life in persons with irritable bowel syndrome: development and validation of a new measure. Dig Dis Sci. 1998; 43:400–411. [PubMed: 9512138]

74. Hahn BA, Kirchdoerfer LJ, Fullerton S, et al. Evaluation of a new quality of life questionnaire for patients with irritable bowel syndrome. Aliment Pharmacol Ther. 1997; 11:547–552. [PubMed: 9218081]

75. Gordon S, Ameen V, Bagby B, et al. Validation of irritable bowel syndrome Global Improvement Scale: an integrated symptom end point for assessing treatment efficacy. Dig Dis Sci. 2003; 48:1317–1323. [PubMed: 12870789]

76. Chan AO, Lam KF, Hui WM, et al. Validated questionnaire on diagnosis and symptom severity for functional constipation in the Chinese population. Aliment Pharmacol Ther. 2005; 22:483–488. [PubMed: 16128687]

77. Woolery M, Carroll E, Fenn E, et al. A constipation assessment scale for use in pediatric oncology. J Pediatr Oncol Nurs. 2006; 23:65–74. [PubMed: 16476780]

78. Varma MG, Wang JY, Berian JR, et al. The constipation severity instrument: a validated measure. Dis Colon Rectum. 2008; 51:162–172. [PubMed: 18172725]

79. Wang JY, Hart SL, Lee J, et al. A valid and reliable measure of constipation-related quality of life. Dis Colon Rectum. 2009; 52:1434–1442. [PubMed: 19617757]

80. Bongers ME, Benninga MA, Maurice-Stam H, et al. Health-related quality of life in young adults with symptoms of constipation continuing from childhood into adulthood. Health Qual Life Outcomes. 2009; 7:20. [PubMed: 19254365]

81. Voskuijl WP, van der Zaag-Loonen HJ, Ketel IJ, et al. Health related quality of life in disorders of defecation: the Defecation Disorder List. Arch Dis Child. 2004; 89:1124–1127. [PubMed: 15557046]

82. Knowles CH, Eccersley AJ, Scott SM, et al. Linear discriminant analysis of symptoms in patients with chronic constipation: validation of a new scoring system (KESS). Dis Colon Rectum. 2000; 43:1419–1426. [PubMed: 11052520]

83. Altomare DF, Spazzafumo L, Rinaldi M, et al. Set-up and statistical validation of a new scoring system for obstructed defaecation syndrome. Colorectal Dis. 2008; 10:84–88. [PubMed: 17441968]

84. Marquis P, De La Loge C, Dubois D, et al. Development and validation of the patient assessment of constipation quality of life questionnaire. Scand J Gastroenterol. 2005; 40:540–551. [PubMed: 16036506]

85. Frank L, Kleinman L, Farup C, et al. Psychometric validation of a constipation symptom assessment questionnaire. Scand J Gastroenterol. 1999; 34:870–877. [PubMed: 10522604]

86. Downing GM, Kuziemsky C, Lesperance M, et al. Development and reliability testing of the Victoria Bowel Performance Scale (BPS). J Pain Symptom Manage. 2007; 34:513–522. [PubMed: 17664054]

87. Jones R, Coyne K, Wiklund I. The gastro-oesophageal reflux disease impact scale: a patient management tool for primary care. Aliment Pharmacol Ther. 2007; 25:1451–1459. [PubMed: 17539985]

88. Liu JY, Woloshin S, Laycock WS, et al. Symptoms and treatment burden of gastroesophageal reflux disease: validating the GERD assessment scales. Arch Int Med. 2004; 164:2058–2064. [PubMed: 15477443]

89. Puhan MA, Guyatt GH, Armstrong D, et al. Validation of a symptom diary for patients with gastro-oesophageal reflux disease. Aliment Pharmacol Ther. 2006; 23:531–541. [PubMed: 16441474]

90. Velanovich V. The development of the GERD-HRQL symptom severity instrument. Dis Esophagus. 2007; 20:130–134. [PubMed: 17439596]

91. Jones R, Junghard O, Dent J, et al. Development of the GerdQ, a tool for the diagnosis and management of gastro-oesophageal reflux disease in primary care. Aliment Pharmacol Ther. 2009; 30:1030–1038. [PubMed: 19737151]

92. Chan Y, Ching JY, Cheung CM, et al. Development and validation of a disease-specific quality of life questionnaire for gastro-oesophageal reflux disease: the GERD-QOL questionnaire. Aliment Pharmacol Ther. 2010; 31:452–460. [PubMed: 19878152]

93. Holtmann G, Chassany O, Devault KR, et al. International validation of a health-related quality of life questionnaire in patients with erosive gastro-oesophageal reflux disease. Aliment Pharmacol Ther. 2009; 29:615–625. [PubMed: 19183145]

94. Pare P, Meyer F, Armstrong D, et al. Validation of the GSFQ, a self-administered symptom frequency questionnaire for patients with gastroesophageal reflux disease. Can J Gastroenterol. 2003; 17:307–312. [PubMed: 12772004]

95. Revicki DA, Wood M, Wiklund I, et al. Reliability and validity of the Gastrointestinal Symptom Rating Scale in patients with gastroesophageal reflux disease. Qual Life Res. 1998; 7:75–83. [PubMed: 9481153]

96. Shikiar R, Flood E, Siddique R, et al. Development and validation of the gastroesophageal reflux disease treatment satisfaction questionnaire. Dig Dis Sci. 2005; 50:2025–2033. [PubMed: 16240210]

97. Malaty HM, O'Malley KJ, Abudayyeh S, et al. Multidimensional measure for gastroesophageal reflux disease (MM-GERD) symptoms in children: a population-based study. Acta Paediatr. 2008; 97:1292–1297. [PubMed: 18510718]

98. Teixeira JP, Mosquera V, Flores A. Long-term outcomes of quality of life after laparoscopic Nissen fundoplication. Hepatogastroenterology. 2009; 56:80–84. [PubMed: 19453033]

99. De La Loge C, Trudeau E, Marquis P, et al. Responsiveness and interpretation of a quality of life questionnaire specific to upper gastrointestinal disorders. Clin Gastroenterol Hepatol. 2004; 2:778–786. [PubMed: 15354278]

100. Kim J, Keininger DL, Becker S, et al. Simultaneous development of the Pediatric GERD Caregiver Impact Questionnaire (PGCIQ) in American English and American Spanish. Health Qual Life Outcomes. 2005; 3:5. [PubMed: 15651991]

101. Gill SK, Maltepe C, Koren G. The effect of heartburn and acid reflux on the severity of nausea and vomiting of pregnancy. Can J Gastroenterol. 2009; 23:270–272. [PubMed: 19373420]

102. Dan D, Seetahal S, Naraynsingh V. Laparoscopic Nissen fundoplication for improved gastrointestinal symptoms and quality of life. West Indian Med J. 2009; 58:8–12. [PubMed: 19565992]

103. Radajewski R, Hazebroek EJ, Berry H, et al. Short-term symptom and quality-of-life comparison between laparoscopic Nissen and Toupet fundoplications. Dis Esophagus. 2009; 22:84–88. [PubMed: 19018854]

104. Zeman Z, Rozsa S, Tihanyi T, et al. Psychometric documentation of a quality-of-life questionnaire for patients undergoing antireflux surgery (QOLARS). Surgl Endosc. 2005; 19:257–261.

105. Shaw MJ, Talley NJ, Beebe TJ, et al. Initial validation of a diagnostic questionnaire for gastroesophageal reflux disease. Am J Gastroenterol. 2001; 96:52–57. [PubMed: 11197287]

106. Belafsky PC, Postma GN, Koufman JA. Validity and reliability of the reflux symptom index (RSI). J Voice. 2002; 16:274–277. [PubMed: 12150380]

107. Bardhan KD. Look--but also listen! ReQuest: a new dimension-oriented GERD symptom scale. Drugs Today (Bare). 2004; 40(Suppl A):15–19.

108. Macran S, Wileman S, Barton G, et al. The development of a new measure of quality of life in the management of gastro-oesophageal reflux disease: the Reflux questionnaire. Qual Life Res. 2007; 16:331–343. [PubMed: 17033907]

109. Coyne KS, Wiklund I, Schmier J, et al. Development and validation of a disease-specific treatment satisfaction questionnaire for gastro-oesophageal reflux disease. Aliment Pharmacol Ther. 2003; 18:907–915. [PubMed: 14616154]

110. Allen ML, Leslie K, Jansen N. Validation of the postoperative nausea and vomiting intensity score in gynaecological patients. Anaesth Intensive Care. 2011; 39:73–78. [PubMed: 21375094]

111. Phipps E, Braitman LE, Stites S, et al. Quality of life and symptom attribution in long-term colon cancer survivors. J Eval Clin Pract. 2008; 14:254–258. [PubMed: 18284521]

112. Nakamura M, Kido Y, Yano M, et al. Reliability and validity of a new scale to assess postoperative dysfunction after resection of upper gastrointestinal carcinoma. Surg Today. 2005; 35:535–542. [PubMed: 15976949]
113. Ikenaga N, Nishihara K, Iwashita T, et al. Long-term quality of life after laparoscopically assisted distal gastrectomy for gastric cancer. J Laparo-endosc Adv Surg Tech A. 2006; 16:119–123.
114. Portenoy RK, Thaler HT, Kornblith AB, et al. The Memorial Symptom Assessment Scale: an instrument for the evaluation of symptom prevalence, characteristics and distress. Eur J Cancer. 1994; 30A:1326–1336. [PubMed: 7999421]
115. Dupuis LL, Taddio A, Kerr EN, et al. Development and validation of the pediatric nausea assessment tool for use in children receiving antineoplastic agents. Pharmacotherapy. 2006; 26:1221–1231. [PubMed: 16945043]
116. Tokunaga M, Hiki N, Ohyama S, et al. Effects of reconstruction methods on a patients quality of life after a proximal gastrectomy: subjective symptoms evaluation using questionnaire survey. Langenbecks Arch Surg. 2009; 394:637–641. [PubMed: 19066939]
117. Satoh S, Sato K, Habuchi T, et al. Health-related quality of life of ileocecal rectal bladder compared with ileal conduit diversion: a questionnaire survey. Int J Urol. 2002; 9:385–391. discussion 91. [PubMed: 12165020]
118. Scarpa M, Barollo M, Polese L, et al. Quality of life in patients with an ileostomy. Minerva Chir. 2004; 59:23–29. [PubMed: 15111829]
119. Smith GD, Watson R, Palmer KR. Inflammatory bowel disease: developing a short disease specific scale to measure health related quality of life. Int J Nurs Stud. 2002; 39:583–590. [PubMed: 12100869]
120. Martin A, Leone L, Fries W, et al. Quality of life in inflammatory bowel disease. Ital J Gastroenterol. 1995; 27:450–454. [PubMed: 8775474]
121. Loonen HJ, Grootenhuis MA, Last BF, et al. Measuring quality of life in children with inflammatory bowel disease: the impact-II (NL). Qual Life Res. 2002; 11:47–56. [PubMed: 12003055]
122. Walmsley RS, Ayres RC, Pounder RE, et al. A simple clinical colitis activity index. Gut. 1998; 43:29–32. [PubMed: 9771402]
123. Griffiths AM, Nicholas D, Smith C, et al. Development of a quality-of-life index for pediatric inflammatory bowel disease: dealing with differences related to age and IBD type. J Pediatr Gastroenterol Nutr. 1999; 28:S46–S52. [PubMed: 10204526]
124. Turner D, Otley AR, Mack D, et al. Development, validation, and evaluation of a pediatric ulcerative colitis activity index: a prospective multicenter study. Gastroenterology. 2007; 133:423–432. [PubMed: 17681163]
125. Irvine EJ, Zhou Q, Thompson AK. The Short Inflammatory Bowel Disease Questionnaire: a quality of life instrument for community physicians managing inflammatory bowel disease. CCRPT Investigators Canadian Crohn's Relapse Prevention Trial. Am J Gastroenterol. 1996; 91:1571–1578. [PubMed: 8759664]
126. Guassora AD, Kruuse C, Thomsen OO, et al. Quality of life study in a regional group of patients with Crohn disease A structured interview study. Scand J Gastroenterol. 2000; 35:1068–1074. [PubMed: 11099060]
127. Coyne K, Joshua-Gotlib S, Kimel M, et al. Validation of the treatment satisfaction questionnaire for Crohn's disease (TSQ-C). Dig Dis Sci. 2005; 50:252–258. [PubMed: 15745081]
128. Lacasse A, Berard A. Validation of the nausea and vomiting of pregnancy specific health related quality of life questionnaire. Health Qual Life Outcomes. 2008; 6:32. [PubMed: 18471301]
129. Koren G, Boskovic R, Hard M, et al. Motherisk-PUQE (pregnancy-unique quantification of emesis and nausea) scoring system for nausea and vomiting of pregnancy. Am J Obstetr Gynecol. 2002; 186:S228–S231.
130. Magee LA, Chandra K, Mazzotta P, et al. Development of a health-related quality of life instrument for nausea and vomiting of pregnancy. Am J Obstetr Gynecol. 2002; 186:S232–S238.
131. Khanna D, Hays RD, Park GS, et al. Development of a preliminary scleroderma gastrointestinal tract 1.0 quality of life instrument. Arthritis Rheum. 2007; 57:1280–1286. [PubMed: 17907224]

132. Khanna D. Minimally important differences of the UCLA scleroderma clinical trial consortium gastrointestinal tract instrument. J Rheumatol. 2011; 38:1920–1924. [PubMed: 21724699]
133. Baron M. Validation of the UCLA scleroderma clinical trial gaatro-intestinal tract instrument version 2.0 for systemic sclerosis. J Rheumatol. 2011; 38:1925–1930. [PubMed: 21724704]
134. Brunner HI, Johnson AL, Barron AC, et al. Gastrointestinal symptoms and their association with health-related quality of life of children with juvenile rheumatoid arthritis: validation of a gastrointestinal symptom questionnaire. J Clin Rheumatol. 2005; 11:194–204. [PubMed: 16357756]
135. Caplan A, Walker L, Rasquin A. Development and preliminary validation of the questionnaire on pediatric gastrointestinal symptoms to assess functional gastrointestinal disorders in children and adolescents. J Pediatr Gastroenterol Nutr. 2005; 41:296–304. [PubMed: 16131984]
136. Kleinman L, Nelson S, Kothari-Talwar S, et al. Development and psychometric evaluation of 2 age-stratified versions of the Pediatric GERD Symptom and Quality of Life Questionnaire. J Pediatr Gastroenterol Nutr. 2011; 52:514–522. [PubMed: 21464757]
137. Baxter AL, Watcha MF, Baxter WV, et al. Development and validation of a pictorial nausea rating scale for children. Pediatrics. 201l; 127:el542–e1549.
138. Hiller L, Radley S, Mann CH, et al. Development and validation of a questionnaire for the assessment of bowel and lower urinary tract symptoms in women. BJOG. 2002; 109:413–423. [PubMed: 12013162]
139. Markland AD, Richter HE, Burgio KL, et al. Outcomes of combination treatment of fecal incontinence in women. Am J Obstetr Gynecol. 2008; 199:el–e7. 699.
140. Casellas F, Varela E, Aparici A, et al. Development, validation, and applicability of a symptoms questionnaire for lactose malabsorption screening. Dig Dis Sci. 2009; 54:1059–1065. [PubMed: 18716868]
141. Spiegel B, Bolus R, Harris LA, et al. Measuring irritable bowel syndrome patient-reported outcomes with an abdominal pain numeric rating scale. Aliment Pharmacol Ther. 2009; 30:1159–1170. [PubMed: 19751360]
142. Drossman DA, Patrick DL, Whitehead WE, et al. Further validation of the IBS-QOL: a disease-specific quality-of-life questionnaire. Am J Gastroenterol. 2000; 95:999–1007. [PubMed: 10763950]
143. Bardhan KD, Stanghellini V, Armstrong D, et al. International validation of ReQuest in patients with endoscopy-negative gastro-oesophageal reflux disease. Aliment Pharmacol Ther. 2004; 20:891–898. [PubMed: 15479361]
144. Beirens K, Fontaine JR. Development of the Ghent Multidimensional Somatic Complaints Scale. Assessment. 2010; 17:70–80. [PubMed: 19700738]
145. Dubois D, Gilet H, Viala-Danten M, et al. Psychometric performance and clinical meaningfulness of the Patient Assessment of Constipation-Quality of Life questionnaire in prucalopride (RESOLOR) trials for chronic constipation. Neurogastroenterol Motil. 2010; 22:e54–e63. [PubMed: 19761492]
146. Grant AM, Wileman SM, Ramsay CR, et al. Minimal access surgery compared with medical management for chronic gastro-oesophageal reflux disease: UK collaborative randomised trial. BMJ. 2008; 337:a2664. [PubMed: 19074946]
147. Monnikes H, Bardhan KD, Stanghellini V, et al. Evaluation of GERD symptoms during therapy. Part II Psychometric evaluation and validation of the new questionnaire ReQuest in erosive GERD. Digestion. 2004; 69:238–244. [PubMed: 15256830]
148. Molassiotis A, Coventry PA, Strieker CT, et al. Validation and psychometric assessment of a short clinical scale to measure chemotherapy-induced nausea and vomiting: the MASCC antiemesis tool. J Pain Symptom Manage. 2007; 34:148–159. [PubMed: 17509816]
149. Power Z, Campbell M, Kilcoyne P, et al. The Hyperemesis Impact of Symptoms Questionnaire: development and validation of a clinical tool. Int J Nurs Stud. 2010; 47:67–77. [PubMed: 19646694]
150. Cella D, Riley W, Stone A, et al. Initial item banks and first wave testing of the Patient-Reported Outcomes Measurement Information System (PROMIS) Network: 2005–2008. J Clin Epidemiol. 2010; 63:1179–1194. [PubMed: 20685078]

151. Choi SW, Reise SP, Pilkonis PA, et al. Efficiency of static and computer adaptive short forms compared to full-length measures of depressive symptoms. Qual Life Res. 2010; 19:125–136. [PubMed: 19941077]

152. Greenhalgh J. The applications of PROs in clinical practice: what are they, do they work, and why? Qual Life Res. 2009; 18:115–123. [PubMed: 19105048]

153. Beatty PC, Willis GB. Research synthesis: the practice of cognitive interviewing. Public Opin Q. 2007; 71:287–311.

154. Willis, GB. Cognitive Interviewing: A Tool for Improving Questionnaire Design. Sage Publications; Thousand Oaks, CA: 2005.

155. Greenhalgh J, Long AF, Flynn R. The use of patient reported outcome measures in routine clinical practice: lack of impact or lack of theory? Soc Sci Med. 2005; 60:833–843. [PubMed: 15571900]

156. Monnikes H, Bardhan KD, Stanghellini V, et al. Evaluation of GERD symptoms during therapy. Part II. Psychometric evaluation and validation of the new questionnaire ReQuest in erosive GERD. Digestion. 2007; 75(Suppll):41–47. [PubMed: 17489031]

157. Bardhan KD, Stanghellini V, Armstrong D, et al. Evaluation of GERD symptoms during therapy. Part I. Development of the new GERD questionnaire ReQuest. Digestion. 2007; 75(Suppl 1):32–40. [PubMed: 17489030]

158. Kulich KR, Madisch A, Pacini F, et al. Reliability and validity of the Gastrointestinal Symptom Rating Scale (GSRS) and Quality of Life in Reflux and Dyspepsia (QOLRAD) questionnaire in dyspepsia: a six-country study. Health Qual Life Outcomes. 2008; 6:12. [PubMed: 18237386]

159. Wyrwich KW, Tardino VM. A blueprint for symptom scales and responses: measurement and reporting. Gut. 2004; 53(Suppl 4):iv45–iv48. [PubMed: 15082614]

160. Loonen HJ, Grootenhuis MA, Last BF, et al. Quality of life in paediatric inflammatory bowel disease measured by a generic and a disease-specific questionnaire. Acta Paediatr. 2002; 91:348–354. [PubMed: 12022311]

161. Guyatt G, Mitchell A, Irvine EJ, et al. A new measure of health status for clinical trials in inflammatory bowel disease. Gastroenterology. 1989; 96:804–810. [PubMed: 2644154]

162. Labus JS, Mayer EA, Chang L, et al. The central role of gastrointestinal-specific anxiety in irritable bowel syndrome: further validation of the visceral sensitivity index. Psychosom Med. 2007; 69:89–98. [PubMed: 17244851]

163. Patrick DL, Drossman DA, Frederick IO, et al. Quality of life in persons with irritable bowel syndrome: development and validation of a new measure. Dig Dis Sci. 1998; 43:400–411. [PubMed: 9512138]

164. Drossman DA, Patrick DL, Whitehead WE, et al. Further validation of the IBS-QOL: a disease-specific quality-of-life questionnaire. Am J Gastroenterol. 2000; 95:999–1007. [PubMed: 10763950]

165. Schmulson MJ, Ortiz-Garrido OM, Hinojosa C, et al. A single session of reassurance can acutely improve the self-perception of impairment in patients with IBS. J Psychosom Res. 2006; 61:461–467. [PubMed: 17011353]

166. Drossman DA, Morris CB, Schneck S, et al. International survey of patients with IBS: symptom features and their severity, health status, treatments, and risk taking to achieve clinical benefit. J Clin Gastroenterol. 2009; 43:541–550. [PubMed: 19384249]

167. Delaney BC, Qume M, Moayyedi P, et al. Helicobacter pylori test and treat versus proton pump inhibitor in initial management of dyspepsia in primary care: multicentre randomised controlled trial (MRC-CUBE trial). BMJ. 2008; 336:651–654. [PubMed: 18310262]

168. Salem EM, Yar T, Bamosa AO, et al. Comparative study of Nigella Sativa and triple therapy in eradication of Helicobacter Pylori in patients with non-ulcer dyspepsia. Saudi J Gastroenterol. 2010; 16:207–214. [PubMed: 20616418]

169. Kulich KR, Calabrese C, Pacini F, et al. Psychometric validation of the Italian translation of the gastrointestinal symptom-rating scale and quality of life in reflux and dyspepsia questionnaire in patients with gastro-oesophageal reflux disease. Clin Drug Investig. 2004; 24:205–215.

170. Talley NJ, Fullerton S, Junghard O, et al. Quality of life in patients with endoscopy-negative heartburn: reliability and sensitivity of disease-specific instruments. Am J Gastroenterol. 2001; 96:1998–2004. [PubMed: 11467624]

APPENDIX 1

List of searchable terms

1.	Abdominal pain
2.	Achalasia
3.	Acid reflux
4.	Bloat
5.	Bowel control
6.	Celiac sprue
7.	Constipation
8.	Crohn’s
9.	Defecation
10.	Depression
11.	Diarrhea
12.	Dyspepsia
13.	Dysphagia
14.	Emesis
15.	Eosinophilic esophagitis
16.	Evacuation
17.	Fatigue
18.	Fecal incontinence
19.	Flatus
20.	Fullness
21.	Functional bowel disease
22.	Functional GI disorders
23.	Gas
24.	Gastroparesis
25.	GERD
26.	Health behavior
27.	Heartburn
28.	Hiccups
29.	Ileal conduit diversion
30.	Ileostomy
31.	Impact
32.	Inflammatory bowel disease
33.	Infrequent stools
34.	Irritable bowel
35.	Irritable bowel syndrome

36.	Loose stool
37.	Low appetite
38.	Malabsorption
39.	Nausea
40.	Odynophagia
41.	Pain
42.	Pediatric
43.	Postgastrectomy
44.	Pregnancy-related GI symptoms
45.	Quality of life
46.	Radiation enteritis
47.	Reflux
48.	Regurgitation
49.	Restrictions
50.	Scleroderma
51.	Sleep
52.	Soilage
53.	Somatic complaints
54.	Stool
55.	Swallowing
56.	Systemic sclerosis
57.	Throat burn
58.	Urgency
59.	Vomiting

GERD, gastroesophageal reflux disease; GI, gastrointestinal.

Study Highlights

WHAT IS CURRENT KNOWLEDGE

- ✓ There are over 100 patient-reported outcome (PRO) questionnaires in gastroenterology.
- ✓ Clinicians and investigators use PROs, such as health-related quality of life (HRQOL), to guide clinical decision making, conduct research, and seek drug approval.
- ✓ It is important to develop a mechanism for easily identifying, categorizing, and rating these PROs.

WHAT IS NEW HERE

- ✓ We have developed a new, National Institutes of Health (NIH)-supported, online PRO clearinghouse—the GI-PROs database—to assist with selecting GI PROs for clinical and research purposes: www.researchcore.org/gipro/.
- ✓ The GI-PRO database is a publicly available, user-friendly, search engine that is searchable at multiple levels, including PROs, domains, and methodological quality.
- ✓ GI-PROs database does not include clinical outcomes that include biomarkers or physician assessments; instead, it focuses only on outcomes that are *patient reported*.

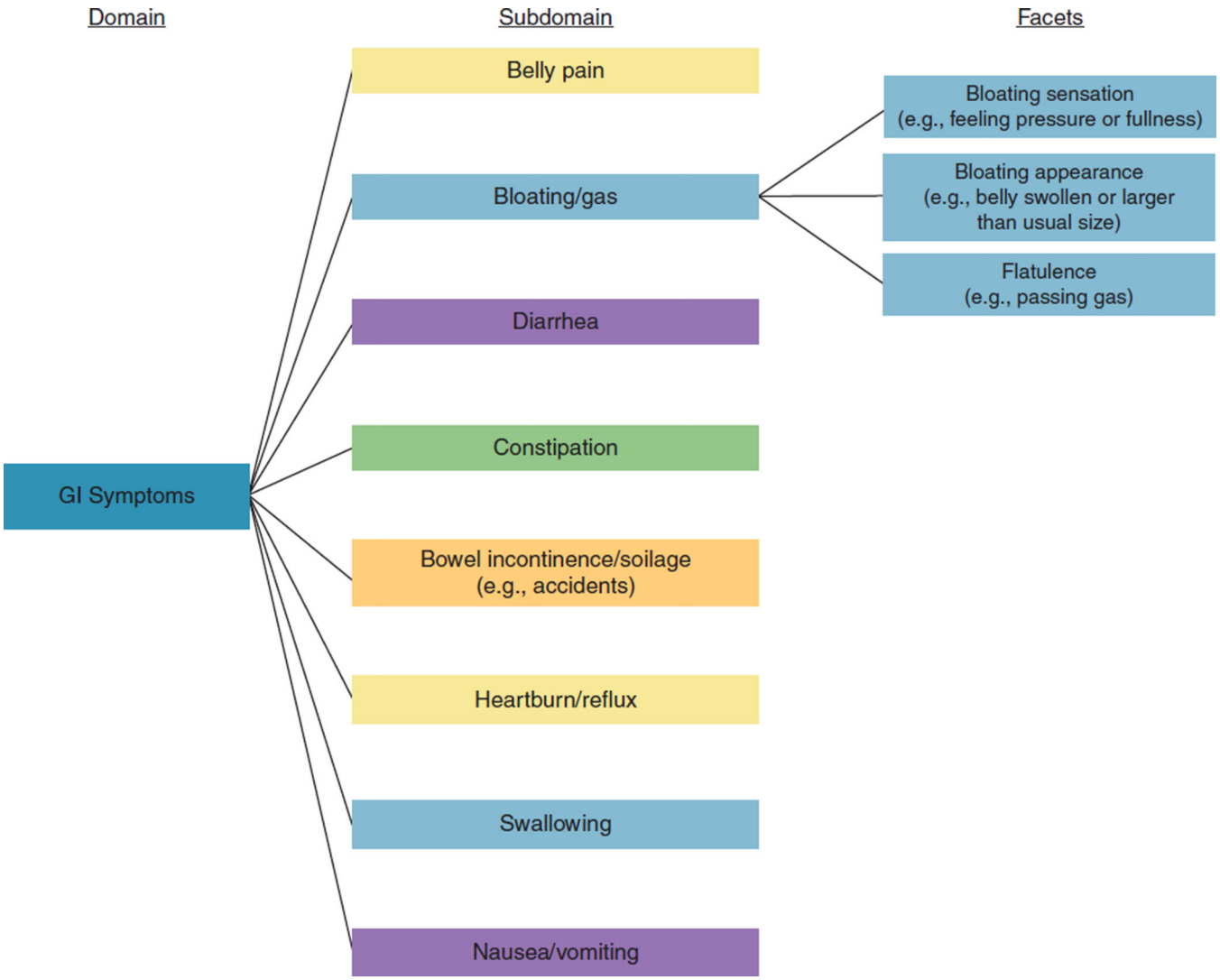

Figure 1.

Conceptual framework of gastrointestinal (GI) symptom bins from items in published GI patient-reported outcome (PRO) instruments.

Figure 2.

Screenshot of GI PRO online library. Users initially view a text box in which any term can be entered that is relevant for a search of gastrointestinal (GI) patient-reported outcome (PRO) instruments. Examples include names of known instruments (e.g., IBS-QOL, PAGI-QOL, IBDQ), names of diseases (e.g., gastroesophageal reflux disease (GERD), dyspepsia, Crohn's, Celiac), symptoms (e.g., bloating, diarrhea, constipation, nausea, pain), or health-related quality-of-life domains (e.g., sleep, fatigue, impact, embarrassment, depression). The example below shows the result of searching for "GERD". Users can select individual instruments to obtain detailed information about them (e.g., inset shows details of ReQuest instrument).

Table 1

Search strategy for GI PROs

Group	Search terms	Significance of grouping
1	PubMed; Cochrane database; Patient-Reported Outcomes and Quality of Life Instruments Database (PROQOUD)	Targeted bibliographic databases
2	"Inflammatory Bowel Diseases"[MeSH] OR "Inflammatory Bowel Disease"[text word] OR "Inflammatory Bowel Diseases"[text word] OR "Ulcerative colitis"[text word] OR "Crohn disease"[text word] OR "Crohn's disease"[text word] OR "Crohns disease"[text word] OR "ileocolitis"[text word] OR "terminal ileitis"[text word] OR "regional ileitis"[text word] OR "granulomatous colitis"[text word] OR "granulomatous enteritis"[text word] OR "regional enteritis" [text word] OR "Irritable Bowel Syndrome"[MESH] OR "Irritable colon "[text word] OR "Mucous colitis"[text word] OR "Mucous colitides "[text word] OR "Irritable bowel syndrome"[text word] OR "Irritable bowel syndromes"[text word] OR "Scleroderma, Systemic"[MESH] OR "Scleroderma"[text word] OR "Systemic Sclerosis" [text word] OR "CREST Syndrome "[text word] OR (("Intestines"[MeSH] OR "Lower Gastrointestinal Tract"[MeSH] OR "Upper Gastrointestinal Tract"[MeSH] OR Gastrointestin*[text word] OR Intestine*[text word] OR Colon*[text word] OR Duodenum*[text word] OR Esophag*[text word] OR Stomach*[text word] OR Ileum[text word] OR Ileocecal[text word] OR Jejunum*[text word] OR Cecum*[text word] OR Anal[text word] OR Anus[text word] OR Rectum*[text word]) AND ("Pain"[MeSH] OR pain*[text word] OR ache[text word] OR aching[text word] OR discomfort*[text word])) OR "Defecation"[MeSH] OR defecat*[text word] OR bloat*[text word] OR "Dilatation, Pathologic"[MeSH] OR disten*[text word] OR "Dyspepsia"[MeSH] OR dyspepsia*[text word] OR indigestion*[text word] OR "Flatulence"[MeSH] OR flatu*[text word] OR "Nausea"[MeSH] OR nausea*[text word] OR nauseous[text word] OR vomit*[text word] OR "Diarrhea"[MeSH] OR "fecal incontinence" [MESH] OR Constipation[MeSH] OR diarrhea*[text word] OR "fecal incontinence"[text word] OR constipat*[text word] OR "stool frequency"[text word] OR "stool frequencies"[text word] OR "feces"[MeSH] OR feces[text word] OR fecal[text word] OR "stool form"[text word] OR "Gastroesophageal Reflux"[MeSH] OR "GERD"[text word] OR "GORD" [text word] OR ("functional bowel "[text word] OR "FBD"[text word] OR "FGID" [text word]) OR ("functional"[text word] AND "bowel"[text word])	Targeted GI keywords and content
3	"psychometrics"[MeSH] OR psychometr*[tiab] OR (cronbach*[tiab] AND (alpha[tiab] OR alphasltiab])) OR "test-retest"[tiab] OR (test[tiab] AND retest[tiab]) OR (reliab*[tiab] AND (test[tiab] OR retest[tiab])) OR "factor analysis"[tiab] OR "factor analyses"[tiab] OR subscale*[tiab] OR (multitraitltiab] AND scaling[tiab] AND (analysis[tiab] OR analyses[tiab])) OR ("minimal clinical important difference"[tiab] OR "minimally clinical important difference"[tiab] OR "Minimal clinically important difference"[tiab] OR "Minimally clinically important difference" [tiab] OR "Minimal important difference"[tiab] OR MCID[tiab] OR MID[tiab]) OR "Item response model"[tiab] OR IRT[tiab] OR Rasch[tiab] OR "Differential item functioning"[tiab] OR DIF[tiab] OR "computer adaptive testing"[tiab] OR "item bank"[tiab] OR "patient reported outcome"[tiab] OR psychosoc*[tiab] OR "Models, Psychological"[MESH] OR "quality of life"[MeSH] OR "Severity of Illness Index"[MESH] OR "severity index"[tiab] OR "activity index"[tiab]	Validated screen for PRO measures (12,16) (modified for current search strategy based on β testing with biomedical librarian)
4	NOT ("animals"[MeSH] NOT "humans"[MeSH]OR Letter [pt] OR Editorial [pt] OR Review [pt] OR News [pt])	Excluded study types and content

GI, gastrointestinal; PRO, patient-reported outcome.

The four search groups were combined as follows: (1 AND 2 AND 3 NOT 4). The search focused exclusively on English-language publications.

Table 2

Checklist for methodological quality scoring

1. Item generation (content validity)
a. 0=No patient or provider item generative input
b. 1=Only provider input (e.g., focus groups, expert panel, review of draft items)
c. 2=Only patient input (e.g., focus groups, open-ended content elicitation survey, or interview)
d. 3=Both patient and provider inputs obtained
2. Evaluating draft items (cognitive interviews; same interviewee can be used for all the items in the survey)
a. 0=No cognitive interviews reported
b. 1=Some cognitive interviews but < 5 per item or unable to determine number
c. 2=5–9 cognitive interviews per item
d. 3=10 cognitive interviews per item
3. Reliability of measures (internal consistency reliability preferred over test-retest; median of all estimates)
a. 0=Not reported
b. 0=<0.70
c. 1=0.70–0.79
d. 2=0.80–0.89
e. 3=0.90
4. Reliability for multiple English-language subgroups (includes item response theory information curves)
a. 0=Not reported
b. 1=Reported
5. Construct validity (a *priori* hypotheses are compared with empirical associations)
a. 0=No data reported
b. 1=Cross-sectional data support *a priori* hypotheses
c. 2=Some longitudinal support for *a priori* hypotheses
d. 3=Extensive support (cross-sectional and longitudinal support in multiple subgroups)
6. Interpretsbility (what do score *differences* mean—e.g., minimal important differences, responsiveness to change, cutpoints)
a. 0=No information on how to interpret scores provided
b. 1=lnformation from cross-sectional differences in scores
c. 2=Some information from longitudinal differences in scores
d. 3=Extension information on interpretation of scores
7. Response burden (estimated time to complete survey; grade-level readability estimates such as Flesch-Kincaid or Lexile)
a. 0=No information on time to complete or English-language readability
b. 1=lnformation on time to complete only
c. 1=lnformation on English-language readability only
d. 2=lnformation on both time to complete and English-language readability
e. 3=Time to complete is <15min and English-language readability reported
8. Mode of administration (e.g., mail, phone, web—do *not* include different language versions as modes)
a. 0=0nlv one mode of administration (mode:_____)
b. 1=Multiple modes of administration (modes:_____)

The checklist evaluated patient-reported outcomes (PROs) using all published articles with information on psychometric properties of the instrument. There are eight checklist items and the possible score range is 0–20.

Availability of the instrument in languages other than English is not factored in because the primary focus is use of the instrument in English language. Other coding approaches (e.g., EMPRO and COSMIN (COnsensus-based Standards for the selection of health status Measurement Instruments)) give credit for availability of language translations.

Table 3

Names of psychological and social subscales across GI PROs identified from literature search

Behaviors	**Activity limitations**
	Dietary habits
	Clothing
	Leisure activities
Cognitions	Body image
	Expectations
	Health perception
	Knowledge and control
	Satisfaction
	Role
	Worries and concerns
	Burden of illness
	Coping/behavior
	Self-perception
Emotions	Embarrassment
	Depression
	Emotional well-being
	Mental health
Psychosocial impact	Employment status
	HRQOL
	Productivity
	Interference with daily activities
	Lifestyle
	Personal function
	Disability
	Medical care utilization
	Physician relationships
	Personal relationships
	Sexual function

GI, gastrointestinal; HRQOL, health-related quality of life; PRO, patient-reported outcome.

Table 4

Results of quality scoring

Number of instruments	126
Mean score±s.d.	8.88±4.19
Total number of items	2,372
Item generation	24% No provider/patient feedback 14% Patient feedback only 27% Provider feedback only 35% Patient and provider feedback
Cognitive interviews	46% No cognitive interviews 9% Less than 5 cognitive interviews 2% 5–9 cognitive interviews 42% 10 or more cognitive interviews
Reliability[a]	22% Not reported 10% < 0.70 21% 0.70–0.79 24% 0.80–0.89 23% 0.90 or higher
Reliability/info for multiple English language subgroups	88% Not reported 12% Reported
Construct validity	15% No data reported 55% Cross-sectional data support 25% Some longitudinal support 5% Extensive support
Interpretability	41% No information 28% Some info from cross-sectional differences 17% Some info from longitudinal differences 13% Extensive information
Response burden	73% No information provided 16% Info on time to complete only 4% Info on English-language readability only 2% Info on time to complete and English-language readability 5% Time to complete <15min and English-language readability reported

All results given in percent of total instruments reported.

[a] Reliability of measures was graded by the median of all estimates of internal consistency (see table 2 for the complete checklist for methodological quality scoring).

Table 5

Highest scoring (0–20) instruments by disease category

Disease	PRO	Quality score
GERD	ReQUEST (143,156,157)	18
	GSAS (31,70,95,158)	15
	PAGI-QOL (37,159)	15
Constipation	PAC-QOL (84,145)	15
IBD	IMPACT (81,121,160)	16
	IBDQ (126,161)	14
Fecal incontinence	Comprehensive Fecal Incontinence Questionnaire (50)	18
IBS	VSI (67,141,162)	14
	IBS-QOL (163–166)	14
Dyspepsia	Leeds (43,167,168)	15
	QOLRAD (40,169,170)	14

GERD, gastroesophageal reflux disease; IBD, inflammatory bowel disease; IBS, irritable bowel syndrome; PRO, patient-reported outcome.