see related editorial on page 1557

Impact of National Institutes of Health Gastrointestinal PROMIS Measures in Clinical Practice: Results of a Multicenter Controlled Trial

Christopher V. Almario, MD, MSHPM[1,2,3], William D. Chey, MD[4], Dinesh Khanna, MD, MSc[5], Sasan Mosadeghi, MD, MS[3], Shahzad Ahmed, MD[3], Elham Afghani, MD, MPH[1], Cynthia Whitman, MPH[1,3], Garth Fuller, MS[1,3], Mark Reid, PhD[2,3], Roger Bolus, PhD[3], Buddy Dennis, PhD[6], Rey Encarnacion, BS[6], Bibiana Martinez, MPH[1,2,3], Jennifer Soares, MSPH[1,2,3], Rushaba Modi, MD[3], Nikhil Agarwal, MD[3], Aaron Lee, MD[3], Scott Kubomoto, MD[3], Gobind Sharma, MD[3], Sally Bolus, MS[3] and Brennan M.R. Spiegel, MD, MSHS[1,2,3]

OBJECTIVES: **The National Institutes of Health (NIH) created the Patient Reported Outcomes Measurement Information System (PROMIS) to allow efficient, online measurement of patient-reported outcomes (PROs), but it remains untested whether PROMIS improves outcomes. Here, we aimed to compare the impact of gastrointestinal (GI) PROMIS measures vs. usual care on patient outcomes.**

METHODS: **We performed a pragmatic clinical trial with an off-on study design alternating weekly between intervention (GI PROMIS) and control arms at one Veterans Affairs and three university-affiliated specialty clinics. Adults with GI symptoms were eligible. Intervention patients completed GI PROMIS symptom questionnaires on an e-portal 1 week before their visit; PROs were available for review by patients and their providers before and during the clinic visit. Usual care patients were managed according to customary practices. Our primary outcome was patient satisfaction as determined by the Consumer Assessment of Healthcare Providers and Systems questionnaire. Secondary outcomes included provider interpersonal skills (Doctors' Interpersonal Skills Questionnaire (DISQ)) and shared decision-making (9-item Shared Decision Making Questionnaire (SDM-Q-9)).**

RESULTS: **There were 217 and 154 patients in the GI PROMIS and control arms, respectively. Patient satisfaction was similar between groups ($P>0.05$). Intervention patients had similar assessments of their providers' interpersonal skills (DISQ 89.4±11.7 vs. 89.8±16.0, $P=0.79$) and shared decision-making (SDM-Q-9 79.3±12.4 vs. 79.0±22.0, $P=0.85$) vs. controls.**

CONCLUSIONS: **This is the first controlled trial examining the impact of NIH PROMIS in clinical practice. One-time use of GI PROMIS did not improve patient satisfaction or assessment of provider interpersonal skills and shared decision-making. Future studies examining how to optimize PROs in clinical practice are encouraged before widespread adoption.**

Am J Gastroenterol 2016; 111:1546–1556; doi:10.1038/ajg.2016.305; published online 2 August 2016

INTRODUCTION

Patients often seek care because they experience symptoms that negatively impact health-related quality of life. Healthcare providers must elicit, measure, and interpret patient symptoms as part of their clinical evaluation. To assist with this goal, researchers have developed and validated a wide range of patient-reported outcomes (PROs) across diseases, with a focus on chronic illnesses (1–3). These PROs, which measure any aspect of a patient's biopsychosocial health that comes directly from the patient, may help direct care and improve outcomes. When PROs are collected

[1]Division of Gastroenterology, Cedars-Sinai Medical Center, Los Angeles, California, USA; [2]Division of Gastroenterology, VA Greater Los Angeles Healthcare System, Los Angeles, California, USA; [3]Cedars-Sinai Center for Outcomes Research and Education (CS-CORE), Los Angeles, California, USA; [4]Division of Gastroenterology, University of Michigan, Ann Arbor, Michigan, USA; [5]Division of Rheumatology, University of Michigan, Ann Arbor, Michigan, USA; [6]UCLA Computing Technology Research Laboratory (CTRL), Los Angeles, California, USA. **Correspondence:** Brennan M.R. Spiegel, MD, MSHS, Division of Gastroenterology, Cedars-Sinai Medical Center, Pacific Theatres Building, 116 North Robertson Boulevard, 4th Floor, Los Angeles, California 90048, USA. E-mail: Brennan.Spiegel@cshs.org
Received 8 March 2016; accepted 22 June 2016

systematically, efficiently, and in the right place at the right time, they may enhance the patient–provider relationship at the center of chronic disease care, improve communication, and help make shared decisions (4–6).

However, despite the promise of using PROs to guide patient care, there are important challenges to applying PROs in routine practice (7–12). For example, it can be time consuming to collect PROs from patients and securely transmit the data into the electronic health record (EHR), making it untenable for use in busy practices. There are also many PROs to choose from, with a lack of measurement standards across questionnaires. Furthermore, clinicians note that it can be difficult to understand and act upon PRO scores. When coupled with limited evidence from previous research that administering PROs truly impacts patient outcomes (2), these challenges limit widespread use of PROs in clinical practice; most providers instead opt for informal measurement of symptoms and function.

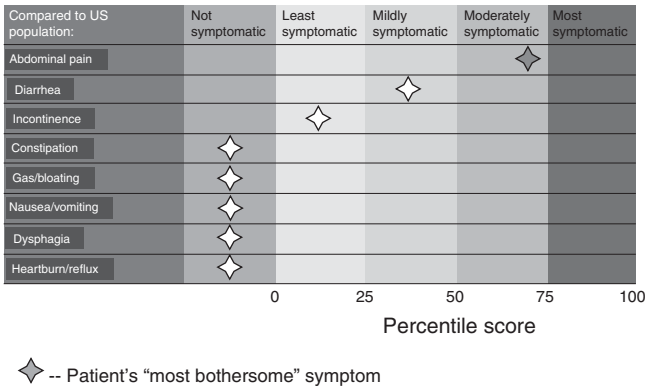
In this context, the National Institutes of Health (NIH) created the Patient Reported Outcomes Measurement Information System (PROMIS) in 2004 with the goal of developing, validating, and disseminating a toolbox of publicly available PROs that cover the breadth and depth of the human illness experience while overcoming technical challenges of applying PROs in practice (http: //www.nihpromis.org) (13). Using modern psychometric techniques, such as item response theory and computerized adaptive testing (14,15), PROMIS offers state-of-the-art psychometrics, establishes common-language benchmarks for symptoms across conditions, and identifies clinical thresholds for action and meaningful clinical improvement or decline. PROMIS questionnaires are administered electronically and efficiently, allowing implementation in busy clinical settings. Because of the extraordinary burden of illness from digestive diseases, the PROMIS consortium added a gastrointestinal (GI) item bank, which our group developed (16). Using the NIH PROMIS framework, we constructed and validated eight GI PROMIS symptom scales using data from over 2,000 subjects (16–18).

However, despite over a decade of NIH PROMIS development, it remains unclear whether implementing GI PROMIS, let alone any PROMIS measures, can improve patient outcomes vs. usual care. In this study, we conducted an NIH-supported multicenter-controlled trial of PROMIS vs. usual care in clinical practice. Specifically, we used GI PROMIS measures in diverse patients with active GI symptoms, collected the results via a patient–provider e-portal, and presented the data at the point of care. We hypothesized that compared with usual care, use of GI PROMIS would enhance the patient–provider interaction, leading to improved patient satisfaction and higher patient assessment of provider interpersonal skills and shared decision-making.

METHODS

Study overview

We performed a pragmatic, multicenter clinical trial comparing use of validated GI PROMIS questionnaires (16) vs. usual care in diverse patients with active GI symptoms, including those with abdominal pain, bowel incontinence, bloating/gas, constipation, diarrhea, dysphagia, heartburn/reflux, and nausea/vomiting. We administered the GI PROMIS questionnaires through a secure, online, patient–provider e-portal (see **Appendix Figure 1** for sample screenshots). The portal collected the PRO data and converted responses into a symptom "heat map" (**Figure 1**) that visually compared each patient's symptoms against the general US population (16,18). Both patients and providers could view this heat map on the portal before and during the clinic visit.



HPI: Mr. Smith is a 34-year-old male who reports a history of Celiac disease and now presents with abdominal pain. The pain first started 8 months ago, and typically lasts for 2 hours at a time. Over the past week, the pain occurred once a day. He describes the pain as "burning" and "gnawing", says it is located in the epigastrium, and reports the pain has been "quite severe" and "quite a bit bothersome" in the past week. It does not radiate. It is associated with eating food. It typically occurs around 10–30 minutes after starting to eat. It usually comes on suddenly. It is not associated with bowel movements. The pain is somewhat relieved by reducing stress. The pain does not awaken him from sleep. He does not report early satiety. He does not report diabetes, gallstones, GERD, pancreatitis, or peptic ulcer. He does not take aspirin or NSAIDs.

He also reports diarrhea and bowel incontinence. The patient does not report dysphagia, heartburn, bloating, constipation, nausea, or vomiting.

He does not report blood in his bowel movements, black stools, vomiting blood, unintended weight loss, diminished appetite, or fevers. He has no history of abdominal surgeries. There is a family history of colorectal cancer.

Figure 1. Sample "heat map" report of gastrointestinal Patient Reported Outcomes Measurement Information System (GI PROMIS) scores and history of present illness (HPI). Patients complete PROMIS items on the e-portal, and the results are converted into a GI PROMIS symptom heat map and HPI. Patients' PROMIS scores are compared with the general US population with benchmarks to add interpretability to the scores, similar to a lab test. Both the heat map and HPI are viewable on the e-portal for both the patient and healthcare provider before the clinic visit.

To enhance clinical applicability of GI PROMIS, the e-portal autocomposed a complete GI history of present illness (HPI) report triggered off the PROMIS symptoms. Patients were guided through a set of questions measuring the timing, severity, frequency, location, quality, and character of each reported GI PROMIS symptom, along with relevant comorbidities, family history, and alarm features (19,20). Once the questions were completed, the information was transformed into a full narrative GI HPI that accompanied the PROMIS heat map (**Figure 1**). In a previous head-to-head trial comparing GI PROMIS-directed computerized HPIs vs. physician HPIs, we found that the computerized HPIs were rated by blinded reviewers to be of higher quality and more thorough, complete, succinct, and relevant (19). However,

the previous trial did not measure the impact of the PROMIS on patient outcomes.

By tying GI PROMIS scores to a focused HPI, using specific GI symptoms with benchmarked interpretation, and directly presenting the results to the provider at the point of care, we attempted to optimize the impact of using PROMIS. In this manner, the current study sought to overcome traditional critiques of using PROs in clinical practice: i.e., technical difficulties of transmitting to the EHR, interpretability, data visualization issues, and clinical actionability.

Study design, patients, and setting

We used a pragmatic, off-on study design alternating weekly between the PROMIS intervention and control arms. Patients who visited the following clinics were eligible for the study: (i) Cedars-Sinai Medical Center general GI clinic; (ii) West Los Angeles Veterans Affairs (WLAVA) Medical Center general GI clinic; (iii) University of Michigan functional GI and motility clinic; and (iv) University of Michigan scleroderma clinic (selected because scleroderma patients have a high prevalence of GI symptoms). The Cedars-Sinai and WLAVA GI clinics are academic teaching practices staffed by GI attending physicians; the initial evaluation in these clinics were primarily conducted by GI specialty fellows, internal medicine residents, or physician assistants. Conversely, attending physicians primarily staffed the GI and scleroderma clinics at the University of Michigan.

We enrolled patients, aged ≥18 years, who were scheduled for an initial visit or had not been seen in the clinic within the past 8 months. Patients were also required to read and write English and possess basic point-and-click computing skills.

During the control weeks, patients were treated according to all customary practices. In the intervention weeks, eligible patients were mailed a letter 1 week before their appointment inviting them to log on to the e-portal to complete GI PROMIS. Eligible intervention patients who did not complete PROMIS before their visit were also approached during the day of their appointment by research staff and again invited to access the e-portal on a clinic computer before seeing their physician. Clinic providers were informed to access the e-portal and view the GI PROMIS symptom heat map for patients who completed PROMIS. In keeping with our pragmatic approach to the study, providers were not mandated to use the PROMIS data or PROMIS-directed HPI. Rather, providers were allowed to make individual decisions on how to use the PROMIS data report, if at all.

Within 24 h of completing the clinic visit, patients were sent the postvisit questionnaires to measure their satisfaction with the visit as well as their assessment of their providers' interpersonal skills and shared decision-making. This study was approved by the Institutional Review Boards (IRBs) at all sites (Cedars-Sinai IRB Pro00041476; University of Michigan IRB HUM00063094; WLAVA IRB PCC no. 2013-111563).

Primary and secondary outcomes

The primary outcome was patient satisfaction as measured by the Consumer Assessment of Healthcare Providers and Systems Clinician and Group Survey 2.0 (CG-CAHPS) (21). Because the CG-CAHPS is a global assessment of patients' satisfaction with their medical care over the past year, we used selected items that were applicable for assessing patient satisfaction after a single visit. Patients were reminded to answer the questions thinking about their most recent visit to the GI or scleroderma clinic. The answer options for most selected CG-CAHPS items were "Yes, definitely," "Yes, somewhat," and "No." We used a "top-box" approach, which is commonly used when reporting CG-CAHPS data (22); a positive response included only "Yes, definitely," while negative responses included "Yes, somewhat" or "No."

Our secondary outcomes were patient assessments of provider interpersonal skills and shared decision-making. Patients completed the Doctors' Interpersonal Skills Questionnaire (DISQ) to assess their provider's interpersonal skills (23). The DISQ comprised 12 items, each scored on a 5-point scale, where 1="Poor" and 5="Excellent." We converted each item to a 100-point scale and averaged the scores for the 12 items to calculate an overall interpersonal skills score.

We used the 9-item Shared Decision Making Questionnaire (SDM-Q-9) to assess patient shared decision-making (24). The SDM-Q-9 contained nine items, each scored on a 6-point scale, where 1="Completely Disagree" and 6="Completely Agree." Similar to the DISQ, we converted each item to a 100-point scale and averaged the nine scores to calculate an overall shared decision-making score. Again, for both DISQ and SDM-Q-9, patients were informed to answer the questions thinking about their most recent GI or scleroderma clinic visit.

Covariates

We also collected information on potentially confounding patient- and provider-level variables. Patient-level factors included age, gender, and race/ethnicity. We also collected provider-level factors, including site of care and provider level of training.

Sample size calculation

Our primary objective was to measure differences in CG-CAHPS provider rating scores between groups. Although CG-CAHPS is widely used and accepted as a measure of patient satisfaction with outpatient visits, we are unaware of data measuring the minimally clinically important difference on the scale. Therefore, the sample size was calculated to achieve an effect size of 0.5 (a half standard deviation difference) in mean CG-CAHPS provider rating scores between groups—an effect size that is moderate and generally correlates with the minimally clinically important difference (25,26). Assuming a two-tailed 5% significance level with a power of 80%, the minimum sample size needed to show an effect size of 0.5 was 64 patients per group.

Statistical and sensitivity analyses

Statistical analyses were performed using Stata 13.1 (StataCorp LP, College Station, TX). A two-tailed P value <0.05 was considered significant. Our primary analyses were performed from the intention-to-treat perspective. For intervention patients who completed GI PROMIS, but did not return the postvisit out-

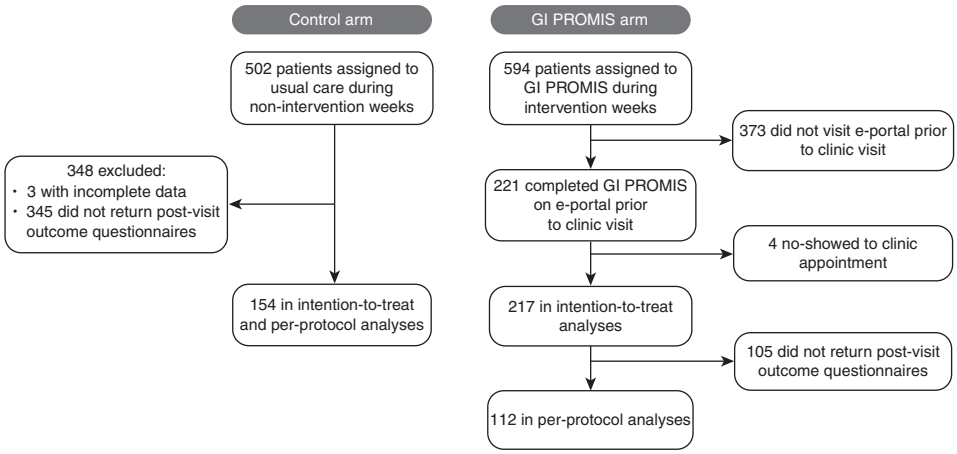


Figure 2. Flow diagram of enrolled patients. For the control group, intention-to-treat and per-protocol analyses included those who returned the postvisit outcome questionnaires. For the gastrointestinal Patient Reported Outcomes Measurement Information System (GI PROMIS) arm, the intention-to-treat analyses included those who completed GI PROMIS and who showed for their visit; missing outcome data was imputed to the corresponding mean value calculated from controls for each item. Per-protocol analyses for the GI PROMIS arm only included those who completed GI PROMIS and the postvisit outcome questionnaires.

come surveys, we assumed their outcomes (CG-CAHPS provider rating, DISQ, and SDM-Q-9) were no different than controls. Specifically, the missing outcome data for this group was imputed to the corresponding mean value calculated from controls for each item. Because this assumption biases towards the null, we also performed a sensitivity analysis using a per-protocol approach where we excluded patients without follow-up data.

For bivariate analyses, we used the two-sample *t*-test and χ^2 test to compare means and proportions, respectively, between groups. We performed a multivariable logistic regression model to identify patient characteristics that were independent predictors of completing PROMIS before the clinic visit.

We used linear regression to generate an adjusted *P*-value and to evaluate differences in CG-CAHPS provider ratings between groups while adjusting for potential confounding patient- and provider-level factors. We used similar approaches when comparing the remaining CG-CAHPS items (Firth logistic regression) and the DISQ and SDM-Q-9 items (linear regression).

RESULTS

Study population

Figure 2 shows patient flow through the clinical trial. Overall, 502 patients were assigned to the control arm and 3 (0.6%) had missing demographic data. Of the 499 with complete covariate data, 154 (30.9%) completed the postvisit outcome questionnaires. Significant differences were seen between completers and non-completers with respect to age, gender, and site of care; no difference in race/ethnicity was noted between groups (**Appendix Table 1**). **Table 1** presents the demographics of patients in the control arm.

For the intervention group, 594 (0% missing demographic data) were invited to complete GI PROMIS before their clinic visit. Among those invited, 221 (37.2%) accessed the e-portal and completed the questionnaires. A majority of the patients who completed PROMIS attended their clinic appointment (217/221; 98.2%). Of the 217 individuals who completed PROMIS and attended their clinic visit, 112 (51.6%) completed the postvisit outcome assessments. Significant differences were seen in age and site of care between those who did and did not return the surveys; no differences were seen in gender and race/ethnicity between groups (**Appendix Table 2**). In **Table 1**, we list the demographics of those in the GI PROMIS arm.

Predictors of completing GI PROMIS

Table 2 shows the results from the multivariable regression on completion of GI PROMIS before the clinic visit. Age and gender were not independent predictors of completing PROMIS. African Americans were less likely to access the e-portal vs. whites (odds ratio 0.44; 95% confidence interval 0.26, 0.74); no differences were seen between whites and the remaining racial/ethnic groups (Latino, Asian, Other/Unknown). Conversely, patients seen at the University of Michigan GI clinic were more likely to complete PROMIS vs. patients at the WLAVA GI clinic (odds ratio 7.96; 95% confidence interval 4.19, 15.1).

Primary and secondary outcomes

Intention-to-treat analyses. **Table 3** presents the CG-CAHPS provider rating scores for the GI PROMIS and control arms in the intention-to-treat analysis. After adjusting for confounders, we found no difference in provider rating between groups. **Tables 4** and **5** list differences in provider interpersonal skills and shared decision-making, respectively. We found no difference in DISQ scores between the PROMIS and control arms. Both groups also had similar shared decision-making scores as by the SDM-Q-9.

Per-protocol analyses. Because the intention-to-treat analysis biases results towards the null, we also performed a per-protocol analysis. Here, there were 154 individuals in the control group and

Table 1. Patient demographics

Variable	Control arm (*n*=154)	GI PROMIS arm (*n*=217)[a]	*P* value
Age (years)	58.7±15.8	54.1±16.3	0.007
Male	67 (43.5%)	109 (50.2%)	0.20
Race/ethnicity			0.24
Caucasian	100 (64.9%)	151 (69.6%)	
African American	32 (20.8%)	28 (12.9%)	
Asian	5 (3.3%)	8 (3.7%)	
Latino	9 (5.8%)	11 (5.1%)	
Other/unknown	8 (5.2%)	19 (8.8%)	
Site of care			0.08
West Los Angeles VA GI clinic	55 (35.7%)	77 (35.5%)	
Cedars-Sinai GI clinic	13 (8.4%)	11 (5.1%)	
University of Michigan GI clinic	77 (50.0%)	100 (46.1%)	
University of Michigan scleroderma clinic	9 (5.8%)	29 (13.4%)	
Provider level of training			0.29
GI or rheumatology attending	89 (57.8%)	138 (63.6%)	
GI fellow	37 (24.0%)	52 (24.0%)	
Internal medicine resident or GI PA	28 (18.2%)	27 (12.4%)	

GI, gastrointestinal; PA, physician assistant; PROMIS, Patient Reported Outcome Measurement Information System; VA, Veterans Affairs.
Data are presented as mean±s.d. or *n* (%).
Columns may not add up to 100% owing to rounding.
[a]Four of the 221 individuals who completed PROMIS no-showed for their clinic appointment.

Table 2. Predictors of completing PROMIS

Variable	Completed GI PROMIS (*n*=221)	OR (95% CI)[a]
Age (years)	—	1.00 (0.98, 1.01)
Gender		
Male	113 (29.1%)	Reference
Female	108 (52.7%)	1.25 (0.72, 2.16)
Race/ethnicity		
Caucasian	152 (48.6%)	Reference
African American	28 (18.4%)	0.44 (0.26, 0.74)
Asian	9 (60.0%)	2.38 (0.76, 7.43)
Latino	13 (19.7%)	0.52 (0.26, 1.05)
Other/unknown	19 (39.6%)	1.03 (0.52, 2.03)
Site of care		
West Los Angeles VA GI clinic	81 (23.4%)	Reference
Cedars-Sinai GI clinic	11 (29.7%)	1.02 (0.44, 2.37)
University of Michigan GI clinic	100 (79.4%)	7.96 (4.19, 15.1)
University of Michigan scleroderma clinic	29 (34.1%)	1.07 (0.53, 2.14)

CI, confidence interval; GI, gastrointestinal; OR, odds ratio; PROMIS, Patient Reported Outcome Measurement Information System; VA, Veterans Affairs.
Data are presented as *n* (%).
[a]The multivariable logistic regression included all variables in the table.

112 in the GI PROMIS arm. There were no differences between groups with respect to age, gender, race/ethnicity, site of care, or provider level of training (all $P>0.05$).

The results were similar: there was no difference in CG-CAHPS provider ratings between groups (control 8.93±1.65 vs. GI PROMIS 8.84±1.64; adjusted $P=0.76$). Both groups had similar patient satisfaction scores for the remaining CG-CAHPS items (**Table 3**). Patient assessment of provider interpersonal skills (**Appendix Table 3**) and shared decision-making (**Appendix Table 4**) were also similar between both arms.

DISCUSSION

To our knowledge, this is the first controlled trial evaluating the impact of PROMIS on patient outcomes in clinical practice. Despite theoretical benefits of measuring GI PROs to drive clinical decision-making, we found no differences in patient satisfaction or assessment of provider interpersonal skills and shared decision-making between those in the NIH GI PROMIS and control arms. These results suggest that simply measuring GI PROMIS scores may be insufficient to meaningfully improve patients' interaction with the healthcare system. These results are also in line with a systematic review that found inconsistent benefits of applying PROs in clinical practice (2).

There are several possible explanations for our negative results. First, despite the wide use of CG-CAHPS for assessing patient satisfaction, the minimally clinically important difference for the scale is unknown; it is possible that the study was underpowered to detect a significant and meaningful difference between groups. Because of this issue, we calculated the sample size to achieve a moderate effect size of 0.5, which prior research found generally correlates with the minimally clinically important difference (25,26). Second, patients reported high levels of satisfaction in both arms of this trial. This result may have led to a "ceiling effect"; it is possible that PROMIS on its own may not offer incremental improvements among patients who are already satisfied with their provider. Third, only 30.9% and 51.6% of patients in the control and intervention arms, respectively, completed the postvisit outcome questionnaires; we cannot know if outcomes would be different in survey non-responders. Fourth, we focused on a proximal outcome of patient satisfaction after a single clinic visit; it is possible that longitudinal use of GI PROMIS (i.e., to track GI symptom improvement and response to therapies) may have led to improved patient satisfaction over time. Last, in keeping with the study's pragmatic approach, we did not mandate or assess the use of PROMIS scores

Table 3. Patient satisfaction assessment

CG-CAHPS item	Control arm	GI PROMIS arm	Adjusted *P* value
Provider explained things in a way that was easy to understand[a]	122/139 (87.8%)	91/101 (90.1%)	0.75[b]
Provider listened carefully[a]	126/139 (90.7%)	92/101 (91.1%)	0.90[b]
Provider gave patient easy to understand information about health questions or concerns[a,c]	108/132 (81.8%)	85/97 (87.6%)	0.25[b]
Provider seemed to know the important information about patient's medical history[a,c]	95/132 (72.0%)	72/97 (74.2%)	0.52[b]
Provider showed respect for what patient had to say[a,c]	121/132 (91.7%)	91/97 (93.8%)	0.18[b]
Provider spent enough time with patient[a]	121/139 (87.1%)	93/101 (92.1%)	0.23[b]
Provider rating (0–10 scale)[d]	8.93±1.65	8.88±1.15	0.94[e]

CG-CAHPS, Consumer Assessment of Healthcare Providers and Systems Clinician and Group Survey; GI, gastrointestinal; PROMIS, Patient Reported Outcome Measurement Information System.

Data are presented as *n* (%) or mean±s.d.

[a]Per-protocol analysis (*n*=240; there was incomplete CG-CAHPS data for 26 patients).

[b]The Firth logistic regression model adjusted for patient- (age, gender, race/ethnicity) and provider-level factors (site of care, level of training).

[c]These questions were not required for the 11 patients who stated that they did not talk with their provider about any health questions or concerns.

[d]Intention-to-treat analysis (*n*=345; there was incomplete data for this item for 26 patients).

[e]The linear regression model adjusted for patient- (age, gender, race/ethnicity) and provider-level factors (site of care, level of training).

Table 4. Patient assessment of provider interpersonal skills

DISQ item	Control arm (*n*=144)	GI PROMIS arm (*n*=209)	Adjusted *P* value[a]
Overall satisfaction with provider	89.7±17.7	88.9±13.4	0.67
Warmth of provider's greeting	90.3±16.4	89.5±12.2	0.72
Ability to listen to the patient	90.4±17.2	89.5±12.7	0.56
Adequacy of explanations to patient	89.4±16.7	88.9±12.7	0.63
Extent of reassurance provided to patient	87.8±18.9	87.6±13.9	0.86
Confidence in provider's ability	90.6±17.5	90.3±13.2	0.77
Opportunity for patient to express concerns and fears	89.7±17.7	88.9±13.5	0.68
Respect shown to patient	91.3±16.3	92.5±9.9	0.33
Time given for visit	88.5±18.9	88.5±13.3	0.86
Consideration of patient's personal situation in treatment or advice	89.7±18.5	88.8±13.4	0.66
Concern for patient as a person	89.9±16.9	89.3±12.7	0.68
Recommendation of provider to friends	90.3±18.5	89.4±13.4	0.67
Average DISQ score	89.8±16.0	89.4±11.7	0.79

DISQ, Doctors' Interpersonal Skills Questionnaire; GI, gastrointestinal; PROMIS, Patient Reported Outcome Measurement Information System.

Data are presented as mean±s.d.

DISQ scores are on a 100-point scale. Complete DISQ data was unavailable for 18 patients.

[a]The linear regression model adjusted for patient- (age, gender, race/ethnicity) and provider-level factors (site of care, level of training).

or PROMIS-directed HPI reports by providers. It is possible that some clinicians did not use the report and managed the patient according to their customary practices. While we could have tested the efficacy of GI PROMIS in a tightly controlled setting by mandating that all providers use PROMIS and incorporate it into their patient assessment, we instead sought to test the effectiveness of PROMIS in a setting that more resembles the "real world."

Our study also has limitations with respect to external generalizability. We only evaluated patients with GI symptoms, thus we cannot know whether using other PROMIS questionnaires, such as those for fatigue, physical function, or pain, among many others, would also fail to show a difference vs. usual care. Moreover, our trial was conducted solely in clinics affiliated with academic universities. It is possible that outcomes may be different when GI PROMIS is used in non-university-based clinics, but that must be formally tested and it is the subject of our future research.

Despite these limitations, we found no differences between groups. Even with *post hoc* analyses searching for differences on an item-by-item basis, the groups were equivalent. This is consistent with existing literature that administering PROs, although conceptually appealing, often fails to meaningfully improve patient outcomes vs. usual care (2). Notably, we attempted to overcome this problem by tying GI PROMIS to a full narrative GI HPI, offering the reports on a computer interface viewable in the clinic, making the results available both before and during the clinic to patients and their providers, and visualizing the scores with a heat map that displays percentile scores vs. the general US population. Furthermore, we tested a focused use case where clinical benefit should be evident—measuring GI symptoms in patients presenting with disorders affecting the GI tract. Despite these multiple efforts to bolster the potential of GI PROMIS, and our enthusiasm for PROMIS as consortium investigators, we found no difference between groups.

Moreover, despite offering patients access to the e-portal 1 week before their visit, only one-third of patients completed the PROMIS assessments. A likely contributing factor was the "untethered" nature of the e-portal used in this study, as it was not integrated

Table 5. Patient assessment of shared decision-making

SDM-Q-9 item	Control arm (*n*=118)	GI PROMIS arm (*n*=185)	Adjusted *P* value[a]
Disclosure that a decision needs to be made	81.8±22.5	82.5±13.0	0.82
Formulation of equality of partners	79.4±23.7	79.8±14.4	0.81
Equipoise statement	77.8±25.6	79.2±13.4	0.85
Informing on the options' benefits and risks	78.7±25.1	78.2±14.7	0.41
Investigation of patient's understanding and expectations	83.2±23.3	82.9±13.2	0.75
Identification of preferences	76.0±25.5	76.3±15.0	0.98
Negotiation	75.8±25.8	75.7±15.7	0.87
Shared decision	76.7±26.3	76.9±15.8	0.84
Arrangement of follow-up	81.2±24.6	82.1±12.9	0.81
Average SDM-Q-9 score	79.0±22.0	79.3±12.4	0.85

GI, gastrointestinal; PROMIS, Patient Reported Outcome Measurement Information System; SDM-Q-9, 9-item Shared Decision Making Questionnaire.
Data are presented as mean±s.d.
SDM-Q-9 scores are on a 100-point scale. Complete SDM-Q-9 data was missing for 68 patients.
[a]The linear regression model adjusted for patient- (age, gender, race/ethnicity) and provider-level factors (site of care, level of training).

into the EHR. However, we approached non-completers in the clinic itself and offered help to complete GI PROMIS on a clinic computer while waiting for the doctor, yet most still were uninterested. Notably, our low uptake is similar to findings from Wagner *et al.* (27) who tested the feasibility of using PROMIS through a "tethered" e-portal among women receiving gynecologic oncology outpatient care. They found that only 37% of PROMIS assessment requests sent via their EHR portal were completed by patients (27). In addition to system-level issues, patient-level factors may have also contributed to the low intervention uptake. While a systematic review found that patients generally have positive attitudes towards e-portals, issues including security concerns, preconceived beliefs about technology, among others, continue to pose important barriers for widespread e-portal adoption (28).

In our study, we also noted differential uptake of the GI PROMIS intervention by patient characteristics. Namely, we found that African Americans were 56% less likely to complete PROMIS on the e-portal before their visit (adjusted P=0.002). There was also a trend towards lower use of PROMIS among Latinos compared with whites, but this difference did not quite reach statistical significance (adjusted P=0.07). These findings are consistent with a number of past reports that also found racial/ethnic disparities in e-portal use (29–31). Efforts to better understand and to address these disparities are critical, as the increasing prevalence of e-portals and other digital health interventions may continue to widen the healthcare gap between whites and minorities.

There were also different rates of GI PROMIS uptake among the four clinical sites. For example, patients at the University of Michigan GI clinic were eight times more likely to use PROMIS than patients at the other sites. The reason behind this is unclear. It is possible that the University of Michigan GI clinic cared for patients that were more "tech-savvy" and willing to use the e-portal. It is also possible that physicians at this clinic were stronger champions of PROMIS, or have a different bond with their patients than those at other clinics; this could not be directly measured. These differences indicate that cultural differences among clinical settings may influence use of PROs such as PROMIS.

Although GI PROMIS did not appreciably improve patient-centric outcomes, there are other potential benefits to using PROMIS that were not assessed for as part of this study. For instance, we did not measure provider satisfaction; it is possible that clinicians with access to the GI PROMIS reports were more satisfied with the clinic encounter. Similarly, we did not evaluate clinic visit efficiency. Having the PROMIS scores and PROMIS-directed HPI in hand before seeing the patient in the exam room may have allowed clinicians to conduct a more efficient and meaningful clinic visit, and may also have reduced charting and documenting time. These are areas that are the subjects of future research.

Even though our findings are "negative," they are still relevant for the field of PRO science. The results of this study may inform future research and policy on how best to implement GI PROMIS and other PROs in clinical practice. For example, the passage of the Medicare Access and CHIP Reauthorization Act in 2015 provided the Centers for Medicare and Medicaid an opportunity to update the Medicare EHR Incentive Programs, otherwise known as "Meaningful Use." One of the aims of the next Meaningful Use iteration is to reward providers for the outcomes that technology helps them achieve with patients (32). It remains to be seen how these outcomes will be defined as well as the role of PROs, but it will be important for policy makers to recognize that EHR PRO collection alone may be insufficient to improve patient outcomes.

In summary, this is the first multicenter controlled trial evaluating the impact of PROMIS on patient outcomes in clinical practice. We found that use of NIH GI PROMIS did not improve patient satisfaction or assessment of provider interpersonal skills and shared decision-making. These negative findings may help guide investigators and policy makers in optimizing use of PROs in future clinical practice.

ACKNOWLEDGMENTS

We thank and acknowledge Kenya Hunter and Jennifer Serrano for their support in the conduct of the study.

CONFLICT OF INTEREST

Guarantor of the article: Brennan M.R. Spiegel, MD, MSHS.
Specific author contributions: Christopher V. Almario, MD, MSHPM: Planning and conducting the study, collecting and interpreting data, drafting the manuscript, approval of final draft submitted. William D. Chey, MD: Planning and conducting the study, interpreting data, drafting the manuscript, approval of final draft submitted. Dinesh Khanna, MD, MSc: Planning and conducting the study, interpreting data, drafting the manuscript, approval of final draft submitted. Sasan Mosadeghi, MD: Conducting the study, collecting data, drafting the manuscript, approval of final draft

submitted. Shahzad Ahmed, MD: Conducting the study, collecting data, approval of final draft submitted. Elham Afghani, MD, MPH: Conducting the study, collecting data, approval of final draft submitted. Cynthia Whitman, MPH: Planning and conducting the study, collecting data, approval of final draft submitted. Garth Fuller, MS: Interpreting data, drafting the manuscript, approval of final draft submitted. Mark Reid, PhD: Interpreting data, drafting the manuscript, approval of final draft submitted. Roger Bolus, PhD: Planning the study, interpreting data, approval of final draft submitted. Buddy Dennis, PhD: Planning and conducting the study, approval of final draft submitted. Rey Encarnacion, BS: Planning and conducting the study, approval of final draft submitted. Bibiana Martinez, MPH: Planning and conducting the study, approval of final draft submitted. Jennifer Soares, MSPH: Planning and conducting the study, approval of final draft submitted. Rushaba Modi, MD: Planning and conducting the study, approval of final draft submitted. Nikhil Agarwal, MD: Planning and conducting the study, approval of final draft submitted. Aaron Lee, MD: Planning and conducting the study, approval of final draft submitted. Scott Kubomoto, MD: Planning and conducting the study, approval of final draft submitted. Gobind Sharma, MD: Planning and conducting the study, approval of final draft submitted. Sally Bolus, MS: Planning and conducting the study, approval of final draft submitted. Brennan M.R. Spiegel, MD, MSHS: Planning and conducting the study, interpreting data, drafting the manuscript, approval of final draft submitted.

Financial Support: This study was supported by an NIH/NIAMS research grant (U01 AR057936-05). Dr Almario was supported by a Career Development Award from the American College of Gastroenterology. The PROMIS-triggered HPI generator was developed under a separate grant from Ironwood Pharmaceuticals.

Potential competing interests: Drs Chey and Spiegel are principals in My Total Health.

Study Highlights

WHAT IS CURRENT KNOWLEDGE

- ✓ The National Institutes of Health (NIH) Patient Reported Outcomes Measurement Information System (PROMIS) is a toolbox of highly reliable, precise patient-reported outcome measures that cover the breadth and depth of human health and illness.
- ✓ Owing to the extraordinary burden of digestive diseases, the NIH PROMIS consortium developed and validated gastrointestinal (GI)-specific PROMIS measures.
- ✓ It is unclear whether implementing GI PROMIS, let alone any PROMIS measures, can improve patient outcomes vs. usual care.

WHAT IS NEW HERE

- ✓ This is the first multicenter controlled trial evaluating the impact of PROMIS on patient outcomes in clinical practice.
- ✓ One-time use of GI PROMIS did not improve patient satisfaction or assessment of provider interpersonal skills and shared decision-making.

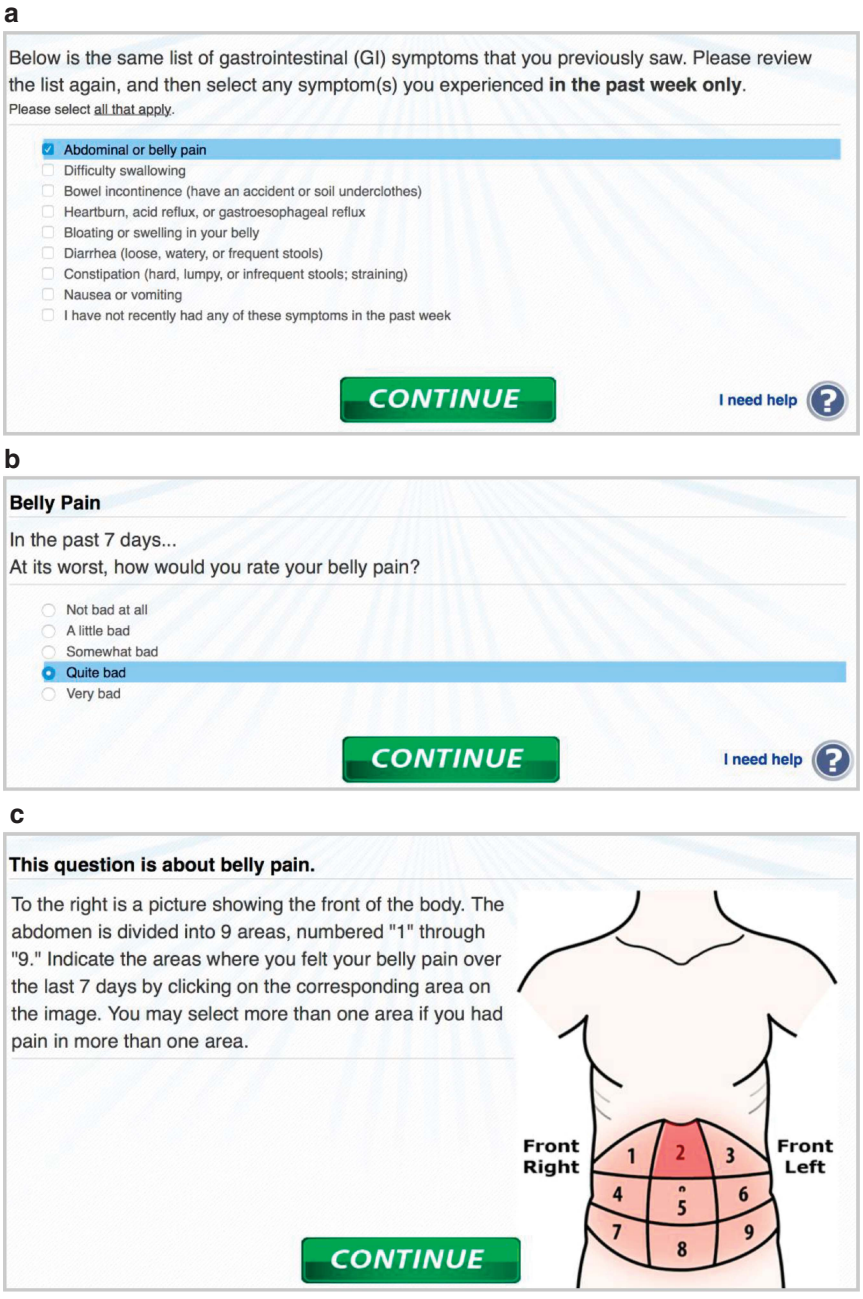
REFERENCES

1. Marshall S, Haywood K, Fitzpatrick R. Impact of patient-reported outcome measures on routine practice: a structured review. J Eval Clin Pract 2006;12:559–68.
2. Valderas JM, Kotzeva A, Espallargues M *et al.* The impact of measuring patient-reported outcomes in clinical practice: a systematic review of the literature. Qual Life Res 2008;17:179–93.
3. 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(Suppl 1):S22–31.
4. Detmar SB, Muller MJ, Schornagel JH *et al.* Role of health-related quality of life in palliative chemotherapy treatment decisions. J Clin Oncol 2002; 20:1056–62.
5. Detmar SB, Muller MJ, Schornagel JH *et al.* Health-related quality-of-life assessments and patient-physician communication: a randomized controlled trial. JAMA 2002;288:3027–34.
6. Neumann M, Edelhauser F, Kreps GL *et al.* Can patient-provider interaction increase the effectiveness of medical treatment or even substitute it?—an exploration on why and how to study the specific effect of the provider. Patient Educ Couns 2010;80:307–14.
7. 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;60:833–43.
8. Greenhalgh J. The applications of PROs in clinical practice: what are they, do they work, and why? Qual Life Res 2009;18:115–23.
9. Lohr KN, Zebrack BJ. Using patient-reported outcomes in clinical practice: challenges and opportunities. Qual Life Res 2009;18:99–107.
10. Gilbert A, Sebag-Montefiore D, Davidson S *et al.* Use of patient-reported outcomes to measure symptoms and health related quality of life in the clinic. Gynecol Oncol 2015;136:429–39.
11. Howell D, Molloy S, Wilkinson K *et al.* Patient-reported outcomes in routine cancer clinical practice: a scoping review of use, impact on health outcomes, and implementation factors. Ann Oncol 2015;26:1846–58.
12. Kroenke K, Monahan PO, Kean J. Pragmatic characteristics of patient-reported outcome measures are important for use in clinical practice. J Clin Epidemiol 2015;68:1085–92.
13. 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(Suppl 1): S3–11.
14. Cook KF, O'Malley KJ, Roddey TS. Dynamic assessment of health outcomes: time to let the CAT out of the bag? Health Serv Res 2005;40(Part 2): 1694–711.
15. Chakravarty EF, Bjorner JB, Fries JF. Improving patient reported outcomes using item response theory and computerized adaptive testing. J Rheumatol 2007;34:1426–31.
16. Spiegel BM, Hays RD, Bolus R *et al.* Development of the NIH Patient-Reported Outcomes Measurement Information System (PROMIS) gastrointestinal symptom scales. Am J Gastroenterol 2014;109:1804–14.
17. Nagaraja V, Hays RD, Khanna PP *et al.* Construct validity of the Patient Reported Outcomes Measurement Information System (PROMIS(R)) gastrointestinal symptom scales in systemic sclerosis. Arthritis Care Res (Hoboken) 2014;66:1725–30.
18. Spiegel BM. Patient-reported outcomes in gastroenterology: clinical and research applications. J Neurogastroenterol Motil 2013;19:137–48.
19. Almario CV, Chey W, Kaung A *et al.* Computer-generated vs. physician-documented history of present illness (HPI): results of a blinded comparison. Am J Gastroenterol 2015;110:170–9.
20. Almario CV, Chey WD, Iriana S *et al.* Computer versus physician identification of gastrointestinal alarm features. Int J Med Inform 2015;84: 1111–7.
21. Dyer N, Sorra JS, Smith SA *et al.* Psychometric properties of the Consumer Assessment of Healthcare Providers and Systems (CAHPS(R)) Clinician and Group Adult Visit Survey. Med Care 2012;50(Suppl):S28–34.
22. Agency for Healthcare Research and Quality. CAHPS Database, 2016. Available at: https://www.cahpsdatabase.ahrq.gov/CAHPSIDB/Public/CG/CG_About.aspx (last accessed 6 January 2016).
23. Greco M, Cavanagh M, Brownlea A *et al.* Validation studies of the Doctors' Interpersonal Skills Questionnaire. Educ Gen Pract 1999;10:256–64.
24. Kriston L, Scholl I, Holzel L *et al.* The 9-item Shared Decision Making Questionnaire (SDM-Q-9). Development and psychometric properties in a primary care sample. Patient Educ Couns 2010;80:94–9.

25. Cohen J. A power primer. Psychol Bull 1992;112:155.
26. Norman GR, Sloan JA, Wyrwich KW. Interpretation of changes in health-related quality of life: the remarkable universality of half a standard deviation. Med Care 2003;41:582–92.
27. Wagner LI, Schink J, Bass M *et al.* Bringing PROMIS to practice: brief and precise symptom screening in ambulatory cancer care. Cancer 2015;121:927–34.
28. Goldzweig CL, Orshansky G, Paige NM *et al.* Electronic patient portals: evidence on health outcomes, satisfaction, efficiency, and attitudes: a systematic review. Ann Intern Med 2013;159:677–87.
29. Goel MS, Brown TL, Williams A *et al.* Disparities in enrollment and use of an electronic patient portal. J Gen Intern Med 2011;26:1112–6.
30. Roblin DW, Houston TKII, Allison JJ *et al.* Disparities in use of a personal health record in a managed care organization. J Am Med Inform Assoc 2009;16:683–9.
31. Sarkar U, Karter AJ, Liu JY *et al.* Social disparities in internet patient portal use in diabetes: evidence that the digital divide extends beyond access. J Am Med Inform Assoc 2011;18:318–21.
32. Slavitt A, DeSalvo K. EHR Incentive Programs: Where We Go Next, 2016. Available at: https://blog.cms.gov/2016/01/19/ehr-incentive-programs-where-we-go-next/ (last accessed 5 February 2016).

APPENDIX FIGURE 1

Sample screenshots from the patient–provider e-portal. The screen in (**a**) asks patients to select among the eight gastrointestinal PROMIS symptoms that they have experienced in the past week. Screens in (**b**) and (**c**) show sample questions included in the abdominal pain PROMIS survey.



APPENDIX TABLE 1

COMPARISON OF POSTVISIT OUTCOME QUESTIONNAIRES COMPLETERS VS. NON-COMPLETERS IN THE CONTROL ARM

Variable	Non-completers (*n*=345)	Completers (*n*=154)	*P* value
Age (years)	52.8±17.0	58.7±15.8	<0.001
Male	216 (62.6%)	67 (43.5%)	<0.001
Race/ethnicity			0.24
Caucasian	198 (57.4%)	100 (64.9%)	
African American	74 (21.5%)	32 (20.8%)	
Asian	12 (3.5%)	5 (3.3%)	
Latino	43 (12.5%)	9 (5.8%)	
Other/unknown	18 (5.2%)	8 (5.2%)	
Site of care			<0.001
West Los Angeles VA GI clinic	188 (54.5%)	55 (35.7%)	
Cedars-Sinai GI clinic	15 (4.4%)	13 (8.4%)	
University of Michigan GI clinic	100 (29.0%)	77 (50.0%)	
University of Michigan scleroderma clinic	42 (12.2%)	9 (5.8%)	

GI, gastrointestinal; VA, Veterans Affairs.
Data are presented as mean±s.d. or *n* (%).

APPENDIX TABLE 2

COMPARISON OF POSTVISIT OUTCOME QUESTIONNAIRES COMPLETERS VS. NON-COMPLETERS IN THE GI PROMIS INTERVENTION ARM

Variable	Non-completers (*n*=105)	Completers (*n*=112)	*P* value
Age (years)	51.6±16.1	56.5±16.3	0.03
Male	50 (47.6%)	59 (52.7%)	0.46
Race/ethnicity			0.26
Caucasian	77 (73.3%)	74 (66.1%)	
African American	11 (10.5%)	17 (15.2%)	
Asian	6 (5.7%)	2 (1.8%)	
Latino	4 (3.8%)	7 (6.3%)	
Other/unknown	7 (6.7%)	12 (10.7%)	
Site of care			0.002
West Los Angeles VA GI clinic	37 (35.2%)	40 (35.7%)	
Cedars-Sinai GI clinic	0 (0%)	11 (9.8%)	
University of Michigan GI clinic	48 (45.7%)	52 (46.4%)	
University of Michigan scleroderma clinic	20 (19.1%)	9 (8.0%)	

GI, gastrointestinal; PROMIS, Patient Reported Outcome Measurement Information System; VA, Veterans Affairs.
Data are presented as mean±s.d. or *n* (%).

APPENDIX TABLE 3

PATIENT ASSESSMENT OF PROVIDER INTERPERSONAL SKILLS IN PER-PROTOCOL ANALYSIS

DISQ item	Control arm (*n*=144)	GI PROMIS arm (*n*=104)	Adjusted *P* value[a]
Overall satisfaction with provider	89.7±17.7	88.3±19.0	0.54
Warmth of provider's greeting	90.3±16.4	88.8±17.4	0.54
Ability to listen to the patient	90.4±17.2	88.7±18.0	0.38
Adequacy of explanations to patient	89.4±16.7	88.3±18.0	0.54
Extent of reassurance provided to patient	87.8±18.9	87.5±19.7	0.99
Confidence in provider's ability	90.6±17.5	90.0±18.7	0.80
Opportunity for patient to express concerns and fears	89.7±17.7	88.3±19.2	0.59
Respect shown to patient	91.3±16.3	93.8±14.0	0.18
Time given for visit	88.5±18.9	88.7±18.9	0.90
Consideration of patient's personal situation in treatment or advice	89.7±18.5	88.1±19.0	0.55
Concern for patient as a person	89.9±16.9	89.0±18.0	0.67
Recommendation of provider to friends	90.3±18.5	88.7±19.1	0.54
Average DISQ score	89.8±16.0	89.0±16.7	0.73

DISQ, Doctors' Interpersonal Skills Questionnaire; GI, gastrointestinal; PROMIS, Patient Reported Outcome Measurement Information System.
Data are presented as mean±s.d.
DISQ scores are on a 100-point scale. Complete DISQ data was unavailable for 18 patients.
[a]The linear regression model adjusted for patient- (age, gender, race/ethnicity) and provider-level factors (site of care, level of training).

APPENDIX TABLE 4

PATIENT ASSESSMENT OF SHARED DECISION-MAKING IN PER-PROTOCOL ANALYSIS

SDM-Q-9 item	Control arm (*n*=118)	GI PROMIS arm (*n*=80)	Adjusted *P* value[a]
Disclosure that a decision needs to be made	81.8±22.5	82.5±19.8	0.91
Formulation of equality of partners	79.4±23.7	79.4±21.9	0.73
Equipoise statement	77.8±25.6	81.5±20.2	0.38
Informing on the options' benefits and risks	78.7±25.1	77.5±22.4	0.50
Investigation of patient's understanding and expectations	83.2±23.3	82.7±20.1	0.83
Identification of preferences	76.0±25.5	76.5±22.9	0.99
Negotiation	75.8±25.7	75.6±24.0	0.86
Shared decision	76.7±26.3	77.1±24.1	0.95
Arrangement of follow-up	81.2±24.6	83.5±19.6	0.45
Average SDM-Q-9 score	79.0±22.0	79.6±18.9	1.0

GI, gastrointestinal; PROMIS, Patient Reported Outcome Measurement Information System; SDM-Q-9, 9-item Shared Decision Making Questionnaire.
Data are presented as mean±s.d.
SDM-Q-9 scores are on a 100-point scale. Complete SDM-Q-9 data were missing for 68 patients.
[a]The linear regression model adjusted for patient- (age, gender, race/ethnicity) and provider-level factors (site of care, level of training).

see related article on page 1546

Using Patient-Reported Outcome Measures in Gastroenterology: PROMISed Land or Road to Nowhere?

David J. Gracie, BSc, MBChB, MRCP[1,2] and Alexander C. Ford, MBChB, MD, FRCP[1,2]

Abstract: Incorporating patient-reported outcomes (PROs) into clinical practice is advocated by some. However, the benefits remain uncertain. Almario *et al.* examined the impact of a gastrointestinal (GI) version of the patient-reported outcomes measurement information system (PROMIS) on patient satisfaction, perception of doctors' interpersonal skills, and the likelihood of shared decision-making. Patients were allocated to complete GI PROMIS prior to their outpatient appointment, or usual management. Overall, uptake of GI PROMIS was poor and there was no difference in any outcome measure between those completing the questionnaire and those receiving usual management, suggesting PROs may be of limited utility in this setting.

Am J Gastroenterol 2016; 111:1557–1558; doi:10.1038/ajg.2016.415

A patient-reported outcome (PRO) is "any report that comes directly from a patient about a health condition or its treatment, without interpretation of the patient's response by a clinician or anyone else" (1). PROs comprise individual symptoms, or clusters of symptoms, that aim to capture the patient's illness experience, and may help health-care providers to better understand symptoms from the patient's perspective (2). Their use is advocated as an aid to the assessment of disease activity in conditions such as inflammatory bowel disease (IBD) and irritable bowel syndrome (IBS), but also as an adjunct to standard clinical practice, in order to improve doctor–patient relations and patient satisfaction. Furthermore, the US Food and Drug Administration now support the use of PROs as end points in clinical trials of novel pharmacological therapies in gastrointestinal (GI) diseases (3).

The National institutes of Health created the patient-reported outcomes measurement information system (PROMIS) in 2004 (ref. 4), which was modified for use in GI disease in 2014 (ref. 5). This data collection tool comprises short questionnaires that are easily administered within the constraints of routine clinical practice, with the aim of providing a standardized set of PROs that can be used across the entirety of GI clinical and research practice. However, although the concept of integrating PROs into standard care is over a decade old, their impact on clinical outcomes remains poorly studied.

In this issue of *American Journal of Gastroenterology*, Almario *et al.* (6) report data on the effect of the use of PROs on patient satisfaction, as well as the assessment of care provider interpersonal skills and shared decision-making at an initial visit in GI outpatient clinics. The study was a pragmatic clinical trial, where patients in the intervention group received a letter inviting them to log on to an e-portal in order to complete a GI version of PROMIS (GI PROMIS), whereas those in the control group received usual management. Based on GI PROMIS responses, a symptom "heat map" and automated history of presenting illness was generated, which was available for the attending physician and patient to review prior to, and during, the consultation. Following clinic review, participants were asked to complete post-visit questionnaires evaluating their satisfaction, their care provider's interpersonal skills, and their perceptions about ability to engage in shared decision-making.

In total, 594 patients were assigned to the intervention group, of whom only 221 (37.2%) accessed the e-portal and completed the GI PROMIS questionnaire. African Americans were significantly less likely to engage than White Caucasians (18.4% vs. 48.6%, odds ratio (OR)=0.44; 95% confidence interval (CI) 0.26–0.74), and there was also a trend toward lower engagement in Latinos (19.7%, OR=0.52; 95% CI 0.26–1.05). Only 112 (18.9%) of those in the intervention arm completed the GI PROMIS questionnaire, attended their clinic appointment, and completed the outcome questionnaire following their visit. In the control group,

[1]Leeds Gastroenterology Institute, St. James's University Hospital, Leeds, UK; [2]Leeds Institute of Biomedical and Clinical Sciences, University of Leeds, Leeds, UK. **Correspondence:** Alexander C. Ford, MBChB, MD, FRCP, Leeds Gastroenterology Institute, St. James's University Hospital, Room 125, 4th floor, Bexley Wing, Beckett street, Leeds LS9 7TF, UK. E-mail: alexf12399@yahoo.com
Received 12 July 2016; accepted 28 July 2016

502 patients were invited to participate. Of these, three (0.6%) had incomplete demographic data and were, therefore, excluded. In total, only 154 (30.9%) of the remaining 499 patients completed their post-visit questionnaires. Patients who were assigned to the intervention arm and who completed the GI PROMIS questionnaire were younger than the 154 in the control group who provided post-visit data (54.1 vs. 58.7 years, P=0.007).

Comparison of outcome questionnaire responses between the intervention and control groups was undertaken using both an intention-to-treat analysis, where GI PROMIS responders who did not complete the outcome questionnaires were assumed to have the same responses as the control arm, and a per-protocol analysis. Overall, in both intention-to-treat and per-protocol analyses, the use of GI PROMIS had no effect on patient satisfaction, their opinions on the care provider's interpersonal skills, or their perception of the ability to engage in shared decision-making, when compared with control patients receiving usual management (P≥0.05 for all comparisons).

Strengths of this study include it being conducted as part of usual clinical practice, meaning that the results are generalizable to other tertiary-care referral populations, and the use of validated questionnaires to assess the end points of interest. However, there are also several limitations. The poor response rate to the invitation to complete the GI PROMIS questionnaire is a major issue as, if patients are not willing to engage with this system, then any potential benefits from its use will be reduced. Of those invited, <40% completed GI PROMIS, with the odds of African American invitees completing the questionnaire significantly lower than those of White Caucasians. Furthermore, GI PROMIS responders were significantly younger than controls, suggesting either that there was selection bias during recruitment into the intervention group, or that responders to GI PROMIS were younger than non-responders, thereby skewing the demographic comparison between the two groups, and potentially limiting the implementation of this intervention to young, White, "tech-savvy" individuals.

Based on these results alone, the use of PROs does not appear to improve patient satisfaction, perception of doctors' interpersonal skills, or the likelihood of shared decision-making between clinician and patient. However, the poor response rate to outcome questionnaires (51.6% in a self-selected intervention group and 30.7% in the control group) limits the validity of these findings. Furthermore, it may be that the standard of communication skills and shared decision-making in the four tertiary-care referral centers used in this study was already extremely high, thereby reducing the likelihood of identifying a benefit of the PROMIS intervention. It would be premature, therefore, to assume that there are no benefits from applying this intervention in primary or secondary care.

Aside from the outcomes incorporated in the present study, the use of PROs may be of benefit in circumstances other than those examined by the authors. Although PROMIS was designed to aid improvement in patient outcomes, the impact of using it on other end points could be studied. A reduction in mean consultation length in those patients who had completed GI PROMIS could improve efficiency and reduce costs in busy outpatient departments, but this was not examined. Moreover, the use of PROs in patients with chronic GI disorders may be of greater interest, as longitudinal alterations in PROs may prove a useful surrogate measure of the natural history of symptoms in chronic GI diseases such as IBS and IBD, as has been highlighted in other specialties (7). Furthermore, incorporating PROs into long-term care pathways may aid the development of doctor–patient relationships over several consultations thereby improving the patient satisfaction, perception of doctors' interpersonal skills, and the likelihood of shared decision-making between clinician and patient.

In summary, this study demonstrates no beneficial effect of GI PROMIS on patient satisfaction, care provider interpersonal skills, or shared decision-making in gastroenterology outpatient clinics in tertiary care. However, inherent limitations of the study design and methodology employed, and the population participating, restrict the applicability of the findings. It may be that engagement with GI PROMIS would be greater in patients with chronic GI disorders, with beneficial effects arising as a consequence. The use of such tools in the longitudinal assessment of conditions, including IBS and IBD, is also of interest, although the sensitivity and specificity of individual PROs used in the formation of them, to date, has been shown to be poor in predicting disease activity (8–10).

CONFLICT OF INTEREST

Guarantor of the article: Alexander C. Ford, MBChB, MD, FRCP.
Specific author contributions: D.J.G. and A.C.F. conceived and drafted the manuscript. D.J.G. and A.C.F. have approved the final draft of the manuscript.
Potential competing interests: None.
Financial support: None.

REFERENCES

1. Burke LB, Kennedy DL, Miskala PH *et al.* The use of patient-reported outcome measures in the evaluation of medical products for regulatory approval. Clin Pharmacol Ther 2008;84:281–3.
2. Spiegel BM. Patient-reported outcomes in gastroenterology: clinical and research applications. J Neurogastroenterol Motil 2013;19:137–48.
3. Williet N, Sandborn WJ, Peyrin-Biroulet L. Patient-reported outcomes as primary end points in clinical trials of inflammatory bowel disease. Clin Gastroenterol Hepatol 2014;12:1246–56.e6.
4. 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:S3–11.
5. Spiegel BM, Hays RD, Bolus R *et al.* Development of the NIH Patient-Reported Outcomes Measurement Information System (PROMIS) gastrointestinal symptom scales. Am J Gastroenterol 2014;109:1804–14.
6. Almario CV, Chey WD, Khanna D *et al.* Impact of National Institutes of Health Gastrointestinal PROMIS® Measures in Clinical Practice: results of a Multicenter Controlled Trial. Am J Gastroenterol 2016;111:1546–56 (this issue).
7. Li Z, Thompson LA, Gross HE *et al.* Longitudinal associations among asthma control, sleep problems, and health-related quality of life in children with asthma: a report from the PROMIS((R)) Pediatric Asthma Study. Sleep Med 2016;20:41–50.
8. Gracie DJ, Williams CJ, Sood R *et al.* Poor correlation between clinical disease activity and mucosal inflammation, and the role of psychological comorbidity, in inflammatory bowel disease. Am J Gastroenterol 2016;111:541–51.
9. Targownik LE, Sexton KA, Bernstein MT *et al.* The relationship among perceived stress, symptoms, and inflammation in persons with inflammatory bowel disease. Am J Gastroenterol 2015;110:1001–12.
10. Jharap B, Sandborn WJ, Reinisch W *et al.* Randomised clinical study: discrepancies between patient-reported outcomes and endoscopic appearance in moderate to severe ulcerative colitis. Aliment Pharmacol Ther 2015;42:1082–92.