

A pilot study of patient-centered outcome assessment using PROMIS for patients undergoing colorectal surgery

Traci L. Hedrick¹ · Amy M. Harrigan¹ · Robert H. Thiele² · Charles M. Friel¹ · Benjamin D. Kozower³ · George J. Stukenborg⁴

Received: 10 August 2016 / Accepted: 17 April 2017 / Published online: 24 April 2017
© The Author(s) 2017. This article is an open access publication

Abstract

Purpose Few studies have assessed patient-reported outcomes following colorectal surgery. The absence of this information makes it difficult to inform patients about the near-term effects of surgery, beyond outcomes assessed by traditional clinical measures. This study was designed to provide information about the effects of colorectal surgery on physical, mental, and social well-being outcomes.

Methods The NIH Patient-Reported Outcomes Measurement Information System (PROMIS®) Assessment Center was used to collect patient responses prior to surgery and at their routine postoperative visit. Four domains were selected based on patient consultation and clinical experience: depression, pain interference, ability to participate in social roles and activities, and interest in sexual activity. Multilevel random coefficient models were used to assess the change in scores

during the follow-up period and to assess the statistical significance of differences in trends over time associated with key clinical measures.

Results In total, 142 patients were consented, with 107 patients completing pre- and postoperative assessments (75%). Preoperative assessments were typically completed 1 month prior to surgery (mean 29.5 days before, SD = 19.7) and postoperative assessments 1 month after surgery (mean 30.7 days after, SD = 9.2), with a mean of 60.3 days between assessment dates. Patients demonstrated no statistically significant changes in scores for pain interference (−0.18 points, $p = 0.80$) or the ability to participate in social roles and activities (0.44 points, $p = 0.55$), but had significant decreases in depression scores between pre- and postoperative assessments (−1.6 points, $p = 0.03$) and near significant increases in scores for interest in sex (1.5 points, $p = 0.06$). Pain interference scores for

Results from this research have been presented in part at the PROMIS Health Organization Inaugural Conference, on May 15, 2015, in Philadelphia PA, USA, and at the American Society of Colon and Rectal Surgeons Annual Scientific Meeting on May 1, 2016 in Los Angeles, CA, USA.

Electronic supplementary material The online version of this article (doi:10.1007/s00520-017-3718-4) contains supplementary material, which is available to authorized users.

✉ Traci L. Hedrick
TH8Q@hscmail.mcc.virginia.edu

Amy M. Harrigan
ACM6A@hscmail.mcc.virginia.edu

Robert H. Thiele
RHT7W@virginia.edu

Charles M. Friel
CMF2X@hscmail.mcc.virginia.edu

Benjamin D. Kozower
kozowerb@wudosis.wustl.edu

George J. Stukenborg
gstukenborg@virginia.edu

¹ Department of Surgery, University of Virginia Health System, Charlottesville, VA, USA

² Department of Anesthesiology, University of Virginia Health System, Charlottesville, VA, USA

³ Division of Cardiothoracic Surgery, Department of Surgery, Washington University, St. Louis, MO, USA

⁴ School of Medicine, University of Virginia Health System, P.O. Box 800717, Charlottesville, VA 22908-0821, USA

patients with neoadjuvant chemotherapy significantly increased (3.5 points, $p = 0.03$). Scores for the interest in sex domain decreased (worsened) for patients with oncologic etiology (-3.7 points, $p = 0.03$). No other differences in score trends by patient characteristics were large enough to be statistically significant at the $p < 0.05$ threshold.

Conclusion These data suggest that the majority of patients quickly return to baseline physical, mental, and social function following colorectal surgery. This information can be used preoperatively to counsel patients about the typical impact of colorectal surgery on quality of life.

Keywords Patient reported outcomes · PROMIS · Colorectal surgery · Enhanced recovery after surgery

Introduction

Mortality, morbidity, and length of stay are key clinical outcomes for patients following colorectal surgery. Outcomes that reflect the patients' subjective perspective, including depression, pain interference, and an individual's ability to function socially, are also of great significance, though these outcomes are less frequently investigated [1].

These other physical, mental, and social function domains of health are examples of patient-reported outcomes (PROs), which taken together comprise the overall quality of life obtained following treatment [2]. The collection and evaluation of PROs are of growing interest for the assessment of the overall quality of life obtained by patients following treatment and also for guiding patient decision-making and symptom management [3]. Implementation of methods for assessing PROs in surgical care has advanced most quickly for total joint replacement patients [4, 5]. The Centers for Medicare and Medicaid (CMS) recently announced plans to collect preoperative and postoperative PROs for patients with lower extremity joint replacement procedures, as a component of payment reforms for Medicare beneficiaries [6]. New initiatives that will require the collection of PROs for patients with other commonly performed surgical procedures are anticipated, primarily through the modification of existing surgery patient registries [7].

A more complete understanding of patient outcomes experienced following colorectal surgery requires the assessment of PROs. There are several different types of measurement tools for evaluating patient-centered outcomes. A systematic review of the PROs that have been used in colorectal cancer surgical studies indicates that there is significant heterogeneity in methods for measuring psychological, symptomatic, and functional domains [8].

Standardized tools for the measurement of physical, mental, and social function domains are available from the Patient-Reported Outcomes Measurement Information System

(PROMIS) [9, 10]. The development of the PROMIS assessments was funded through the National Institutes of Health (NIH) Common Fund Initiative, with the purpose of providing standardized resources for use in clinical research and practice. PROMIS provides validated tools for the measurement of physical, mental, and social function domains that can be customized to capture selected domains of primary concern for specific types of patients.

Few studies have assessed these types of measures in colorectal surgery patients, and the absence of this information makes it difficult to inform patients about the near-term effects of surgery, beyond outcomes assessed by traditional clinical measures. This study was designed to provide information about the effects of colorectal surgery on physical, mental, and social well-being outcomes. We hypothesized that the PROMIS system could be utilized successfully to capture PROs in this patient population.

Methods

Study population

A 22-month (July 10, 2013 to April 30, 2015) prospective observational study was undertaken at the University of Virginia under informed consent in keeping with local Institutional Review Board standards (IRB #16097). Adult patients (18 years of age and older) undergoing elective colorectal procedures within an enhanced recovery after surgery protocol by one of two colorectal surgeons were consented and prospectively followed by a clinical research coordinator (CRC) [11]. Patients with emergent procedures, patients less than 18 years of age, prisoners, pregnant women, and patients unable to provide consent were excluded.

PRO assessment

The PROMIS Assessment Center software was used to collect and score PROs for patients before and after surgery [12]. The assessment center software is internet-based, with the collection and secure storage of PROMIS measures and other content implemented on a study-specific website. A customized assessment instrument was developed to capture patient reported information addressing four domains: depression, pain interference, interest in sex, and ability to participate in social roles and activities. These four domains were selected for use in the study based on discussions with patients and physicians, who provided examples of psychological, symptom, and functional domains of primary concern. The assessment was limited to the collection of information for these four domains in order to minimize

the total time required to complete the assessment during the patient's clinic visit to 10 min or less.

Each domain was assessed using Computer-Adaptive Testing (CAT) algorithms that present a series of health domain questions tailored to each patient, with subsequent questions selected based on a respondent's answers to prior questions [13]. The CAT-based assessments yield more precise measurement than assessments with a fixed series of questions, while requiring fewer question items than fixed length question series, and also provide improved precision at the extremes of the domain continuum [14].

The selected PROMIS domains were scored using question items with five ordinal response options (1 to 5) for each question in the domain [15]. Total scores were calculated for each domain, for each patient assessment. PROMIS domain scores are standardized to a mean of 50 and standard deviation of 10, reflecting the distribution of scores reported for a large sample representative of the 2000 US General Census [16]. For example, a score of 40 represents a score that is one standard deviation below the reference population mean score.

Patients were identified in the outpatient clinic and approached for informed consent after the decision was made to proceed with surgery. Consenting patients accessed the assessment during clinic visits using tablet computers with wireless access to the PROMIS website [17]. Patients completed an assessment preoperatively as part of their routine preoperative workup appointment. The postoperative assessment was obtained as part of the routine postoperative follow-up appointment, which is typically scheduled for three or more weeks after discharge from the hospital following surgery.

Clinical research coordinators provided brief instruction on the use of the tablet computer to all patients and provided direct assistance with the assessment process to patients upon request. The availability of direct assistance was implemented as part of the protocol. Prior research demonstrated that while most patients would be able to complete the assessment without assistance, some patients would require direct assistance with the use of the tablet computer, with the assessment software, or with the assessment questions [17].

Statistical analysis

Longitudinal analysis of changes in PROs during the perioperative period was conducted using the multilevel random coefficient regression model (Fig. 1) [18–20]. Since measurement of the four domains is conducted both before and after surgery for each subject, the data are hierarchically structured with responses correlated among patients. Patient assessment dates also varied with regard to the lengths of time before and after surgery. The multilevel random coefficient model approach accommodates both the within subject correlation

and the variation in time of assessment. Patient PRO scores are estimated as a function of time using separate intercepts and separate slopes for each patient, and the model can also include adjustments for differences in functions attributable to patient level covariates. The multilevel random coefficient model yields fixed effect estimates of the overall mean PRO score at selected time points, the difference in mean PRO score at selected time points by patient level characteristic, the change in score over time from preoperative to postoperative assessment, and the difference in the change over time by patient level characteristic.

The estimated mean scores and differences in mean scores associated with patient-level characteristics (fixed effects) were estimated with reference to the date of surgery, which is a shared longitudinal reference point for all patients. Multilevel random coefficient models were estimated to assess the overall change during the perioperative period for each of the four PRO domains. Separate multilevel random coefficient models were estimated to assess the significance of differences in these changes over time associated with patient demographics, preoperative characteristics, and postoperative outcomes. The statistical significance of the estimated fixed effects was assessed by the *F* test statistic for type 3 tests, using the a priori selected threshold value of $p < 0.05$. Figure 1 presents an example formulation of the multilevel random coefficient model.

The statistical models measure the change from pre- to post-assessment as constant over time, beginning at the time of the preoperative assessment. This aspect of the models has the potential to diminish the impact of surgery on the magnitude of the observed change in symptoms assessed at follow-up. We assessed the sensitivity of the model results to this assumption by recalculating results for each model, assuming that the preoperative assessment occurred on the date of surgery. The original analysis results were compared to the simulated results to determine whether any of the original model results were meaningfully different, assuming that the changes occurred between the date of surgery and the date of the postoperative assessment.

Sample size requirements for the study were estimated for a repeated measures ANOVA analysis of hypothesized levels of change in PROMIS scores during the perioperative period associated with a between-subject factor [21]. The Hotelling-Lawley *F* test was used to estimate sample sizes required to assess a hypothesized factor effect of a 10 unit difference in post-surgery scores from a median score of 50 prior to surgery, at $\alpha = 0.05$ and power = 0.80, with correlation of residuals = 0.5, for standard deviations of post-surgery scores of 10 or 15 units. Sample size calculations demonstrated that from 34 to 74, patients were needed to achieve power = 0.80 for a post-surgery score difference of 10 units associated with a given factor, at standard deviations from 10 units to 15 units, respectively.

Fig. 1 Example level 1 and level 2 components of the combined equation represented in the multilevel random coefficient model

$$\begin{aligned} \text{Level 1 (Day): PRO Score} &= \beta_{0j} + \beta_{1j}(\text{Day}) + r_{ij} \\ \text{Level 2 (Patient): } \beta_{0j} &= \gamma_{00} + \gamma_{01}(\text{Sex}) + u_{0j} \\ \text{Level 2 (Patient): } \beta_{1j} &= \gamma_{10} + \gamma_{11}(\text{Sex}) + u_{1j} \end{aligned}$$

$$\text{Combined: } \gamma_{00} + \gamma_{01}(\text{Sex}) + \gamma_{10}(\text{Day}) + \gamma_{11}(\text{Sex} * \text{Day}) + [u_{0j} + u_{1j}(\text{Day}) + r_{ij}]$$

Fixed effect components

γ_{00} mean PRO score at time of surgery (Day = 0) for reference value of patient covariate (Sex)
 γ_{01} change in PRO score (per Day) for reference value of patient covariate (Sex)
 γ_{10} difference in mean PRO score (Day = 0) for specified value of patient covariate (Sex)
 γ_{11} difference in change in PRO score (per Day) for specified value of patient covariate (Sex)

Random effect: $[u_{0j} + u_{1j}(\text{Day}) + r_{ij}]$

Results

In total, 142 patients were consented during the study period. Of these, 9 patients failed to complete the preoperative assessment, and 26 patients failed to complete the postoperative assessment. The primary reasons for failure to complete assessments were missed postoperative visits due to readmissions or other complications, cancelled surgeries, and technical difficulties with the availability of wireless access to the online assessment. There was one immediate postoperative death: An elderly patient with bilateral stomas committed suicide on postoperative day 5 within 24 h of discharge. Both preoperative and postoperative assessments were available for 107 patients, or 75% of the total number of colorectal surgery patients originally consented. All patients with both preoperative and postoperative assessments were included in the final study population. Figure 2 illustrates the flow of consented patients who completed or did not complete the pre-procedure and post-procedure assessments.

Differences in the distribution of demographic, preoperative, and postoperative characteristics between the subgroup of patients who did not complete the procedure assessment and the final study population were assessed for statistical significance using Fischer’s exact test, with significance defined at the $p < 0.05$ threshold. Patients who completed only the pre-procedure assessment were significantly different with regard to race ($p = 0.0067$), but were not significantly different with regard to the proportion male ($p = 0.6662$) or payor status

($p = 0.1179$). Surgery was cancelled following completion of the preoperative assessment for seven of the 26 patients without a post-procedure assessment.

Among the 19 patients with surgery and no post-procedure assessment, there were no statistically significant differences from the final study population in frequency of preoperative stoma, oncologic etiology, inflammatory bowel disease, diverticulitis, or receipt of neoadjuvant chemotherapy or radiation. There also were no significant differences in the frequency of surgical site infection, return to operating room, or postoperative stoma. Patients with surgery and no post-procedure assessment however had significantly lower frequencies of laparoscopic surgery (26 vs. 53%, $p = 0.0449$), were more likely to be readmitted following hospital discharge (38 vs. 5%, $p < 0.0001$), and were more likely to incur any morbidity within 30 days of surgery (37 vs. 11%, $p = 0.0096$). Complete results for the comparison of patients with and without post-procedure assessments are provided in a supplemental table available online.

Table 1 lists the demographic, preoperative clinical characteristics, and postoperative outcomes measured for the final study population of 107 patients with both pre- and post-procedure assessments. The majority of patients reported their race as “White” (85%), and females accounted for slightly more than one half (55%) of the study population. Preoperative stomas were reported in 15%, inflammatory bowel disease in 19%, and diverticulitis in 21%. Oncologic cases represented 43%, with neoadjuvant chemotherapy

Fig. 2 Illustrates the flow of pre- and post-procedure assessments collected for consented patients and details the number of patients who completed assessments and who did not complete assessments at each stage

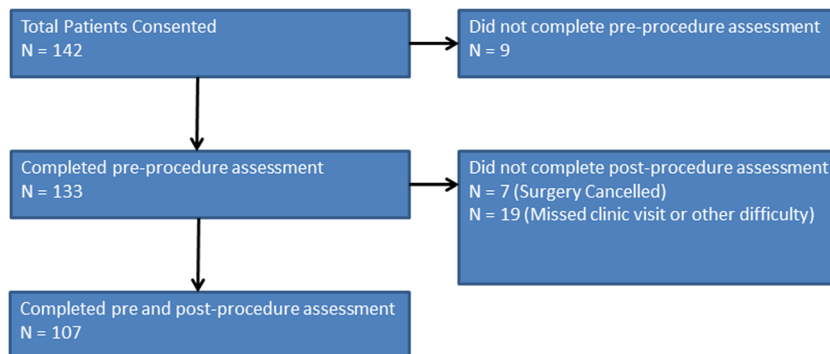


Table 1 Study population characteristics

	Frequency	Percentage
Total number of patients	107	100.00
Demographics		
Sex: male	48	44.86
Race: Asian	1	0.97
Race: African-American	10	9.34
Race: White	91	85.05
Race: not reported	5	4.67
Insurance: Medicaid and indigent	14	13.08
Insurance: Medicare	40	37.38
Insurance: private	53	49.53
Preoperative characteristics		
Preoperative stoma: colostomy	11	10.28
Preoperative stoma: ileostomy	5	4.67
Preoperative stoma: none	91	85.05
Oncologic: yes	46	42.99
Inflammatory bowel disease: yes	20	18.69
Diverticulitis: yes	23	21.50
Neoadjuvant chemotherapy: yes	27	25.23
Neoadjuvant radiation: yes	25	23.36
Surgery		
Laparoscopic: yes	57	53.27
Postoperative outcomes		
Surgical site infection: superficial	5	4.67
Surgical site infection: deep	1	0.93
Surgical site infection: none	101	94.39
Returned to operating room: yes	3	2.80
Postoperative stoma: colostomy	11	10.28
Postoperative stoma: ileostomy	30	28.04
Postoperative stoma: none	66	61.68
Readmitted: yes	5	4.67
Any morbidity with 30 days: yes	12	11.21

reported for 25%, and neoadjuvant radiation for 23%. Postoperative stomas were reported for 38%. Postoperative complications included superficial or deep surgical site infections for 5%, return to operating room for 3%, readmission for 5%, and the occurrence of any morbidity within 30 days for 11%.

Table 2 lists the distributional characteristics of the PRO domain scores, the assessment completion times, the number of domain items completed, and the number of days elapsed between assessments. The mean time required to complete the assessment was 4.7 min. Patients on average completed the preoperative assessment 1 month prior to surgery (mean 29.5 days before, SD = 19.7) and completed the postoperative assessment 1 month after surgery (mean 30.7 days after, SD = 9.2). The mean duration between the preoperative and postoperative assessment dates was 60.3 days.

Overall, the study population means for each PRO domain score were near the US general population calibrated reference means of 50, with preoperative mean scores for depression (49.9), interest in sex (47.4), pain interference (56.0), and ability to participate in social roles and activities (47.8) and with postoperative mean scores for depression (48.2), interest in sex (49.8), pain interference (56.0), and ability to participate in social roles and activities (49.4). However, the means of the preoperative pain interference scores and postoperative pain interference scores were both six points higher than the reference US general population standard.

Prior research comparing PROMIS domain scores to clinically relevant symptom severity levels demonstrates that increments of one standard deviation can be used to define reliable threshold values for grouping patients by symptom severity, such as 50–59 (mild), 60–69 (moderate), and ≥ 70 (severe) for pain and 55–64 (mild), 65–74 (moderate), and ≥ 75 (severe) for depression [22]. Comparison of patients grouped at these threshold values demonstrates a large improvement in the distribution of depression scores following surgery, with 12.8% of patients reporting moderate-to-severe depression at the preoperative assessment, and 4.0% reporting moderate-to-severe depression at the postoperative assessment. None of the other three domains demonstrated this large shift in score distributions. For example, the overall distribution of pain interference scores changed only slightly following surgery, with 36.5% of patients reporting moderate-to-severe pain interference at the preoperative assessment, and 33.6% reporting moderate-to-severe depression at the postoperative assessment. The proportion of patient PRO scores for each domain grouped within the ranges of 5 unit increments is provided in a supplemental table. Median, mean, and standard deviation values for each of the four PRO domains, for both the preoperative and postoperative periods, are provided for each of the demographic characteristics, preoperative clinical characteristics, and postoperative outcomes in a supplemental table available online.

The trend of change in scores during the perioperative period was assessed using the fixed effect estimates obtained by the multilevel random coefficient models. Models were estimated to assess the change in scores for each of the four PRO domains, both overall and for each demographic, preoperative clinical characteristic, and postoperative outcome. Table 3 provides a summary of the model estimated postoperative change in the mean scores, overall and for each group characteristic, along with the estimated statistical significance of the estimated change in scores. Complete results for each of the multilevel random coefficient models are provided in a supplemental table.

Overall scores for the depression domain significantly decreased (improved) over the perioperative period, with a mean estimated decline of -1.6 points per month ($p = 0.03$). Scores for the interest in sex domain increased, with a mean estimated

Table 2 PROMIS assessment characteristics

	Median	Mean	Standard deviation
Minutes to complete total assessment	4.00	4.67	4.11
Minutes to complete depression domain	1.00	0.98	0.85
Minutes to complete interest in sex domain	1.00	0.75	1.10
Minutes to complete pain interference domain	1.00	1.53	1.81
Minutes to complete ability to participate in social roles and activities domain	1.00	1.10	1.56
Depression domain items completed (range 1 to 12)	4.00	6.59	3.36
Interest in sex domain items completed (range 1 to 4)	4.00	3.96	0.34
Pain interference domain items completed (range 1 to 12)	4.00	5.71	3.20
Ability to participate in social roles and activities domain completed (range 1 to 12)	4.00	5.44	2.78
Days from preoperative assessment to procedure	29.00	29.53	19.72
Days from procedure to postoperative assessment	27.00	30.77	9.17
Days from preoperative to postoperative assessment	56.00	60.29	22.63
Preoperative depression score	49.90	52.20	11.68
Postoperative depression score	48.20	48.27	10.30
Preoperative interest in sex score	47.40	46.41	10.58
Postoperative interest in sex score	49.80	48.36	11.59
Preoperative pain interference score	56.00	55.41	10.90
Postoperative pain interference score	56.00	54.87	10.55
Preoperative ability to participate in social roles and activities score	47.80	48.44	10.98
Postoperative ability to participate in social roles and activities score	49.40	49.24	10.73

increase of 1.5 points per month ($p = 0.06$); however, the increase was not below the threshold standard for statistical significance. No statistically significant changes were demonstrated over the perioperative period in either the pain interference domain scores (-0.18 points, $p = 0.80$) or the ability to participate in social roles and activities domain scores (0.44 points, $p = 0.55$). Figure 3 presents plots of the model

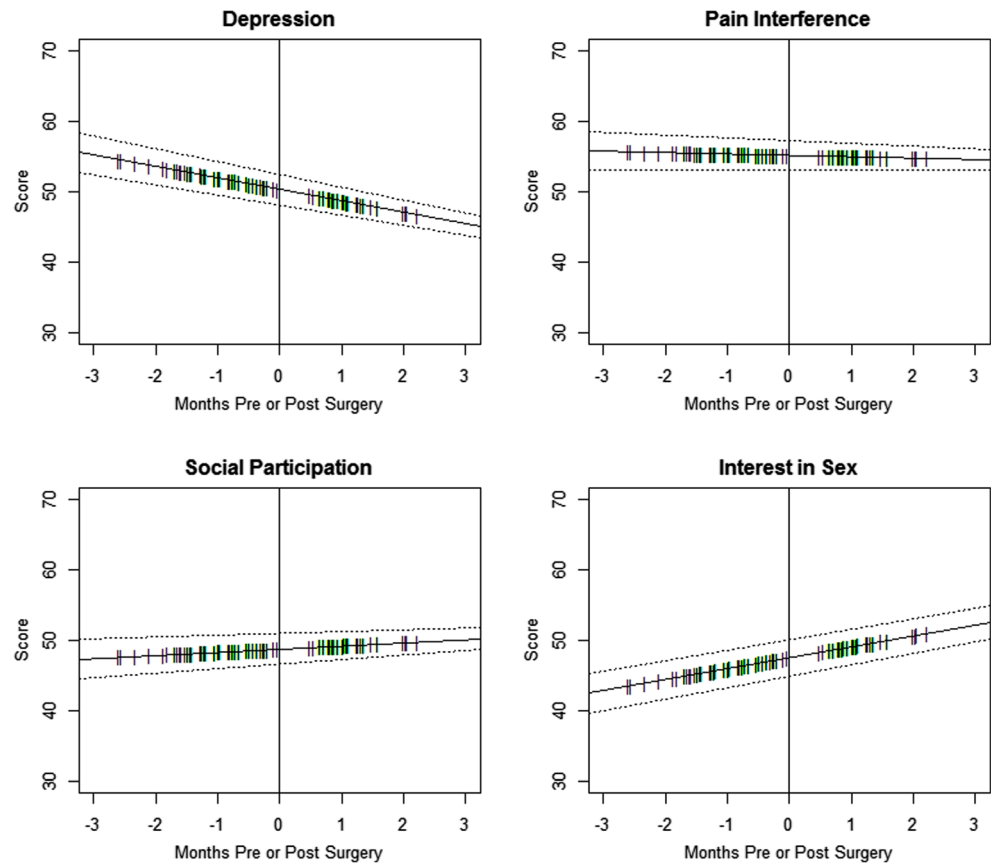
estimated functions obtained for the overall change in scores during the perioperative period, for each of the four PRO domains.

Few patient characteristics were associated with statistically significant differences in the overall pattern of change in PRO domain scores following surgery. Cancer-related patient characteristics were associated with significant changes in pre-

Table 3 Estimated change in PROMIS Scores per month, overall, and by group differences

Change in score per month	PROMIS domains							
	Depression		Interest in sex		Pain interference		Ability to participate in social roles and activities	
	Est.	p value	Est.	p value	Est.	p value	Est.	p value
Overall	-1.62	0.029	1.55	0.060	-0.18	0.803	0.44	0.547
Female vs. male	2.08	0.163	1.43	0.394	1.43	0.345	-2.06	0.180
Race all other vs. White	2.37	0.212	1.05	0.586	-1.81	0.349	1.12	0.578
Insurance other vs. private	-1.56	0.474	3.94	0.166	1.99	0.388	-0.59	0.810
Medicare vs. private	1.46	0.325	0.35	0.836	-1.68	0.304	0.63	0.699
Pre-op colostomy vs. none	-3.26	0.129	0.14	0.960	-2.38	0.278	2.10	0.349
Pre-op ileostomy vs. none	-2.73	0.364	3.83	0.208	-2.96	0.360	2.25	0.476
Oncologic vs. other	-1.01	0.494	-3.70	0.030	0.25	0.869	-0.24	0.871
IBD vs. none	-3.20	0.100	1.04	0.662	0.39	0.850	1.72	0.406
Diverticulitis vs. none	1.18	0.509	1.16	0.600	-1.09	0.540	0.87	0.625
Chemo vs. none	-0.05	0.976	-1.21	0.519	3.49	0.032	-2.83	0.084
Radiation vs. none	-0.56	0.733	-1.07	0.568	3.24	0.051	-1.69	0.312
Laparoscopic vs. other	2.21	0.113	1.78	0.300	0.24	0.871	-0.86	0.565
SSI vs. none	-3.56	0.183	2.44	0.661	-4.07	0.164	4.06	0.136
Operating return vs. not	-5.79	0.127	-0.85	0.900	-7.29	0.091	5.94	0.113
Postop colostomy vs. none	-3.41	0.135	3.28	0.299	0.30	0.901	0.77	0.749
Postop ileostomy vs. none	-3.14	0.055	-3.04	0.113	1.60	0.346	1.84	0.294
Readmitted vs. not	0.23	0.946	-3.87	0.353	0.88	0.793	0.76	0.854
Any morbidity vs. none	-1.38	0.501	-1.84	0.461	-1.22	0.571	1.85	0.394

Fig. 3 Panel plot depicting the linear change in patient-reported outcome measure domain scores during the perioperative period. In each plot, the *solid line* depicts the multilevel random coefficient model estimated linear function for the monthly change in scores, with *parallel dotted lines* indicating the 95% confidence interval for the linear function estimates. *Hatch marks* plotted along the linear function identify points in time when assessments were recorded for individual patients



and post-procedure PRO scores. Pain interference scores for patients with neoadjuvant chemotherapy significantly increased (worsened) over the perioperative period (+3.5 points, $p = 0.03$). Increased scores for pain interference reported for patients with neoadjuvant radiation were near statistical significance (+3.2 points, $p = 0.05$). Scores for the interest in sex domain decreased (worsened) for patients with oncologic etiology (-3.7 points, $p = 0.03$), compared to other patients.

Postoperative ileostomy was the only other patient characteristic associated with differences in scores large enough to be statistically significant at the $p < 0.05$ threshold. Decreased depression scores were reported by patients with postoperative ileostomy, in comparison to patients without postoperative stomas, were nearly statistically significant (-3.1, $p = 0.06$). Figure 4 presents plots of the estimated change in score per month during the perioperative period, for each of the four examples of group differences that were statistically significant or nearly significant at the $p < 0.05$ threshold.

The observed results were not sensitive to the model assumption that change from pre- to post-assessment was constant over time. Each domain and group characteristic combination associated with a statistically significant change ($p < 0.05$) was demonstrated to also have a statistically significant change in the sensitivity analysis, where the preoperative assessment was assumed to have occurred on the day of surgery. Only one domain and group characteristic combination

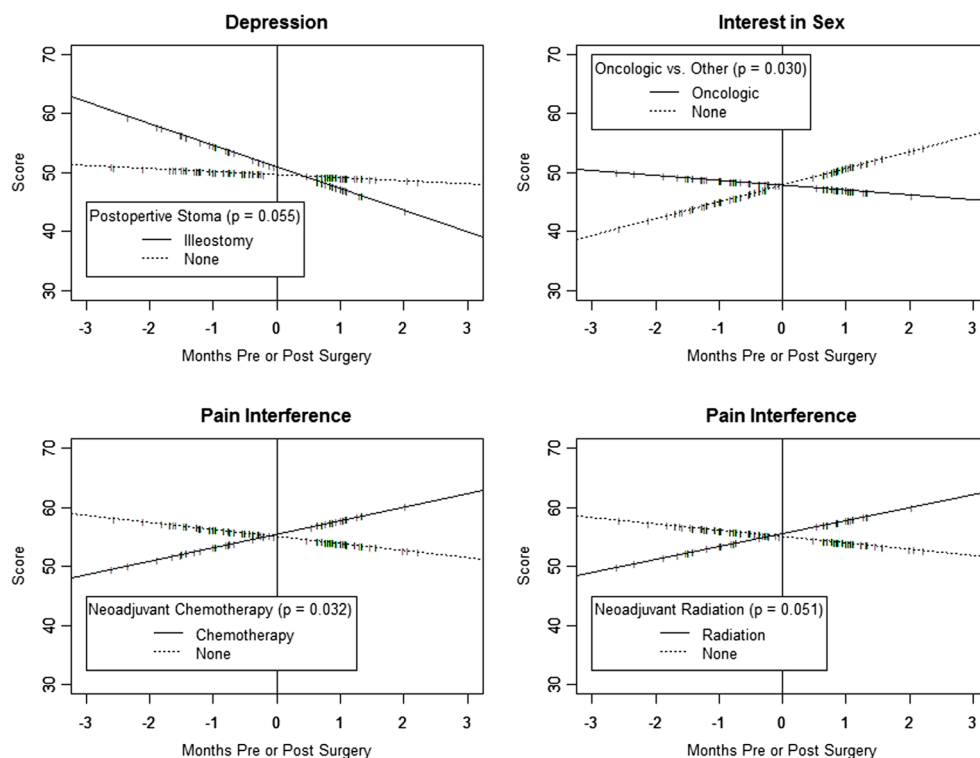
that was demonstrated to not be statistically significant in the original analysis (pain interference ~ return to operating room, $p = 0.091$) was statistically significant in the sensitivity analysis simulation ($p = 0.048$). Complete results for the sensitivity analysis conducted for each of the multilevel random coefficient models are provided in a supplemental table.

Discussion

This study demonstrates that PRO assessment can be conducted using the PROMIS assessment center software tools for patients undergoing major abdominal surgery in under 5 min during the routine clinic visit. The ready collection and availability of this information could be used to improve and inform patient care. We found that patients undergoing major colorectal surgery demonstrated no statistically significant change overall in scores for pain interference or social participation during the perioperative period, demonstrated a statistically significant decrease overall in depression scores between their pre and postoperative assessments (-1.62 points, $p = 0.03$), and demonstrated nearly statistically significant increases in scores for interest in sex (+1.55 points, $p = 0.06$).

However, cancer patients had statistically significant declines in scores for several PROs, compared to other patients. Post-procedure pain interference scores for patients with

Fig. 4 Panel plot depicting the difference in the linear change in selected patient-reported outcome measure domain scores during the perioperative period, for four example groups. In each plot, the *solid and dotted lines* depict the multilevel random coefficient model estimated linear functions for the monthly change in scores, for specified groups. *Hatch marks* plotted along the linear function identify points in time when assessments were recorded for individual patients in each compared group



neoadjuvant chemotherapy (+3.5 points, $p = 0.03$) and for patients with neoadjuvant radiation (+3.2 points, $p = 0.05$) were significantly increased, and scores for the interest in sex domain decreased (worsened) for patients with oncologic etiology (-3.7 points, $p = 0.03$). These effect sizes are within the range of detectable improvements or declines in scores, and they indicate clinically significant magnitudes of change. Prior studies demonstrate that PROMIS score changes in the range from 3 to 5 points are sufficient for identifying clinically meaningful changes in pain interference and other PROMIS domains measured over time [23–25].

Other studies have assessed quality-of-life outcomes following colorectal surgery. Most recently, Brown et al., using EORTC QLQ-C30/CR38 as part of the MRC-CLASICC trial, reported that the development of complications had a negative impact on long-term quality-of-life outcomes [26]. However, to our knowledge, we are the first group to report patient-reported outcomes using the PROMIS instrument in patients following major abdominal surgery. The PROMIS system is advantageous because it provides clinicians and researchers with a set of instruments that serves as the scientific foundation for patient-centered research as prioritized by the NIH. It enables researchers to customize the assessment measurements to suit the needs of the individual project while also providing a validated instrument that can be understood and shared across disciplines.

Several limitations attend this research. The study was powered to assess the significance of large differences in post-surgery scores between groups, equivalent to a difference

of one standard deviation from the US general population mean. Prior research comparing PROMIS domain scores to clinically relevant symptom severity levels demonstrates that increments of one standard deviation can be used to define reliable threshold values for grouping patients by symptom severity, such as 50–59 (mild), 60–69 (moderate), and ≥ 70 (severe) for pain and 55–64 (mild), 65–74 (moderate), and ≥ 75 (severe) for depression [22].

While the study is adequately powered to assess the significance of large differences, assessing the significance of smaller differences associated with select patient level characteristics will require larger study populations. Given the unique nature of neoadjuvant and surgical treatment in rectal cancer patients, future studies are needed in this patient population to fully elucidate the impact of surgery on PROs. The study population included all available patients consented and receiving surgery during the 22-month study period. However, a larger study population recruited over a longer time interval would yield better estimates. The study population was recruited from a single large tertiary hospital within a standardized ERAS program, and the study population demonstrated relatively low rates of complications following surgery. The results may not be generalizable to patients receiving care in other settings or traditional care pathways. It is also important to note that all patients underwent elective surgery and were relatively healthy at baseline. These results are not generalizable to patients undergoing emergent colorectal surgery.

Post-procedure assessments were obtained for 85% of patients following surgery. However, there is a potential for a

selection bias if patients who were struggling in the postoperative period did not complete the survey. We assessed this potential for selection bias by comparing the demographic, preoperative, and post-procedure characteristics of the subset of patients with surgery but without a post-procedure assessment to the final study population. We found that patients without post-procedure assessments were more likely to have reported any morbidity 30 days following surgery and were more likely to have been readmitted following discharge. Patients without post-procedure assessments also had significantly lower frequencies of laparoscopic surgery, which is associated with fewer complications and quicker recovery than open surgical procedures [27]. These comparisons suggest that post-procedure assessments were not obtained for some patients because of follow-up schedule changes and other exigencies related to complications and readmissions following surgery.

Another limitation of the study is that assessment results were obtained only for the immediate perioperative period. Longer term follow-up of patients, while requiring additional resources and different protocols for scheduling assessment data collection, would provide important information about patient outcomes. Examining patient outcomes over longer periods of time following surgery may also produce different results regarding the relationships between demographic, pre-procedure, and post-procedure events. Extending the follow-up period to include additional longitudinal assessment points could also reduce the number patients with missing post-procedure assessments, by providing additional opportunities to collect PROs at later follow-up appointments.

Finally, we focused the pilot study on four domains (depression, pain interference, interest in sex, and ability to function socially) that were selected following discussions with stakeholders. We selected only four domains in order to minimize the time required to complete the assessment. At the conclusion of the study, we learned that most patients were able to complete the assessment in less than 5 min. Several other domains are of great importance, including information about gastrointestinal symptoms, fatigue, pain intensity, and anxiety. Our results suggest that the number of domains included in the assessment protocol could be expanded to include several additional domains of interest, without increasing the duration of the assessment beyond 10 min for most patients.

Conclusions

Recovery from surgery is a complicated process that has effects on the physical, emotional, and social domains of health. Patients may experience a variety of emotions, ranging from depression, due to the loss of autonomy, to anxiety at not knowing what will happen next. There are also social

challenges that result from the loss of independence following major surgery. These data suggest that the majority of patients quickly return to baseline physical, mental, and social function following colorectal surgery. This information is essential for preoperative patient counseling about the typical impact of colorectal surgery on quality of life. Further studies are needed to validate our results in a broader population, to provide information about other domains of key interest, and to more closely examine the effect of individual characteristics (such as diagnosis and surgical type) and the development of complications on postoperative changes in PROs among colorectal surgery patients.

Acknowledgements PROMIS® was funded with cooperative agreements from the National Institutes of Health (NIH) Common Fund Initiative (U54AR057951, U01AR052177, U54AR057943, U54AR057926, U01AR057948, U01AR052170, U01AR057954, U01AR052171, U01AR052181, U01AR057956, U01AR052158, U01AR057929, U01AR057936, U01AR052155, U01AR057971, U01AR057940, U01AR057967, U01AR052186). The contents of this article uses data developed under PROMIS. These contents do not necessarily represent an endorsement by the US Federal Government or PROMIS. See www.nihpromis.org for additional information on the PROMIS initiative.

Contributions Each author has made substantial contributions to conception and design, or acquisition of data, or analysis and interpretation of data.

Compliance with ethical standards

Conflict of interest The authors declare that there are no conflicts of interest to report.

Dr. Stukenborg retains full control of all primary data used in this analysis and the authors agree to allow the journal to review the data if requested. The authors have no possible conflicts of interest in the manuscript, including no financial, consultant, institutional, and other relationships that might lead to bias.

Open Access This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

1. Temple L, Fuzesi S, Patil S (2009) The importance of determining quality of life in clinical trials. *Surgery* 145:622–626
2. Basch E, Torda P, Adams K (2013) Standards for patient-reported outcome-based performance measures. *JAMA* 310:139–140
3. Basch E, Abernethy AP, Mullins CD, Reeve BB, Smith ML, Coons SJ, Sloan J, Wenzel K, Chauhan C, Eppard W, Frank ES, Lipscomb J, Raymond SA, Spencer M, Tunis S (2012) Recommendations for incorporating patient-reported outcomes into clinical comparative effectiveness research in adult oncology. *J Clin Oncol* 30:4249–4255
4. Franklin PD, Lewallen D, Bozic K, Hallstrom B, Jiranek W, Ayers DC (2014) Implementation of patient-reported outcome measures

- in U.S. Total joint replacement registries: rationale, status, and plans. *J Bone Joint Surg Am* 96(Suppl 1):104–109
5. Berliner JL, Brodke DJ, Chan V, SooHoo NF, Bozic KJ (2016) John Chamley award: preoperative patient-reported outcome measures predict clinically meaningful improvement in function after THA. *Clin Orthop Relat Res* 474:321–329
 6. Federal Register. Vol. 80, No. 134. Tuesday, July 14, 2015. Proposed rules. Medicare program; Comprehensive care for joint replacement payment model for Acute Care Hospitals Furnishing Lower Extremity Joint Replacement Services. Section III.C.5.b.(2)
 7. Bilimoria KY, Cella D, Butt Z (2014) Current challenges in using patient-reported outcomes for surgical care and performance measurement: everybody wants to hear from the patient, but are we ready to listen? *JAMA Surgery* 149:505–506
 8. McNair A, Whistance RN, Forsythe RO, Rees J, Jones JE, Pullyblank AM, Avery K, Brookes ST, Thomas MG, Sylvester PA, Russell A, Oliver A, Morton D, Kennedy R, Jayne DG, Huxtable R, Hackett R, Dutton SJ, Coleman MG, Card M, Brown J, Blazeby JM (2015) Synthesis and summary of patient-reported outcome measures (PROMs) to inform the development of a core outcome set in colorectal cancer surgery. *Color Dis* 17:217–229
 9. Garcia SF, Cella D, Clauser SB, Flynn KE, Lad T, Lai JS, Reeve BB, Smith AW, Stone AA, Weinfurt K (2007) Standardizing patient-reported outcomes assessment in cancer clinical trials: a patient-reported outcomes measurement information system initiative. *J Clin Oncol* 25:5106–5112
 10. Cella D, Riley W, Stone A, Rothrock N, Reeve B, Yount S, Amtmann D, Bode R, Buysse D, Choi S, Cook K, Devellis R, DeWalt D, Fries JF, Gershon R, Hahn EA, Lai JS, Pilkonis P, Revicki D, Rose M, Weinfurt K, Hays R (2010) The patient-reported outcomes measurement information system (PROMIS) developed and tested its first wave of adult self-reported health outcome item banks: 2005–2008. *J Clin Epidemiol* 63:1179–1194
 11. Thiele RH, Rea KM, Turrentine FE, Friel CM, Hassinger TE, McMurry TL, Goudreau BJ, Umapathi BA, Kron IL, Sawyer RG, Hedrick TL (2015) Standardization of care: impact of an enhanced recovery protocol on length of stay, complications, and direct costs after colorectal surgery. *J Am Coll Surg* 220:430–443
 12. Gershon RC, Rothrock N, Hanrahan R, Bass M, Cella D (2010) The use of PROMIS and assessment center to deliver patient-reported outcome measures in clinical research. *Journal of applied measurement* 11:304–314
 13. Revicki DA, Cella DF (1997) Health status assessment for the twenty-first century: item response theory, item banking and computer adaptive testing. *Qual Life Res* 6:595–600
 14. Hung M, Stuart AR, Higgins TF, Saltzman CL, Kubiak EN (2014) Computerized adaptive testing using the PROMIS physical function item bank reduces test burden with less ceiling effects compared with the short musculoskeletal function assessment in orthopaedic trauma patients. *J Orthop Trauma* 28:439–443
 15. PROMIS Scoring Manuals. Patient reported outcomes measurement information system. Assessment Center. Accessed on 19 Jan 2016
 16. Liu H, Cella D, Gershon R, Shen J, Morales LS, Riley W, Hays RD (2010) Representativeness of the patient-reported outcomes measurement information system internet panel. *J Clin Epidemiol* 63:1169–1178
 17. Stukenborg GJ, Blackhall L, Harrison J, Barclay JS, Dillon P, Davis MA, Mallory M, Read P (2014) Cancer patient-reported outcomes assessment using wireless touch screen tablet computers. *Qual Life Res* 23:1603–1607
 18. Laird NM, Ware JH (1982) Random-effects models for longitudinal data. *Biometrics* 38:963–974
 19. Rajeswaran J, Blackstone EH (2015) Patient-reported outcomes and importance of their appropriate statistical analyses. *J Thorac Cardiovasc Surg* 150:461–462
 20. Detry MA, Ma Y (2016) Analyzing repeated measurements using mixed models. *JAMA* 315:407–408
 21. Guo Y, Logan HL, Glueck DH, Muller KE (2013) Selecting a sample size for studies with repeated measures. *BMC Med Res Methodol* 13:100
 22. Cella D, Choi S, Garcia S, Cook KF, Rosenbloom S, Lai J-S, Tatum DS, Gershon R (2014) Setting standards for severity of common symptoms in oncology using the PROMIS item banks and expert judgment. *Qual Life Res* 23:2651–2661
 23. Jensen RE, Moynour CM, Potosky AL, Lobo T, Hahn EA, Hays RD, Cella D, Smith AW, Wu XC, Keegan TH, Paddock LE, Stroup AM, Eton DT (2016) Responsiveness of 8 patient-reported outcomes measurement information system (PROMIS) measures in a large, community-based cancer study cohort. *Cancer* 123:327–335
 24. Yost KJ, Eton DT, Garcia SF, Cella D (2011) Minimally important differences were estimated for six patient-reported outcomes measurement information system—cancer scales in advanced-stage cancer patients. *J Clin Epidemiol* 64:507–516
 25. Hays RD, Spritzer KL, Fries JF, Krishnan E (2015) Responsiveness and minimally important difference for the patient-reported outcomes measurement information system (PROMIS) 20-item physical functioning short form in a prospective observational study of rheumatoid arthritis. *Ann Rheum Dis* 74:104–107
 26. Brown SR, Mathew R, Keding A, Marshall HC, Brown JM, Jayne DG (2014) The impact of postoperative complications on long-term quality of life after curative colorectal cancer surgery. *Ann Surg* 259:916–923
 27. King PM, Blazeby JM, Ewings P, Longman RJ, Kipling RM, Franks PJ, Sheffield JP, Evans LB, Soulsby M, Bulley SH, Kennedy RH (2006) The influence of an enhanced recovery programme on clinical outcomes, costs and quality of life after surgery for colorectal cancer. *Color Dis* 8:506–513