




Original Article

Corresponding Author

Kern Singh

 <https://orcid.org/0000-0002-6118-7273>

Department of Orthopaedic Surgery, Rush University Medical Center, 1611 W. Harrison St, Suite #300, Chicago, IL 60612, USA

Email: kern.singh@rushortho.com

Received: November 4, 2020

Revised: February 1, 2021

Accepted: July 5, 2021

What Can Legacy Patient-Reported Outcome Measures Tell Us About Participation Bias in Patient-Reported Outcomes Measurement Information System Scores Among Lumbar Spine Patients?

Conor P. Lynch, Elliot D.K. Cha, Caroline N. Jadczyk, Shruthi Mohan, Cara E. Geoghegan, Kern Singh

Department of Orthopaedic Surgery, Rush University Medical Center, Chicago, IL, USA

Objective: Patient-Reported Outcomes Measurement Information System (PROMIS) is a validated tool for assessing patient-reported outcomes in spine surgery. However, PROMIS is vulnerable to nonresponse bias. The purpose of this study is to characterize differences in patient-reported outcome measure scores between patients who do and do not complete PROMIS physical function (PF) surveys following lumbar spine surgery.

Methods: A prospectively maintained database was retrospectively reviewed for primary, elective lumbar spine procedures from 2015 to 2019. Outcome measures for Patient Health Questionnaire-9 (PHQ-9), visual analogue scale (VAS) back & leg, Oswestry Disability Index (ODI), and 12-item Short Form health survey physical composite summary (SF-12 PCS) were recorded at both preoperative and postoperative (6 weeks, 12 weeks, 6 months, 1 year, 2 years) timepoints. Completion rates for PROMIS PF surveys were recorded and patients were categorized into groups based on completion. Differences in mean scores at each timepoint between groups was determined.

Results: Eight hundred nine patients were included with an average age of 48.1 years. No significant differences were observed for all outcome measures between PROMIS completion groups preoperatively. Postoperative PHQ-9, VAS back, VAS leg, and ODI scores differed significantly between groups through 1 year (all $p < 0.05$). SF-12 PCS differed significantly only at 6 weeks ($p = 0.003$).

Conclusion: Patients who did not complete PROMIS PF surveys had significantly poorer outcomes than those that did in terms of postoperative depressive symptoms, pain, and disability. This suggests that patients completing PROMIS questionnaires may represent a healthier cohort than the overall lumbar spine population.

Keywords: Patient-reported outcomes, Patient-Reported Outcomes Measurement Information System, Lumbar spine



This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<https://creativecommons.org/licenses/by-nc/4.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Copyright © 2022 by the Korean Spinal Neurosurgery Society

INTRODUCTION

As the frequency of spinal procedures has steadily increased, so too has the use of patient-reported outcome measures (PROMs). Legacy PROMs such as the Oswestry Disability Index (ODI),

visual analogue scale (VAS) and the 12-item Short Form health survey (SF-12) measure perceptions of pain, disability, and physical function (PF), but are dated in their ability to provide more personalized assessment. More recent metrics such as the Patient-Reported Outcomes Measurement Information System

(PROMIS) utilize computer adaptive testing which customizes questions based on previous responses and provides more efficient and focused assessments of patient-reported outcomes (PROs). Additionally, PROMIS surveys have demonstrated minimal floor and ceiling effects, and their use among spine patients is well-validated.¹⁻⁴

Although PROs are of utmost importance to accurately track postoperative improvement, noncompliance is nearly inevitable with any self-report measure and bias may be thereby introduced. Participation bias (also known as nonresponse bias) can occur when there are significant differences between respondents and nonrespondents and may lead to an inaccurate representation of the population at large. Previous studies have demonstrated several important differences between respondents and nonrespondents to PROM surveys. Parrish et al.⁵ examined demographic and perioperative variables as predictors of survey completion and reported that patients of African-American or Hispanic race and those with radicular pain were less likely to complete surveys. Conversely, older individuals and patients with more severe depressive symptoms were more likely to complete PROMIS PF questionnaires. Furthermore, several other studies demonstrated higher completion rates among older individuals, and those with postoperative complications, whereas male sex, younger age, lower socioeconomic status and non-White race were reported as predictors of decreased compliance.⁶⁻⁸

The variability in demographics and perioperative characteristics that may predict respondent and nonrespondent status may have implications for the outcomes experienced by these patients as well. If key differences in PROs exist between respondents and nonrespondents, the data obtained by these surveys may be unrepresentative of the patient population as a whole and, if taken at face value, may misguide clinical decision making or lead some patients to receive inappropriate or inadequate care. Given that it is predicated on the *absence* of data, participation bias is inherently difficult to quantify. Several orthopedic studies have employed different tactics, such as telephone outreach, to quantify outcomes in nonrespondents.⁹⁻¹¹ One avenue to elucidate PROM trends in nonrespondents that has not been well explored is the use of data from “partial-respondents” who may have completed other PROMs, but not PROMIS PF. The establishment of strong correlations between PROMIS PF and VAS, ODI, and SF-12 at multiple timepoints for both lumbar and cervical spine patients may allow for an alternative route to quantify outcome measures of nonrespondents.^{2,3,12} More specifically, use of completed legacy PROMs to extrapolate important data about potential PROMIS scores could provide insight to the true status of

postoperative PF among nonrespondents. As the popularity and applications of PROMIS surveys continue to expand, it becomes more important than ever to quantify the impact of participation bias on their results. This study aims to explore the extent of participation bias for PROMIS PF in a cohort of lumbar spine patients by analyzing differences in legacy PROM scores between PROMIS respondents and nonrespondents.

MATERIALS AND METHODS

1. Patient Population

Prior to study onset, this study was approved by the Institutional Review Board of Rush University Medical Center (ORA #14051301) and written informed consent were obtained from patients. A private registry of prospectively maintained spine surgery data was retrospectively reviewed for patients that underwent primary, elective lumbar spine procedures, which included fusions, decompressions, and discectomies between the dates of May 2015 and July 2019. Revision procedures along with surgeries indicated for traumatic, infectious, or malignant etiologies were excluded.

2. Data Collection

The following patient demographic characteristics were collected: age, sex, body mass index (BMI), preoperative smoking status, diabetic status, American Society of Anesthesiologists (ASA) physical status classification, Charlson Comorbidity Index (CCI), ethnicity, and insurance/payment type received. Preoperative spinal pathologies were classified as herniated nucleus pulposus (new-onset or recurrent), degenerative spondylolisthesis, isthmic spondylolisthesis, and scoliosis. Perioperative characteristics were recorded including operative duration (in minutes), estimated blood loss (EBL; in mL), and postoperative length of stay (in hours). PROMs assessing pain (VAS back and leg), disability (ODI), PF (SF-12 physical component summary [SF-12 PCS]), and depressive symptoms (Patient Health Questionnaire-9 [PHQ-9]) were collected at preoperative and 6-week, 12-week, 6-month, and 1-year postoperative timepoints. All PROMs were completed either during clinic appointments using a hand-held tablet device or remotely using the patients' personal devices through an online portal. Patients completing PROMs during clinic appointments were required to finish surveys before meeting with clinicians to avoid any biases.

3. Statistical Analysis

All statistical tests and calculations were performed using

Stata 16.1 (StataCorp, College Station, TX, USA). Descriptive statistics were reported for patient demographic characteristics, preoperative spinal diagnoses, and perioperative variables. Perioperative variables were reported separately for patients who underwent lumbar fusion and patients who underwent lumbar decompression/discectomy. Outlier analysis was performed to identify patients with operative duration, EBL, or length of stay > 3 standard deviations above or below the mean value. Outliers were excluded to limit the amount of bias introduced by highly atypical cases. Patients were categorized at each timepoint as PROMIS respondents or nonrespondents based on whether they had completed a PROMIS PF survey corresponding to that given timepoint. Chi-square and Student t-test were used to compare demographic and perioperative variables between PROMIS respondents and nonrespondents. Student t-test for independent samples was used to compare scores for each of the other included PROMs between PROMIS respondents and nonrespondents at each timepoint. A p-value of ≤ 0.05 was set as the threshold for statistical significance for all tests.

RESULTS

A total of 827 eligible lumbar spine patients were initially identified. Following removal of outliers, 809 patients were included in final analysis. The overall cohort had a mean age of 48.1 years and a majority were male (66.9%) and nonobese (BMI < 30 kg/m²; 56.7%). The mean ASA classification was 1.9 and mean CCI score was 1.4. Ethnicity ($p < 0.001$) and insurance/payment type ($p < 0.001$) were significantly associated with PROMIS completion status (Table 1). The study cohort included 335 lumbar fusion patients among whom degenerative spondylolisthesis was the most common preoperative spinal pathology (49.0%) and means for perioperative variables were as follows: operative duration 136.6 \pm 45.8 minutes, EBL 52.1 \pm 30.4 mL, and length of stay 32.7 \pm 21.5 hours. The study cohort included 474 lumbar decompression/discectomy patients among whom herniated nucleus pulposus was the most common spinal pathology (82.7%) and means for perioperative variables were as follows: operative duration 46.0 \pm 16.7 minutes, EBL 26.9 \pm 9.2 mL, and length of stay 5.8 \pm 7.6 hours. No perioperative variables significantly differed between PROMIS PF respondents and nonrespondents for either procedure type (Table 2).

No significant preoperative differences in any of the included PROMs were observed between PROMIS respondents and nonrespondents. PHQ-9 scores were significantly more severe for nonrespondents at 6 weeks (3.3 vs. 5.7, $p < 0.001$), 12 weeks

Table 1. Patient demographics

Demographic	Total (n = 809)	Completed PROMIS PF (n = 570)	Incomplete PROMIS PF (n = 239)	p-value [†]
Age (yr)	48.1 \pm 13.3	48.0 \pm 13.2	48.2 \pm 13.6	0.918
Sex				0.603
Female	268 (33.1)	192 (33.7)	76 (31.8)	
Male	541 (66.9)	378 (66.3)	163 (68.2)	
Body mass index (kg/m ²)				0.617
< 30	450 (56.7)	320 (57.3)	130 (55.3)	
≥ 30	344 (43.2)	239 (43.8)	105 (44.7)	
Smoking status				0.506
Nonsmoker	707 (87.4)	501 (87.9)	206 (86.2)	
Smoker	102 (12.6)	69 (12.1)	33 (13.3)	
Diabetes				0.067
Nondiabetic	743 (91.8)	530 (93)	213 (89.1)	
Diabetic	66 (8.2)	40 (7.0)	26 (10.9)	
ASA PS classification	1.9 \pm 0.6	1.8 \pm 0.6	1.9 \pm 0.6	0.244
CCI score	1.4 \pm 1.5	1.3 \pm 1.5	1.4 \pm 1.5	0.823
Ethnicity				< 0.001*
White	562 (69.7)	437 (76.9)	125 (52.5)	
African-American	82 (10.2)	49 (8.6)	33 (13.9)	
Hispanic	107 (13.3)	48 (8.5)	59 (24.8)	
Asian	31 (3.9)	22 (3.9)	9 (3.8)	
Other	24 (3.0)	12 (2.1)	12 (5.0)	
Insurance				< 0.001*
Medicare/Medicaid	25 (3.1)	18 (3.2)	7 (2.9)	
Workers' compensation	212 (26.3)	112 (19.7)	100 (41.8)	
Private	570 (70.6)	438 (77.1)	132 (55.2)	

Values are presented as mean \pm standard deviation or number (%). PROMIS PF, Patient-Reported Outcomes Measurement Information System physical function; ASA PS, American Society of Anesthesiologists physical status; CCI, Charlson Comorbidity Index.

* $p < 0.05$, statistically significant differences. [†]p-values calculated using t-test for independent samples (continuous) or chi-square analysis (categorical).

(3.6 vs. 5.4, $p = 0.005$), 6 months (3.7 vs. 5.3, $p = 0.007$), and 1 year (4.0 vs. 5.7, $p = 0.042$). VAS back pain scores were significantly higher for nonrespondents at 6 weeks (3.1 vs. 3.8, $p = 0.004$), 12 weeks (3.2 vs. 4.1, $p = 0.003$), 6 months (3.3 vs. 4.3, $p = 0.002$), and 1 year (3.2 vs. 4.5, $p = 0.004$). VAS leg pain scores were significantly higher for nonrespondents at 6 weeks (2.8 vs. 3.6, $p = 0.004$), 12 weeks (2.8 vs. 3.4, $p = 0.047$), 6 months (2.8 vs. 3.7, $p = 0.011$), and 1 year (2.7 vs. 3.9, $p = 0.011$). ODI scores indicat-

Table 2. Spinal pathology and operative characteristics

Characteristic	Total	Completed PROMIS PF	Incomplete PROMIS PF	p-value [†]
Lumbar fusion	(n = 335)	(n = 236)	(n = 99)	
Spinal pathology				
Degenerative spondylolisthesis	164 (49.0)	118 (50.0)	46 (45.5)	0.555
Isthmic spondylolisthesis	88 (26.3)	59 (25.0)	29 (29.3)	0.415
Recurrent HNP	45 (13.4)	27 (11.4)	18 (18.2)	0.099
Degenerative scoliosis	35 (10.5)	26 (11.0)	9 (9.1)	0.599
Operative time [‡] (min)	136.6 ± 45.8	137.1 ± 45.7	135.3 ± 46.3	0.749
Estimated blood loss (mL)	52.1 ± 30.4	51.1 ± 29.5	54.6 ± 32.5	0.351
Length of stay (hr)	32.7 ± 21.5	31.5 ± 19.5	35.6 ± 25.4	0.114
Lumbar decompression/discectomy	(n = 474)	(n = 334)	(n = 140)	
Spinal pathology				
HNP	392 (82.7)	270 (80.8)	122 (87.1)	0.098
Central stenosis	362 (76.4)	257 (77.0)	105 (75.0)	0.649
Foraminal stenosis	254 (53.6)	178 (53.3)	76 (54.3)	0.843
Operative time [‡] (min)	46.0 ± 16.7	46.0 ± 15.9	46.0 ± 18.6	0.995
Estimated blood loss (mL)	26.9 ± 9.2	26.5 ± 7.1	27.7 ± 12.9	0.225
Length of stay (hr)	5.8 ± 7.6	5.4 ± 6.9	6.5 ± 9.0	0.158

Values are presented as number (%) or mean ± standard deviation.

PROMIS PF, Patient-Reported Outcomes Measurement Information System physical function; HNP, herniated nucleus pulposus.

[†]p-values calculated using t-test for independent samples (continuous) or chi-square analysis (categorical). [‡]Skin incision to skin closure.

ed significantly more severe disability for nonrespondents at 6 weeks (27.6 vs. 35.0, $p < 0.001$), 12 weeks (25.2 vs. 33.9, $p < 0.001$), 6 months (24.1 vs. 32.3, $p < 0.001$), and 1 year (22.8 vs. 31.7, $p = 0.009$). SF-12 PCS scores were significantly poorer for nonrespondents at 6 weeks (35.9 vs. 32.5, $p = 0.003$), but not at any other timepoint (all $p \geq 0.090$). A summary of postoperative PROM improvement by respondent group can be found in Table 3.

DISCUSSION

Defined as key differences in nonrespondents and respondents to a survey in a given population that may influence overall results, participation bias (also known as nonresponse bias) is a concern for clinical research, particularly those focused on PROMs. Such biases can be influenced both by the rate of response and the degree of difference between respondents and nonrespondents. Given that this bias is predicated on the absence of data, it is particularly difficult to quantify. Previous studies have utilized a variety of different methods to explore participation bias in surgical patients and report a wide range of results. Our analysis indicates significant differences in both the

physical and mental health PROM scores of lumbar spine patients between respondents and nonrespondents to PROMIS PF surveys. These differences raise concerns for nonnegligible participation bias in the PROMIS scores of lumbar spine patients.

The challenging task of quantifying participation bias in PROs has necessitated a good deal of creativity on the part of researchers. One method employed by multiple orthopedic studies involves the use of a relatively generic, mail-based survey to categorize “respondents” and “nonrespondents,” followed by self-reported and more objective clinical data collection at subsequent follow-up appointments. Both Kwon et al.¹³ and Kim et al.¹⁴ conducted such analyses using a mail-based survey assessing satisfaction and functional status in patients undergoing total knee arthroplasty.

Telephone-based outreach has also been utilized by several groups to connect with patients that did not respond to initial survey requests. In a study of patients from the Danish Shoulder Arthroplasty Registry, Polk et al.¹⁵ utilized both postal reminders and telephone contact to increase completion rates of the Western Ontario Osteoarthritis of the Shoulder index from 65% to 82%. Højmark et al.¹¹ also studied nonrespondents to a

Table 3. Outcomes by PROMIS completion status

Variable	Completed PROMIS PF	Incomplete PROMIS PF	p-value [†]
PHQ-9			
Preoperative	6.0 ± 5.7 (448)	6.9 ± 6.6 (144)	0.117
6 Weeks	3.3 ± 4.4 (347)	5.7 ± 6.0 (147)	< 0.001*
12 Weeks	3.6 ± 5.2 (257)	5.4 ± 6.0 (105)	0.005*
6 Months	3.7 ± 5.2 (232)	5.3 ± 6.1 (142)	0.007*
1 Year	4.0 ± 5.7 (189)	5.7 ± 6.0 (65)	0.042*
VAS back			
Preoperative	6.1 ± 2.6 (553)	6.5 ± 3.0 (183)	0.130
6 Weeks	3.1 ± 2.7 (435)	3.8 ± 2.8 (188)	0.004*
12 Weeks	3.2 ± 2.8 (316)	4.1 ± 2.7 (133)	0.003*
6 Months	3.3 ± 2.9 (255)	4.3 ± 2.8 (132)	0.002*
1 Year	3.2 ± 2.8 (179)	4.5 ± 2.9 (49)	0.004*
VAS leg			
Preoperative	5.9 ± 2.8 (556)	6.1 ± 2.8 (182)	0.311
6 Weeks	2.8 ± 2.7 (434)	3.6 ± 3.0 (187)	0.001*
12 Weeks	2.8 ± 2.8 (316)	3.4 ± 2.8 (134)	0.047*
6 Months	2.8 ± 3.0 (255)	3.7 ± 2.8 (132)	0.011*
1 Year	2.7 ± 3.0 (180)	3.9 ± 3.2 (49)	0.011*
Oswestry Disability Index			
Preoperative	41.5 ± 17.2 (556)	42.9 ± 17.8 (182)	0.319
6 Weeks	27.6 ± 18.8 (436)	35.0 ± 21.9 (188)	< 0.001*
12 Weeks	25.2 ± 19.8 (315)	33.9 ± 20.3 (137)	< 0.001*
6 Months	24.1 ± 20.4 (256)	32.3 ± 20.4 (132)	< 0.001*
1 Year	22.8 ± 20.7 (177)	31.7 ± 20.9 (49)	0.009*
SF-12 PCS			
Preoperative	31.1 ± 8.0 (539)	32.1 ± 8.4 (155)	0.159
6 Weeks	35.9 ± 10.0 (424)	32.5 ± 9.3 (122)	0.003*
12 Weeks	37.9 ± 11.1 (329)	35.8 ± 9.0 (71)	0.141
6 Months	39.5 ± 11.4 (268)	37.9 ± 11.6 (86)	0.265
1 Year	40.7 ± 11.3 (241)	37.6 ± 12.7 (47)	0.090

Values are presented as mean ± standard deviation (number).

PROMIS PF, Patient-Reported Outcomes Measurement Information System physical function; PHQ-9, Patient Health Questionnaire-9; VAS, visual analogue scale; SF-12 PCS, 12-item Short Form health survey physical composite score.

*p < 0.05, statistically significant differences. †p-values calculated using t-test for independent samples.

mail-based survey from the Danish national spine database (DaneSpine) at 1 year follow-up by initiating contact through a structured telephone interview. Though this study is one of few to examine PROM participation bias in a cohort of spine patients, Cabitza et al.⁹ also utilized telephone-based follow-up,

but used a slightly less conventional method of characterizing and studying “nonrespondents.” These authors utilized phone-based reminders at 3 separate timepoints, and timed the third reminder such that patient responses had essentially ceased before this final outreach was attempted. Predicated on the idea that patients engaged by the third reminder otherwise would likely not have responded, outcome response data from these patients were used as a “proxy” for “true nonrespondents.”

Our group’s collection of a variety of different PROMs at multiple postoperative intervals allows us the opportunity to use data from “partial-respondents” who complete some PROMs but not others, to extrapolate potential trends for missing surveys. A number of previous studies have demonstrated robust correlations between PROMIS PF and the other physical health-related “legacy” PROMs utilized in our study. In their 2-year PROMIS validation study, Jenkins et al.³ demonstrated strong correlations of PROMIS PF with VAS back, VAS leg, ODI, and SF-12 at both short- and long-term follow-up in patients undergoing transforaminal lumbar interbody fusion. This finding has been similarly reproduced in other studies across a number of procedures including lumbar fusions and microdiscectomies,¹⁶⁻¹⁸ with the exception of Vaishnav et al, who reported a weak correlations of PROMIS PF with SF-12 preoperatively.¹⁹ Based on these well-documented correlations, we can be confident that the completed PROM data we do have may provide useful information regarding the potential PROMIS scores for those that did not complete the PROMIS questionnaire. These relationships indicate that in cases where “legacy” PROM scores differ significantly between PROMIS respondents and nonrespondents, PROMIS scores may differ as well.

We identified 2 key demographic variables that were significantly associated with PROMIS completion. Specifically, the nonrespondent group included a larger proportion of patients who were African-American or Hispanic, and patients who made payments through workers’ compensation. Parrish et al.⁵ previously examined demographic factors associated with PROMIS survey completion and reported similar trends of lower survey completion among African-American and Hispanic spine patients. While their study did not replicate our results regarding workers’ compensation patients, several investigations have reported poorer lumbar surgery outcomes among African-American and workers’ compensation populations.^{20,21} These observed demographic variations in PROMIS response rates may contribute to and/or exacerbate the apparent response bias demonstrated in our results.

Although our analysis demonstrated substantial discrepan-

cies in postoperative PROM scores, preoperative PROM scores did not significantly differ for any measure between PROMIS respondents and nonrespondents. This trend was demonstrated for both mental and physical health measures and may be particularly strong given that the greatest number of participants were included at these preoperative timepoints. Other studies of participation bias have reported similar results, with negligible preoperative differences between respondents and nonrespondents, even when significant differences emerged postoperatively.^{9,14} One potential explanation for this observation is that differences in response rates may be influenced by patient experiences, perceptions, or outcomes of surgery. Perhaps patients hold relatively similar perceptions of surgery at the preoperative timepoint, given that they have all decided to pursue elective procedures, but these perceptions may diverge following varying postoperative outcomes and experiences. In fact, a number of studies have demonstrated that postoperative satisfaction is significantly associated with rates of survey completion.^{10,14}

In contrast with our preoperative results, PROMIS PF nonrespondents reported significantly worse back pain, leg pain, and disability at all postoperative timepoints. In their study of total knee arthroplasty patients, Kim et al.¹⁴ demonstrated poorer mean scores and less postoperative improvement in pain, functionality, and Knee Society knee scores in patients that did not respond to their initial, mail-based survey. Cabitza et al.⁹ also demonstrated poorer pain outcomes among survey nonrespondents in their cohort of hip, knee, and spine patients. However, others, such as Højmark et al.¹¹ and Kwon et al.¹³ reported no significant difference in pain scores between respondents and nonrespondents.

PF, as measured by SF-12 PCS, demonstrated the least postoperative difference between PROMIS respondents and nonrespondents, with the nonresponding group demonstrating worse scores at the 6-week timepoint only. In previous validation studies, SF-12 PCS demonstrated some of the strongest, most consistent correlations with PROMIS PF.^{3,4} Given that these measures are both specifically designed to assess physical functioning, the relative lack of difference in SF-12 PCS scores between PROMIS respondents and nonrespondents may be reassuring in terms of the validity of PROMIS data for drawing conclusions about the entire cohort. Our results conflict again with that of Kwon et al.¹³ and also with Cabitza et al.⁹ in this regard, as these studies both reported significantly poorer SF PF scores among survey nonrespondents.

In addition to the differences, we observed in physical health

measures, patient-reported depressive symptoms, as measured by PHQ-9, were also significantly more severe at all postoperative timepoints for patients that did not complete PROMIS. Literature related to participation bias in measures of depression among surgical patients is quite limited. Cabitza et al.⁹ was one of very few studies to include a measure of mental health status and, in contrast with their results regarding physical health, demonstrated no significant difference in SF mental component summary scores between respondents and nonrespondents. Several previous studies in more general medical populations have also reported minimal effects of participation bias with regard to depressive symptoms or mental health outcomes.^{22,23} Nonetheless, the substantial differences we observed in depressive symptoms is concerning, especially considering evidence for a connection between PHQ-9 scores and physical outcomes in patients undergoing spine surgery.²⁴⁻²⁶

The primary limitation of this study is that our use of legacy PROMs as a proxy for PROMIS scores did not allow us to study “complete nonrespondents” who did not complete any PROM surveys at all. These patients may differ from the 2 groups examined in our study in several important ways, and future studies of participation bias should consider alternative ways to capture outcomes in these patients. Additionally, all procedures in this study were performed by a single attending surgeon at the same academic institution. Therefore, the ability to generalize our results regarding PROMIS nonrespondents to other populations may be limited. A follow-up study using a multicenter design and an innovative method of engaging nonrespondents could be helpful to address these limitations. However, the current study provides a novel analysis of PROM trends in patients that did not complete the PROMIS PF survey, and presents important data regarding the potential for participation bias in this measure among lumbar spine patients.

CONCLUSION

No significant preoperative differences were observed for any of the assessed PROM scores between PROMIS respondents and nonrespondents. PROMIS nonrespondents demonstrated significantly poorer postoperative back pain, leg pain, disability, and depressive symptoms than respondents through 1-year following surgery. PF, as quantified by SF-12 PCS, generally did not differ between respondents and nonrespondents. Our results indicate that some degree of nonresponse bias may exist for PROMIS surveys, leading to a potential underestimation of PF deficits in the overall lumbar spine cohort, particularly at

short-term postoperative timepoints. Efforts should be taken whenever possible to maximize survey completion and the outcomes of nonrespondents should be considered alongside available survey data.

NOTES

Conflict of Interest: The authors have nothing to disclose.

Funding/Support: This study received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Author Contribution: Conceptualization: CL, EC, KS; Data curation: CL, EC, CJ, SM, CG, KS; Formal analysis: CL, EC, KS; Methodology: CL, EC, KS; Project administration: CL, EC, CJ, SM, CG, KS; Visualization: CL, EC, CJ, SM, CG, KS; Writing - original draft: CL, EC, CJ, KS; Writing - review & editing: CL, EC, CJ, SM, CG, KS

ORCID

Conor P. Lynch: 0000-0002-9554-2878

Elliot D.K. Cha: 0000-0001-8118-008X

Kern Singh: 0000-0002-6118-7273

REFERENCES

- Boody BS, Bhatt S, Mazmudar AS, et al. Validation of Patient-Reported Outcomes Measurement Information System (PROMIS) computerized adaptive tests in cervical spine surgery. *J Neurosurg Spine* 2018;28:268-79.
- Patel AA, Dodwad SM, Boody BS, et al. Validation of patient Reported Outcomes Measurement Information System (PROMIS) computer adaptive tests (CATs) in the surgical treatment of lumbar spinal stenosis. *Spine (Phila Pa 1976)* 2018; 43:1521-8.
- Jenkins NW, Parrish JM, Cha EDK, et al. Validation of PROMIS physical function in MIS TLIF: 2-year follow-up. *Spine (Phila Pa 1976)* 2020;45:E1516-22.
- Parrish JM, Jenkins NW, Brundage TS, et al. PROMIS physical function predicts postoperative pain and disability following anterior cervical discectomy and fusion. *Clin Spine Surg* 2020;33:382-7.
- Parrish JM, Jenkins NW, Patel DV, et al. Demographic and perioperative factors associated with Patient-reported Outcomes Measurement Information System (PROMIS) survey completion. *Clin Spine Surg* 2020;33:E519-24.
- Solberg TK, Sørli A, Sjaavik K, et al. Would loss to follow-up bias the outcome evaluation of patients operated for degenerative disorders of the lumbar spine? *Acta Orthop* 2011; 82:56-63.
- Tyser AR, Abtahi AM, McFadden M, et al. Evidence of non-response bias in the Press-Ganey patient satisfaction survey. *BMC Health Serv Res* 2016;16(a):350.
- Hutchings A, Neuburger J, Grosse Frie K, et al. Factors associated with non-response in routine use of patient reported outcome measures after elective surgery in England. *Health Qual Life Outcomes* 2012;10:34.
- Cabitz F, Dui LG, Banfi G. PROs in the wild: assessing the validity of patient reported outcomes in an electronic registry. *Comput Methods Programs Biomed* 2019;181:104837.
- Imam MA, Barke S, Stafford GH, et al. Loss to follow-up after total hip replacement: a source of bias in patient reported outcome measures and registry datasets? *Hip Int* 2014;24: 465-72.
- Højmark K, Støttrup C, Carreon L, et al. Patient-reported outcome measures unbiased by loss of follow-up. Single-center study based on DaneSpine, the Danish spine surgery registry. *Eur Spine J* 2016;25:282-6.
- Parrish JM, Jenkins NW, Nolte MT, et al. A Validation of Patient Health Questionnaire-9 for cervical spine surgery. *Spine (Phila Pa 1976)* 2020;45:1668-75.
- Kwon SK, Kang YG, Chang CB, et al. Interpretations of the clinical outcomes of the nonresponders to mail surveys in patients after total knee arthroplasty. *J Arthroplasty* 2010;25: 133-7.
- Kim J, Lonner JH, Nelson CL, et al. Response bias: effect on outcomes evaluation by mail surveys after total knee arthroplasty. *J Bone Joint Surg Am* 2004;86:15-21.
- Polk A, Rasmussen JV, Brorson S, et al. Reliability of patient-reported functional outcome in a joint replacement registry. A comparison of primary responders and non-responders in the Danish Shoulder Arthroplasty Registry. *Acta Orthop* 2013;84:12-7.
- Haws BE, Khechen B, Guntin JA, et al. Validity of PROMIS in minimally invasive transforaminal lumbar interbody fusion: a preliminary evaluation. *J Neurosurg Spine* 2018;29: 28-33.
- Bhatt S, Boody BS, Savage JW, et al. Validation of Patient-reported Outcomes Measurement Information System computer adaptive tests in lumbar disk herniation surgery. *J Am Acad Orthop Surg* 2019;27:95-103.
- Khechen B, Haws BE, Patel DV, et al. PROMIS physical function score strongly correlates with legacy outcome measures

- in minimally invasive lumbar microdiscectomy. *Spine (Phila Pa 1976)* 2019;44:442-6.
19. Vaishnav AS, McAnany SJ, Iyer S, et al. Psychometric evaluation of Patient-reported Outcomes Measurement Information System physical function computer adaptive testing in minimally invasive lumbar spine surgery: an analysis of responsiveness, coverage, discriminant validity, and concurrent validity. *J Am Acad Orthop Surg* 2020;28:717-29.
 20. Gum JL, Glassman SD, Carreon LY. Is type of compensation a predictor of outcome after lumbar fusion? *Spine (Phila Pa 1976)* 2013;38:443-8.
 21. Elsamadicy AA, Kemeny H, Adogwa O, et al. Influence of racial disparities on patient-reported satisfaction and short- and long-term perception of health status after elective lumbar spine surgery. *J Neurosurg Spine* 2018;29:40-5.
 22. Linnenkamp U, Gontscharuk V, Brüne M, et al. Using statutory health insurance data to evaluate non-response in a cross-sectional study on depression among patients with diabetes in Germany. *Int J Epidemiol* 2020;49:629-37.
 23. Vink JM, Willemsen G, Stubbe JH, et al. Estimating non-response bias in family studies: application to mental health and lifestyle. *Eur J Epidemiol* 2004;19:623-30.
 24. Patel DV, Yoo JS, Khechen B, et al. PHQ-9 score predicts postoperative outcomes following minimally invasive transforaminal lumbar interbody fusion. *Clin Spine Surg* 2019;32:444-8.
 25. Alvi MA, Berg J, Goyal A, et al. Mental illnesses among patients undergoing elective anterior cervical discectomy and fusion: analysis from the National Readmissions Database. *Clin Neurol Neurosurg* 2020;193:105765.
 26. Tuck AN, Scribani MB, Grainger SD, et al. The 9-item Patient Health Questionnaire (PHQ-9): an aid to assessment of patient-reported functional outcomes after spinal surgery. *Spine J* 2018;18:1398-405.