The Spine Journal 19 (2019) 1850−1857

THE SPINE JOURNAL

Clinical Study

Creating value in spine surgery: using patient reported outcomes to compare the short-term impact of different orthopedic surgical procedures

David S. Jevotovsky, BA[a], Caroline P. Thirukumaran, MBBS, MHA, PhD[b], Paul T. Rubery, MD[b,*]

[a] *New York University School of Medicine, 550 1st Ave, New York, NY 10016, USA*
[b] *Department of Orthopedics and Rehabilitation, University of Rochester Medical Center, 601 Elmwood Ave, Box 665, Rochester, NY, 14625, USA*
Received 22 January 2019; revised 30 May 2019; accepted 31 May 2019

Abstract

BACKGROUND CONTEXT: Society increasingly asks Medicine to create "value" for patients. As health-care costs rise, this question will become more important. Debate exists regarding the relative "value" of many surgical procedures, including spinal surgery. Comparison of the relative value that patients experience after different orthopedic procedures is theoretical, but informs the ongoing debate.

METHODS: The Patient Reported Outcome Measurement Information System (PROMIS) assessments for Physical Function, Pain Interference, and Depression are routinely collected in our orthopedic clinics. Patients who underwent lumbar discectomy (DSC) or arthroscopic anterior cruciate ligament reconstruction (ACLR) were retrospectively identified. Data relating to PROMIS domains, patient demographics, and other relevant encounter details were extracted. The primary outcomes were (1) preoperative PROMIS domain scores, (2) scores at a minimum of 40 days postoperatively for DSC patients and 133 days postoperatively for ACLR patients, and (3) the change in scores with surgery. Propensity score matching identified age-, sex-, race-, and comorbidity-matched groups from each cohort. Chi-square tests and nonparametric Kruskal-Wallis tests compared the distribution of outcomes and characteristics. Multivariate linear regression models with interactions between the matched cohort and operative phase estimated the change in the outcomes scores between the two cohorts and controlled for the baseline differences between them.

RESULTS: Before surgery, the DSC cohort had lower physical function, higher pain interference and higher depression scores as compared with the ACLR cohort. This pattern remained postoperatively, indicating less desirable outcomes for DSC patients. However, after controlling for their baseline scores, DSC patients experienced significantly greater improvements after surgery of 3.84 (95% CI 1.08−6.60; p=.01), −4.87 (95% CI −7.52 to −2.23; p<.001), and −2.95 (95% CI −5.70 to −0.21; p=.04) points in their physical function, pain interference, and depression scores, respectively, as compared with ACLR patients.

CONCLUSIONS: Based upon PROMIS assessments at short-term follow-up, DSC patients receive a larger benefit from surgery than ACLR despite the overall less desirable postoperative PROMIS scores in the DSC cohort. This result, while theoretical, informs the debate regarding the comparative value of DSC to patients.

Keywords: ACL reconstruction; Patient reported outcome; PROMIS; Discectomy

FDA device/drug status: Not applicable.

Author disclosures: ***DSJ:*** Nothing to disclose. ***CPT:*** Nothing to disclose. ***PTR:*** Nothing to disclose.

Level of Evidence: III

* Corresponding author. Department of Orthopedics and Rehabilitation, University of Rochester Medical Center, 601 Elmwood Avenue, Box 665, Rochester, NY 14642, USA. Tel.: (585) 275-5167; fax: (585) 756-4721.
E-mail address: Paul_Rubery@urmc.rochester.edu (P.T. Rubery).

https://doi.org/10.1016/j.spinee.2019.05.595
1529-9430/

Introduction

Lumbar discectomy (DSC) is one of the most common spine surgeries in the United States [1,2]. Previous studies have documented the value of this surgery for patients [1,3], however, as many as 40% of patients experience persistent pain, functional disability, or decreased quality of life, and 20% undergo further surgery [4−8]. These divergent outcomes spur discussion regarding the appropriate role and rates for, and the value of, spinal surgery [9]. The data comparing patient reported outcomes of spine patients to other common orthopedic pathologies is limited. Comparison of DSC outcomes with other orthopedic surgeries is an artificial construct but further informs the debate regarding the value of spinal surgery.

Patients who undergo anterior cruciate ligament reconstructions (ACLR) are potentially a good population for comparison with DSC. ACLR is also a common orthopedic surgery in the United States. A recent review found that 129,836 ACL reconstructions were performed in 2006, with ongoing increases in the procedure amongst women and in patients both under 20 and over 40 years of age [10−12]. Studies show that the majority of patients experience good long-term outcomes and that approximately 81% of ACLR patients return to some level of athletic activity [13]. Because the underlying pathology and respective patient complaints differ, any previous comparison or commentary on the relative success or benefit of the respective surgeries was matters of opinion.

Patient reported outcome (PRO) assessments have become a routine part of patient care. The Patient Reported Outcomes Measurement Information System (PROMIS) is a PRO measure that was developed with National Institutes of Health funding beginning in 2002 [14]. The PROMIS consists of validated instruments which can be utilized to assess health domains reflecting numerous aspects of a patient's life [15]. The PROMIS can be administered with computer adaptive testing (CAT) based upon item response theory, which decreases administration time while assuring accuracy [16,17]. PROMIS questionnaires have been developed and validated to measure physical functioning (PROMIS-PF), pain interference (PROMIS-PI) and depression (PROMIS-D) status, as well as many other domains such as anxiety, asthma impact, sleep disturbance, and substance use [18−23]. Studies have shown that PROMIS outperforms traditional patient-reported metrics [20,21,24−26]. Thus, PROMIS assessment scores might be used to compare surgery outcomes within orthopedic surgery. Specifically, PROMIS could be used to calculate the preoperative to postoperative change in scores thereby helping to quantify the impact of the surgery.

To our knowledge, there has been no study published that quantitatively measures the value that patients derive from spinal surgery by comparing it with a group of matched patients who underwent another orthopedic surgery. Thus, our aim was to determine the value of DSC by comparing the change in PROMIS-PF, PROMIS-PI, and PROMIS-D scores of DSC patients with those of ACLR patients. Our hypothesis is that although DSC patients will have worse PROMIS scores preoperatively as compared with ACLR patients, they will derive greater benefit from surgery as compared with ACLR patients.

Patients/methods

Database

We used administrative and medical record data from a tertiary care academic medical center to address the research objective. We used the administrative data to identify the cohorts of interest and to obtain key patient demographics and service details. We used the medical record data to confirm the cohorts and to obtain PROMIS domain scores for the cohorts. Prospective collection of PROMIS scores for all patient visits to the academic orthopedic clinic has been undertaken since March 2015. PROMIS-PF, PROMIS-PI, and PROMIS-D scores are assessed at initial evaluation and follow-up visits.

Patient cohorts

The DSC cohort included patients who underwent elective DSC from 2015 to 2017. One hundred twenty-nine discectomy patients were identified using Current Procedural Terminology code 63030 (posterior extradural laminotomy and excision of herniated intervertebral disk) and were confirmed with medical chart review. Patients indicated for surgery had back pain with radicular pain, a well-correlated magnetic resonance imaging or myelogram finding, and had either failed to respond to 6 weeks of standard conservative care or presented with surgical "red flags" [27]. We further limited the DSC cohort to patients with complete postoperative scores at a minimum of 40 days after and not later than 1-year after surgery. After application of inclusion and exclusion criteria to the original 129 patient sample, 84 DSC patients remained. Surgery was performed by fellowship trained spinal surgeons under general anesthesia. Either microscope or loupe magnification was employed, and the surgery was performed via a unilateral open exposure with minimal laminotomy, fragment removal, and exploration of the disc space for residual fragments.

To study the benefit that patients derive from DSC, it was essential to compare the impact of the surgery on PROMIS scores in the DSC cohort with a control group of patients. Using a hospital-based control group that was similar in observed characteristics to the treatment group (the DSC cohort) but whose disease process was unrelated to the treatment group, would enable us to determine the impact of the surgery. For this purpose and based on reasons presented in the Introduction, we chose patients who underwent ACLR to serve as controls for the DSC cohort.

The ACLR cohort was derived from the database of a separate prospective study assessing the utility of PROMIS in monitoring recovery from ACLR [26]. In that study, skeletally mature patients scheduled to undergo arthroscopic bone-tendon-bone autograft ACLR surgery from 2015 to 2016 were recruited and enrolled preoperatively. These patients were monitored longitudinally with PROMIS scores. The ACLR cohort was further limited to patients over the age of 13 years with postoperative scores at a minimum of 133 days after and not later than 1-year after surgery and excluded patients with balance disorders, degenerative gait disorders, neuromuscular disorders, dementia, depression, or cognitive impairment. This cohort included 210 ACLR patients. The University's Institutional Review Board approved this study (RSRB protocol #37409) No external funds were received in support of this work, and the authors have no relevant conflict of interest disclosures.

Outcomes

The outcomes of interest were PROMIS-PF, PROMIS-PI, and PROMIS-D scores. The PROMIS CAT instruments were accessed at the Assessment Center (www.assessmentcenter.net), including PROMIS-PF (version 1.2), PROMIS-PI (version 1.1), and PROMIS-D (version 1.0). The PROMIS CAT algorithm results in t-scores standardized to a normative US population. The mean score is 50 representing the average in the population and the standard deviation is 10 points. It is notable that a higher PROMIS-PF score signifies increased function, whereas higher PROMIS-PI and PROMIS-D scores signify increased pain interference and depression, respectively. These scores were specified as continuous variables for the analysis.

Lumbar discectomy patients often return to their spine surgeon for a final postoperative visit at 6 weeks, at which time they are often discharged to *return as needed* status. ACLR patients typically have routine follow-up visits through at least 6 months in order for physicians to approve return to sport [28]. PROMIS has been validated to detect significant change at 20-week follow-up for the ACLR patient population [26]. Consequently, this time point was used as the final ACLR postoperative reference time point. Study data for analysis was collected from the final preoperative visit and the postoperative clinic visit closest to 40 days for DSC patients and 133 days for ACLR patients. Thus, for each patient in the study cohort and for each PROMIS domain, we use the baseline/preoperative PROMIS score and the postoperative PROMIS score (measured at the time interval specified above) to study how these scores changed with surgery for each of the two cohorts.

Independent variables

The key explanatory variables were the surgical cohort, the surgical phase, and an interaction between these two variables. The surgical cohort was specified as a binary variable representing the intervention-specific (DSC/ACLR) cohort. The surgical phase was specified as a binary variable representing the preoperative and postoperative stages. The interaction between the surgical cohort and phase was used to analyze whether the change in the PROMIS domain scores with the respective surgeries differed between the DSC and ACLR cohorts. We used the interaction term to examine changes in the PROMIS scores among DSC patients by comparing the postoperative scores to their baseline values (within the DSC cohort) and by comparing the change in scores to that in the ACLR cohort (between the DSC and ACLR cohorts). As an example, for the PROMIS-PF outcome, by using the interaction term, we account for (1) the baseline PROMIS-PF score, the postoperative score, and the change in this score for each of the two cohorts, and (2) compare the change in this score in the DSC cohort with that in the ACLR cohort. By doing this, we can draw inferences about DSC and ACLR patients with similar preoperative PROMIS-PF scores and how these scores change with their surgeries.

We used the following variables for generating propensity scores to account for the differences in the baseline characteristics of the DSC and ACLR cohorts. We used patient age which was specified as a continuous variable, the race was specified as a categorical variable (White, Black, or Other), and the sex was specified as a categorical variable. The number of preoperative comorbidities was identified using the Elixhauser Comorbidity Index [29], and specified as a continuous variable.

Statistical analysis

We used propensity score matching to generate matched groups of DSC and ACLR patients that had similar distributions of patient's age, sex, race, and number of comorbidities preoperatively [30]. For this, we generated propensity scores using Stata's −pscore- command, we checked for the balance of the scores across the DSC and ACLR cohorts within blocks of the propensity score, and we used the nearest neighbor 1:1 matching without replacement [31,32]. For descriptive analysis, we used chi-square tests and nonparametric Kruskal-Wallis tests to compare the distribution of outcomes and characteristics between the two cohorts. We constructed linear regression models with interactions between the matched cohorts and pre-/postoperative phase to study the change in the outcomes scores for the DSC patients in comparison to their baseline scores and the scores for the ACLR cohort. All regression models controlled for the year of surgery and for the clustering of observations within patients and providers. In these regression models, we did not control for patient demographics since these were used in calculating the propensity scores. We further calculated adjusted marginal estimates to determine the adjusted pre- and postoperative PROMIS scores for the two cohorts. The level of significance was set at 0.05. All data management and analyses were performed using Stata statistical software, version 15 (StataCorp. LCC, College Station, TX, USA).

To check for the robustness of the findings, we used another propensity score method—Inverse Probability of Treatment Weighting (IPTW)—to generate weights that are the inverse probability of being in a particular surgical cohort [32,33]. For this, patients in the DSC group were assigned a weight that was the inverse of the propensity score, and patients in the ACLR group were assigned weight that was the inverse of one minus the propensity score. The weights were normalized to one. These weights were included in the regression models which were otherwise similar in specification to the main/matched models. We supplemented findings from the propensity score matching analysis with the IPTW to check for the consistency of the findings from the main analysis. Moreover, the IPTW has the advantage of retaining most observations from the cohort.

Results

Patient characteristics

Table 1 presents the descriptive characteristics for 84 patients in each of the matched DSC and ACLR cohorts. The DSC cohort was 51.19 % female (N=43), 83.33% white (N=70), 9.52% black (N=8), had a mean age of 40.34 years (standard deviation [SD]=13.71) and mean number of comorbidities identified using Elixhauser's algorithm was 0.58 (SD=0.81). There were no significant differences in patient characteristics between the DSC and ACLR cohorts. Appendix Table 1 presents descriptive statistics for all patients in the two cohorts before propensity score matching, and Appendix Table 2 presents the descriptive statistics for patients included in the IPTW models.

Unadjusted perioperative PROMIS scores

The Figure presents the unadjusted perioperative PROMIS scores. In both the pre- and postoperative phases, the mean PROMIS scores for the DSC patients indicated worse outcomes than the ACLR cohort. In comparison to ACLR patients, DSC patients exhibited worse preoperative (34.26 vs. 42.43, p<.001) and postoperative (43.34 vs. 47.67, p<.001) PROMIS-PF scores indicative of lower physical function. For PROMIS-PI, in comparison to ACLR patients, DSC patients had higher preoperative (66.66 vs. 55.38, p<.001) and postoperative (56.94 vs. 50.53, p<.001) scores indicative of greater pain. Finally, DSC patients recorded higher PROMIS-D scores in the preoperative (54.03 vs. 46.47, p<.001) and postoperative (46.20 vs. 42.45) phases.

Table 1
Patient characteristics in propensity score matched cohorts

	DSC	ACLR	*p Value*
Number of patients	84	84	
Female: N (%)	43 (51.19%)	41 (48.80%)	.76
Race: N (%)			
White	70 (83.33%)	73 (86.90%)	.80
Black	8 (9.52%)	6 (7.14%)	
Other	6 (7.14%)	5 (5.95%)	
Age in years: Mean (SD)	40.34 (13.71)	36.70 (12.82)	.08
Number of comorbidities: Mean (SD)	0.58 (0.81)	0.42 (0.68)	.15

DSC, lumbar discectomy; ACLR, anterior cruciate ligament reconstruction; N (%), number of patients (column percentage); mean (SD), Mean (Standard deviation).

Comparing matched postoperative outcomes

Table 2 presents the estimates for the interaction term from the regression models that examine the impact of the surgery on the PROMIS domain scores in the two cohorts and the extent to which this impact varies between the two cohorts. After controlling for relevant confounders and in comparison to the baseline scores, DSC patients experienced a greater increase in PROMIS-PF scores as compared with ACLR patients (Adjusted estimate of interaction term: 3.84; 95% Confidence Interval, CI 1.08−6.60; p=.01). Moreover, DSC patients also experienced greater decline in PROMIS-PI (Adjusted estimate: −4.87; 95% CI: −7.52 to −2.23; p<.001) and greater decline in PROMIS-D scores (Adjusted estimate: −2.95; 95% CI −5.70 to −0.21; p=.04) with surgery. Appendix Table 3 present regression estimates from models using the IPTW strategy. The inferences from these models are consistent with the findings from the main analysis.

Marginal estimates comparing surgical impact

Table 3 presents the adjusted marginal estimates of the perioperative score changes from the regression models. PROMIS-PF perioperative score difference was 9.09 for the DSC group and 5.25 for the ACLR group, indicating a greater increase in physical function in DSC patients in comparison to ACLR patients after controlling for preoperative scores. PROMIS-PI perioperative score differences also showed DSC patients obtained a greater reduction in pain interference after surgery (−11.29 vs. −6.41). Lastly, PROMIS-D scores displayed that DSC patients receive a greater decline in depression after surgery (−7.82 vs. −4.88). Similar inferences were obtained from the marginal estimates following the IPTW analysis.

Discussion

As society begins making difficult choices about the allocation of resources for medical and surgical care, it is incumbent upon surgeons to document the value they provide for their patients. Although there are studies quantifying and evaluating the improvement patients experience after specific orthopedic surgeries, there are few comparing the relative benefit or value patients perceive from different procedures. To date, the most common method for

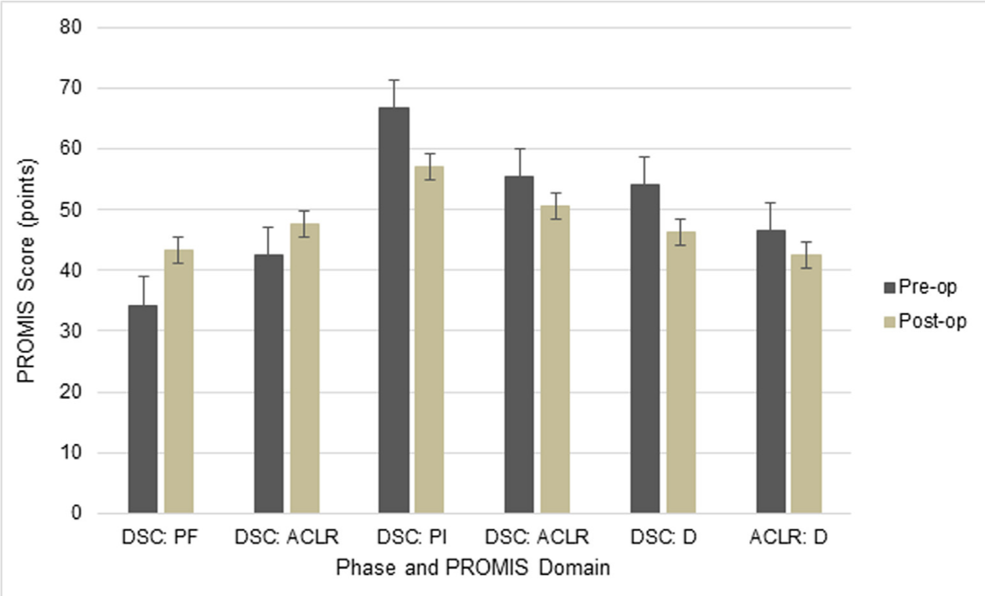

Fig. Unadjusted PROMIS scores before and after surgery of DSC and ACLR cohorts are displayed in the graph. Significant differences (p<.001) were discovered between cohorts for all pre- and postoperative PROMIS scores across physical function, pain interference, and depression domains. PROMIS, Patient-reported outcomes information systems; DSC, lumbar discectomy; ACLR, anterior cruciate ligament reconstruction; PF, physical function; PI, pain interference; D, depression.

Table 2
Propensity score matched estimates from the multivariate models examining the variation in the surgical impact between the two cohorts

Domain	Interaction term estimate	95% CI	p Value
PROMIS-PF	3.84	[1.08, 6.60]	.01
PROMIS-PI	−4.87	[−7.52, −2.23]	<.001
PROMIS-D	−2.95	[−5.70, −0.21]	.04

N, number of patients; CI, confidence interval; PROMIS, patient-reported outcomes information system; PF, physical function; PI, pain interference; D, depression.

Notes: The interaction term estimates are to be interpreted as follows. For example, for PROMIS-PF, after controlling for relevant confounders through propensity score matching and for the pre-operative scores, the mean PROMIS-PF score for the lumbar discectomy cohort was 3.84 points higher than that for the anterior cruciate ligament reconstruction cohort.

comparing the relative value of two surgeries is the Resource Based Relative Value Scale, a method whose design and implementation has been controversial [34,35]. Studies such as ours are another lens through which physicians can evaluate surgical interventions.

Previous studies comparing spine and other orthopedic procedures showed a lack of statistical difference between spine surgery outcomes and total hip or knee replacement, however, many of these studies have utilized the Standard Form-36 (SF-36) or SF-12 [36−38]. Recent work shows that PROMIS outperforms both the SF-36 and SF-12 [35−39]. PROMIS is both more efficient and has lower floor and ceiling effects [39−41]. Another advantage of the PROMIS is that it can be used to compare patient reported outcomes across differing disease states and different surgical procedures [24,42]. This aspect of the instrument is important given the obvious clinical differences between patients undergoing ACLR and DSC. The present study is the first, that we are aware of, that compares procedural outcomes between orthopedic pathologies using PROMIS.

Prior research has shown that ACLR patients report a good quality of life after surgery, even when compared with a matched sample of people from the general population. A systematic review of ACLR literature showed that 65% of patients returned to their preinjury level of sport [13,43]. In contradistinction, up to 40% of spine patients, experience persistent symptoms [4−8]. Our data shows that the ACLR cohort achieved preferable postoperative outcomes with lower absolute postoperative scores, but our data also shows that DSC patients report a larger incremental benefit from the surgery. This finding differs from previous work, and the differential impact of the DSC procedure is interesting in the context of surgeon's efforts at value creation [28,32].

Additionally, the results demonstrate that DSC patients experience greater improvement in physical function, pain interference, and depression levels after surgery when compared with a matched cohort of ACLR patients. Previous studies such as the Spine Patient Outcomes Research Trial (SPORT) and the Maine Lumbar Spine Study demonstrate the similarly large improvements in spine patients [3,44]. However, these studies do not compare spine patients to those with other musculoskeletal complaints. The present study found the magnitude of change after surgical intervention was, on average, larger in the DSC cohort than the ACLR cohort in all investigated PROMIS domains.

Table 3
Adjusted marginal estimates of the perioperative PROMIS scores

PROMIS Scores	PSM analysis			IPTW analysis		
	DSC cohort	ACLR cohort	Difference	DSC cohort	ACLR cohort	Difference
PROMIS-PF						
Preoperative	34.63	42.05	−7.42	34.58	40.62	−6.04
Postoperative	43.72	47.30	−3.58	45.62	48.21	−2.59
Perioperative difference	9.09	5.25	3.84	11.04	7.59	3.45
PROMIS-PI						
Preoperative	66.36	57.24	9.12	66.24	55.56	10.68
Postoperative	55.07	50.83	4.24	53.25	46.91	6.34
Perioperative Difference	−11.29	−6.41	−4.88	−12.99	−8.65	−4.34
PROMIS-D						
Preoperative	53.68	46.60	7.08	53.02	45.73	7.29
Postoperative	45.86	41.72	4.14	44.43	40.00	4.43
Perioperative difference	−7.82	−4.88	−2.94	−8.59	−5.73	−2.86

PROMIS, patient-reported outcomes information systems; DSC, lumbar discectomy; ACLR, anterior cruciate ligament reconstruction; PSM, propensity score matching; IPTW, inverse probability of treatment weighting; PF, physical function, PI, pain interference, D, depression.

Although the complete analysis of any procedure requires long-term follow-up, the immediate surgical result certainly influences surgeon and patient opinion. Few studies exist to compare the relatively early benefits patients perceive from two different procedures.

The results of this study should be interpreted with consideration of certain limitations. The limitations include: selection bias, the short-term follow-up, and a potential bias of comparing a prospective trial data set with a retrospective analysis of prospective clinical data.

The observational cohorts were developed from the population of patients receiving ongoing care in our clinics. The aim was to compare PRO-reported improvement between two different orthopedic surgeries as they are carried out in actual practice. Consequently, the time point at which final PRO were collected differs between the two procedures; however, the time points which were chosen are both typical final follow-up timepoints in the context of the specific surgery being assessed. We cannot comment on outcomes after the study's final follow-up point (40 days for DSC, 133 days for ACLR). It is possible that longer follow-up intervals would have shown different outcomes in the cohorts. However, because our practice is to discharge DSC patients at approximately 40 days to *return as needed* status, mandating a longer follow-up interval may have skewed data by excluding patients with excellent outcomes and biasing the population toward those patients with continued follow-up, potentially indicative of inferior clinical results. Additionally, as the DSC population was collected prospectively in the normal course of our clinic activities, requiring longer term follow-up of patients who were pleased with their outcome would have imposed an unreasonable expense of both time and money on those patients. Finally, the work of Papuga et al. showed that the 20-week follow-up interval after ACLR had excellent correlation with the ultimate outcome [26].

Since DSC patients had worse preoperative scores than ACLR patients, it is possible that the differential DSC score improvement is the result of ceiling effects of the PRO instrument for the ACLR cohort. It is our hope that the use of PROMIS helps mitigate these potential ceiling effects. PROMIS-PF CAT has demonstrated 0.81% and 1.4% ceiling effects in spine and ACLR patients, respectively [39,45]. Bhatt et al. have demonstrated similar low ceiling effects in a population of lumbar discectomy patients for the PROMIS PI. [41] Floor and ceiling effects for PROMIS-PI in the ACLR cohort have not been studied; however, a review of our ACLR data set showed a good distribution of scores suggesting low floor and ceiling effects (data not shown). Floor and Ceiling effect analysis is not available for PROMIS-D in either of these populations. Consistently favorable floor and ceiling effect analyses in other PROMIS studies suggests instrument validity and sensitivity to accurately capture results in our study [24,26,39−41,46,47].

Because the ACLR cohort was enrolled as part of a prospective trial, and the DSC was the result of retrospective review of prospectively collected clinical data, there is the possibility that the systematic differences between the two cohorts may bias our results. However, the process of PROMIS collection in our clinics is uniform and identical, and this combined with the propensity score matching should serve to minimize that effect.

This study shows that in an observation cohort study based upon propensity score matching, ACLR patients have preferable postoperative depression, function, and pain interference scores in comparison with DSC patients; however, DSC patients receive a greater incremental benefit from the surgery. This study provides a novel approach to assessing the relative value that patients receive from two different orthopedic surgeries and as such is a platform for further analysis, discussion, and future research.

Acknowledgment

No external funds or support were used for this study.

Appendix

Appendix Table 1. Table of patient characteristics before propensity score matching

Appendix Table 2. Table of patient characteristics in IPTW Matched Cohorts

Appendix Table 3. Table that illustrates the multivariate regression analysis estimates for age, sex, race and Elixhauser comorbidity index by PROMIS domain

Appendix Table 1
Pre-PSM patient characteristics by surgical intervention

	DSC	ACLR	p
Number of patients	88	210	
Age (years)	40.14 (13.64)	26.84 (12.54)	<.001
Gender (% female)	42 (52.27%)	105 (50%)	.72
Race			
White	73 (82.95%)	185 (88.10%)	.71
Black	8 (9.09%)	14 (6.67%)	
Other	5 (5.68%)	11 (5.24%)	
Number of comorbidities	0.61 (0.65)	0.2 (0.50)	<.001
Follow-up (days)	89.69 (44.86)	169.36 (32.64)	<.001

Values are reported as number (percentage) or mean (standard deviation).

PSM indicated propensity score matching; DSC, lumbar discectomy; ACLR, anterior cruciate ligament reconstruction; PROMIS, patient-reported outcomes measurement information system.

Appendix Table 2
Patient characteristics in IPTW matched cohorts

	DSC	ACLR (IPTW)	p Value
Number of patients	84	185	
Age (years)	40.34 (13.71)	27.05 (12.60)	<.001
Gender (% female)	43 (51.19%)	91(49.19%)	.76
Race (%)			
White:	70 (83.33%)	162 (86.48%)	.65
Black:	8 (9.52%)	13 (7.03%)	
Other:	6 (7.14%)	10 (5.41)	
Number of comorbidities	0.58 (0.81)	0.21 (0.51)	<.001

Values are reported as number (percentage) or mean (standard deviation).

DSC indicates lumbar discectomy; ACLR, anterior cruciate ligament reconstruction.

See patient characteristics by PSM cohorts in Table 1.

References

[1] Weinstein JN, Tosteson TD, Lurie JD, Tosteson AN, Hanscom B, Skinner JS, et al. Surgical vs nonoperative treatment for lumbar disk herniation: the Spine Patient Outcomes Research Trial (SPORT): a randomized trial. JAMA 2006;296:2441–50.

[2] Weinstein JN, Lurie JD, Tosteson TD, Skinner JS, Hanscom B, Tosteson AN, et al. Surgical vs nonoperative treatment for lumbar disk herniation: the Spine Patient Outcomes Research Trial (SPORT) observational cohort. JAMA 2006;296:2451–9.

[3] Atlas SJ, Deyo RA, Keller RB, Chapin AM, Patrick DL, Long JM, et al. The Maine Lumbar Spine Study, Part II. 1-year outcomes of surgical and nonsurgical management of sciatica. Spine (Phila Pa 1976) 1996;21:1777–86.

[4] Dohrmann GJ, Mansour N. Long-term results of various operations for lumbar disc herniation: analysis of over 39,000 patients. Med Princ Pract 2015;24:285–90.

[5] Jansson KA, Nemeth G, Granath F, Jonsson B, Blomqvist P. Health-related quality of life (EQ-5D) before and one year after surgery for lumbar spinal stenosis. J Bone Joint Surg 2009;91:210–6.

[6] Malmivaara A, Slatis P, Heliovaara M, Sainio P, Kinnunene H, Kaukare J, et al. Surgical or nonoperative treatment for lumbar spinal stenosis? A randomized controlled trial. Spine (Phila Pa 1976) 2007;32:1–8.

[7] Mannion AF, Denzler R, Dvorak J, Grob D. Five-year outcome of surgical decompression of the lumbar spine without fusion. Eur Spine J 2010;19:1883–91.

[8] Martin BI, Mirza SK, Comstock BA, Gray DT, Kreuter W, Deyo RA. Reoperation rates following lumbar spine surgery and the influence of spinal fusion procedures. Spine (Phila Pa 1976) 2007;32:382–7.

[9] Hanley EN Jr., Herkowitz HN, Kirkpatrick JS, Wang JC, Chen MN, Kang JD. Debating the value of spine surgery. J Bone Joint Surg 2010;92:1293–304.

[10] Mall NA, Chalmers PN, Moric M, Tanaka MJ, Cole BJ, Bach BR, et al. Incidence and trends of anterior cruciate ligament reconstruction in the United States. Am J Sports Med 2014;42:2363–70.

[11] Group MK, Spindler KP, Huston LJ, Chagin KM, Kattan MW, Reinke EK, et al. Ten-year outcomes and risk factors after anterior cruciate ligament reconstruction: a MOON longitudinal prospective cohort study. Am J Sports Med 2018;46:815–25.

[12] Hughes G, Watkins J. A risk-factor model for anterior cruciate ligament injury. Sports Med 2006;36:411–28.

[13] Ardern CL, Taylor NF, Feller JA, Webster KE. Fifty-five per cent return to competitive sport following anterior cruciate ligament reconstruction surgery: an updated systematic review and meta-analysis including aspects of physical functioning and contextual factors. Br J Sports Med 2014;48:1543–52.

[14] Cella D, Yount S, Rothrock N, Gershon R, Cook K, Reeve B, 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(5 Suppl 1):S3–11.

Appendix Table 3
Estimates from regression models using inverse probability of treatment weighting

	IPTW analysis		
	PROMIS-PF	*PROMIS-PI*	*PROMIS-D*
Postoperative phase	7.59 *** [6.59, 8.59]	−8.65 *** [−9.57, −7.72]	−5.73 *** [−6.29, −5.18]
DSC cohort	−6.04 *** [−7.34, −4.74]	10.68 *** [9.65, 11.72]	7.29 *** [4.65, 9.93]
Interaction	3.44 ** [1.27, 5.62]	−4.34 *** [−5.52, −3.16]	−2.86 *** [−3.53, −2.19]

Values are reported as regression estimate [95% confidence interval].

PROMIS indicates patient-reported outcomes measurement information system; PSM2, propensity score matching; PF, physical function; PI, pain interference; D, depression; IPTW, inverse probability of treatment weighting.

** indicates p<.01, *** indicates p<.001.

[15] Garcia SF, Cella D, Clauser SB, Flynn KE, lad T, Lai JS, et al. Standardizing patient-reported outcomes assessment in cancer clinical trials: a patient-reported outcomes measurement information system initiative. J Clin Oncol 2007;25:5106–12.
[16] Hung M, Nickisch F, Beals TC, Greene T, Clegg DO, Saltzman CL. New paradigm for patient-reported outcomes assessment in foot & ankle research: computerized adaptive testing. Foot Ankle Int 2012;33:621–6.
[17] Hung M, Stuart AR, Higgins TF, Saltzman CL, Kubiak EN. 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 2014;28:439–43.
[18] Forrest CB, Meltzer LJ, Marcus CL, de la Motte A, Kratchman A, Buysse DJ, et al. Development and validation of the PROMIS pediatric sleep disturbance and sleep-related impairment item banks. Sleep 2018.
[19] Hung M, Stuart A, Cheng C, Hon SD, Spiker R, Lawrence B, et al. Predicting the DRAM mZDI using the PROMIS anxiety and depression. Spine (Phila Pa 1976) 2015;40:179–83.
[20] Kendall R, Wagner B, Brodke D, Bousanga J, Voss M, Gu Y, et al. The relationship of PROMIS pain interference and physical function scales. Pain Med (Malden, Mass) 2018;19:1720–4.
[21] Papuga MO, Mesfin A, Molinari R, Rubery PT. Correlation of PROMIS physical function and pain CAT instruments with Oswestry Disability Index and neck disability index in spine patients. Spine 2016;41:1153–9.
[22] Pilkonis PA, Yu L, Dodds NE, Johnston KL, Lawrence SM, Hilton TF, et al. Item banks for substance use from the Patient-Reported Outcomes Measurement Information System (PROMIS((R))): severity of use and positive appeal of use. Drug Alcohol Depend 2015;156:184–92.
[23] Yeatts KB, Stucky B, Thissen D, Irwin D, Varni JW, DeWitt EM, et al. Construction of the Pediatric Asthma Impact Scale (PAIS) for the Patient-Reported Outcomes Measurement Information System (PROMIS). J Asthma 2010;47:295–302.
[24] Brodke DJ, Saltzman CL, Brodke DS. PROMIS for orthopaedic outcomes measurement. J Am Acad Orthop Surg 2016;24:744–9.
[25] Durkin B, Romeiser J, Shroyer AL, Schiller R, Bae J, Davis RP, et al. Report from a quality assurance program on patients undergoing the MILD procedure. Pain Med 2013;14:650–6.
[26] Papuga MO, Beck CA, Kates SL, Schwarz EM, Maloney MD. Validation of GAITRite and PROMIS as high-throughput physical function outcome measures following ACL reconstruction. J Orthop Res 2014;32:793–801.
[27] Herndon CM, Zoberi KS, Gardner BJ. Common questions about chronic low back pain. Am Fam Physician 2015;91:708–14.
[28] Zaffagnini S, Grassi A, Marcheggiani Muccioli GM, et al. Return to sport after anterior cruciate ligament reconstruction in professional soccer players. Knee 2014;21:731–5.
[29] Elixhauser A, Steiner C, Harris DR, Coffey RM. Comorbidity measures for use with administrative data. Med Care 1998;36:8–27.
[30] Austin PC. An Introduction to propensity score methods for reducing the effects of confounding in observational studies. Multivariate Behav Res 2011;46:399–424.
[31] Garrido MM, Kelley AS, Paris J, Roza K, Meier DE, Morrison RS, et al. Methods for constructing and assessing propensity scores. Health Serv Res 2014;49:1701–20.
[32] Inacio MC, Chen Y, Paxton EW, Namba RS, Kurtz SM, Cafri G. Statistics in brief: an introduction to the use of propensity scores. Clin Orthop Relat Res 2015;473:2722–6.
[33] Austin PC, Stuart EA. Moving towards best practice when using inverse probability of treatment weighting (IPTW) using the propensity score to estimate causal treatment effects in observational studies. Stat Med 2015;34:3661–79.
[34] Orr RD, Sodhi N, Dalton SE, Khlopus A, Sultan AA, Chugtai M, et al. What provides a better value for your time? The use of relative value units to compare posterior segmental instrumentation of vertebral segments. Spine J 2018;18:1727–32.
[35] Baadh A, Peterkin Y, Wegener M, Flug J, Katz D, Hoffmann JC. The relative value unit: history, current use, and controversies. Curr Probl Diagn Radiol 2016;45:128–32.
[36] Mokhtar SA, McCombe PF, Williamson OD, Morgan MK, White GJ, Sears WR. Health-related quality of life: a comparison of outcomes after lumbar fusion for degenerative spondylolisthesis with large joint replacement surgery and population norms. Spine J 2010;10:306–12.
[37] Polly DW Jr., Glassman SD, Schwender JD, Shaffrey CI, Branch C, Burkus JK, et al. SF-36 PCS benefit-cost ratio of lumbar fusion comparison to other surgical interventions: a thought experiment. Spine (Phila Pa 1976) 2007;32(11 Suppl):S20–6.
[38] Rampersaud YR, Ravi B, Lewis SJ, Stas V, Barron R, Davey R, et al. Assessment of health-related quality of life after surgical treatment of focal symptomatic spinal stenosis compared with osteoarthritis of the hip or knee. Spine J 2008;8:296–304.
[39] Brodke DS, Goz V, Voss MW, Lawrence BD, Spiker WR, Hung M. PROMIS PF CAT outperforms the ODI and SF-36 physical function domain in spine patients. Spine (Phila Pa 1976) 2017;42:921–9.
[40] Patel AA, Dodwad SM, Boody BS, Bhatt S, Savage JW, Hsu WK, 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.
[41] Bhatt S, Boody BS, Savage JW, Hsu WK, Rothrock NE, Patel AA. Validation of patient-reported outcomes measurement information system computer adaptive tests in lumbar disk herniation surgery. J Am Acad Orthop Surg 2018;27:95–103.
[42] Cella D, Riley W, Stone A, Rothrock N, Reeve B, Yount S, et al. 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 2010;63:1179–94.
[43] Mansson O, Kartus J, Sernert N. Health-related quality of life after anterior cruciate ligament reconstruction. Knee Surg Sports Traumatol Arthrosc 2011;19:479–87.
[44] Lurie JD, Tosteson TD, Tosteson AN, Zhao W, Morgan TS, Abdu WA, et al. Surgical versus nonoperative treatment for lumbar disc herniation: eight-year results for the spine patient outcomes research trial. Spine 2014;39:3–16.
[45] Scott EJ, Westermann R, Glass NA, Hettrich C, Wolf BR, Bollier MJ. Performance of the PROMIS in patients after anterior cruciate ligament reconstruction. Orthop J Sports Med 2018;6:2325967118774509.
[46] Bernstein DN, Kelly M, Houck JR, Ketz JP, Flemnister AS, DiGiovanni BF, et al. PROMIS pain interference is superior vs numeric pain rating scale for pain assessment in foot and ankle patients. Foot Ankle Int 2018:1071100718803314.
[47] Bernstein DN, Bakhsh W, Papuga MO, Menga EN, Rubery PT, Mesfin A. An evaluation of PROMIS in patients with primary or metastatic spine tumors. Spine (Phila Pa 1976) 2018;44:747–52.