Preoperative Patient Reported Outcomes Measurement Information System Scores Assist in Predicting Early Postoperative Success in Lumbar Discectomy

Paul T. Rubery

Jeff Houck

Addisu Mesfin

Robert Molinari

Mark O. Papuga

Preoperative Patient Reported Outcomes Measurement Information System Scores Assist in Predicting Early Postoperative Success in Lumbar Discectomy

Paul T. Rubery, MD,* Jeff Houck, PT, PhD,† Addisu Mesfin, MD,* Robert Molinari, MD,* and Mark O. Papuga, PhD*,‡

Study Design. Retrospective review of prospective data.

Objective. Determine whether patient reported outcome (PRO) data collected prior to lumbar discectomy predicts achievement of a minimal important difference (MID) after surgery. Compare ability of PRO and clinical information to predict achievement of MID in short term follow-up after discectomy.

Summary of Background Data. We investigated the ability of patient reported outcomes measurement information system (PROMIS) and clinical factors at the preoperative time point to determine patients achieving MID after surgery.

Methods. PROMIS physical function (PF), pain interference (PI), and depression (D) scores were assessed at evaluation and follow-up for consecutive visits between February, 2015 and September, 2017. Patients with preoperative scores within 30 days prior to surgery and with scores 40 days or more after surgery who completed all PROMIS domains were included yielding 78 patients. MIDs were calculated using a distribution-based method. A multivariate logistic regression model was created, and the ability to predict achieving MID for each of the PROMIS domains was assessed. Cut-off values and prognostic probabilities were determined for this model and models combining preoperative PROMIS with clinical data.

Results. Preoperative PROMIS scores modestly predict reaching MID after discectomy (areas under the curve [AUC] of 0.62, 0.68, and 0.76 for PF, PI, and D, respectively). Preoperative cut-off scores show patients who have PF and PI scores more than 2 standard deviations, and D more than 1.5 standard deviations worse-off than population mean are likely to achieve MID. The combination of PROMIS with clinical data was the most powerful predictor of reaching MID with AUCs of 0.87, 0.84, and 0.83 for PF, PI, and D.

Conclusion. PROMIS scores before discectomy modestly predict improvement after surgery. Preoperative PROMIS combined with clinical factors was more predictive of achieving MID than either clinical factors or PROMIS alone.

Key words: herniated lumbar disc, lumbar discectomy, patient reported outcome, patient reported outcomes measurement information system.

Level of Evidence: 3

Spine 2019;44:325–333

L umbar intervertebral disc herniation (HNP) is an important cause of back and leg pain and a common indication for surgery.[1–5] Numerous studies have documented the clinical efficacy of discectomy.[6–14] However, not all patients respond positively to surgery, and consequently careful presurgical selection is recommended.

There is great interest in preoperative factors which predict postoperative success. Historical, demographic, social, imaging, and physical examination parameters have been studied with occasionally conflicting results.[15–24] Factors frequently associated with better outcomes include leg pain predominating over back pain, the presence of sensory, motor, or reflex changes, and the presence of a straight leg raising sign. One factor often associated with undesired outcomes is a workmen's compensation injury.[19,25] The absence of consistent links between objective clinical factors and predictable outcomes, results in surgical selection remaining both an "art and science."

Additional clinical tools to predict postoperative results would be valuable.

The patient reported outcomes measurement information system (PROMIS) is a patient reported outcomes assessment tool (PRO) that was developed with National Institutes of Health funding beginning in 2002.[26] The PROMIS consists of validated instruments which can be utilized to assess health domains reflecting numerous aspects of a patient's life.[27] Because PROMIS uses computer adaptive testing (CAT) based upon item response theory administration time is decreased while accuracy remains high.[28,29] The PROMIS has been shown to correlate well with traditional PRO used in the lumbar and cervical spine such as the Oswestry Disability Index (ODI) and the Neck Disability Index (NDI)[30,31] PROMIS may augment the clinical examination by defining the level of severity on specific health domains.[32–35] PROMIS measures are increasingly being used by clinicians in "real time" during clinic visits to track patient progress, and may help identify patients at risk for poor outcome.[33–38]

The purpose of this study was to: (1) compare pre- and postoperative PROMIS physical function (PF), pain interference (PI), and depression (D) scores in patients undergoing lumbar discectomy. (2) To determine the accuracy of using pre-op PROMIS PF, PI, and D in predicting achievement of a postoperative minimally important difference (MID). (3) Assess the accuracy of history and physical exam findings in the same patients for predicting the achievement of postoperative MID. (4) Assess whether a combination of the traditional data with preoperative PROMIS scores leads to an improved ability to predict the achievement of MID.

MATERIALS AND METHODS

This study was approved by the institutional review board (protocol RSRB #62787).

Prospective patient reported outcomes collection is performed for all consecutive patient visits to our orthopedic clinic utilizing a tablet computer (Apple, iPad). PROMIS PF, PI, and D scores are assessed at initial evaluation and at all follow-up visits. Data collected are immediately available within the electronic health record during the clinic visit. For this study, we undertook a review of all clinic visits occurring from February, 2015 through September, 2017. Patients undergoing elective lumbar discectomy were identified by reviewing billing records for CPT 63030, and by subsequent review of the operative report. There were 855 unique visits across 129 patients. Patients with preoperative scores within 30 days prior to surgery and with postoperative scores at least 40 days after who had completed all PROMIS domains were included; yielding 78 patients. All surgeries were performed by three fellowship-trained orthopedic spine surgeons via mini-open or microscope assisted approach.

The PROMIS CAT instruments were accessed through the Assessment Center (www.assessmentcenter.net), including physical function (PF, version 1.2), pain interference, (PI, version 1.1) and depression (D, version 1.0). The PROMIS CAT algorithm results in t scores standardized to a normative US population. The scores range 0 to 100 points, and the mean score of 50 represents the mean for the United States' population. A standard deviation is defined as 10 points. It is notable that a higher PF score signifies increased function, while higher PI and depression scores signify increased pain interference and depression, respectively. Thus, a patient with a PF score of 40 has physical function, as measured by PROMIS, one standard deviation worse off than the mean for the US population at large.

To compare the preoperative PROMIS scores to postoperative PROMIS scores (change in PROMIS scores), a two-way repeated measures ANOVA analysis was used. Factors were time (pre and postoperation) and PROMIS domain. An interaction effect was consistent with an improvement in all scales. In the presence of an interaction, pairwise comparisons were used to determine differences due to time for each PROMIS scale. An alpha level of 0.05 was used for all analyses.

Quantifying improvement after clinical interventions is challenging. The minimally important difference (MID) refers to the smallest meaningful difference in a condition that has implications for patient care.[39,40] A common definition of MID was applied across all scales by using the distribution-based method. The distribution method defines an MID as half of the standard deviation (SD) of each PROMIS domain. The MIDs calculated for PF, PI, and D were +3.75, −3.89, and −6.56, respectively. We defined early postoperative success as a patient reaching MID.

Receiver operating characteristic (ROC) analysis was used to estimate the best cut-off points for each of the PROMIS domains to predict surgical success. The area under the curve (AUC) was determined for each PROMIS scale as a measure of the accuracy of the test. An AUC of 0.5 signifies accuracy similar to chance, whereas an AUC of 1.0 signifies perfect accuracy. An AUC of 0.7 is considered reasonably accurate, whereas an AUC of 0.8 is considered excellent.[41] To improve clinical utility, cutoff values for each PROMIS domain were selected so that a positive test would predict meeting MID for that domain with 90% specificity. Separate cut-off values were selected so that a positive test would predict failing to meet MID with 90% specificity. Likelihood ratios were calculated for each PROMIS scale for meeting and not meeting MID. To improve clinical application, pre- and post-test probabilities were derived for each PROMIS domain to determine how a positive test at the specified cutoff would change the probability of meeting MID postoperatively. A larger change from pre- to post-test probability suggests higher clinical utility of the preoperative variable. To determine adequate sample size 70% of patients were assumed to achieve a MID improvement based upon previous data analysis. With an allocation ratio of 0.7, and alpha of 0.05 and the ability to distinguish an AUC of 0.675 or higher from an AUC of 0.5 with 80% power a sample of at least 65 participants was calculated.[42]

Chart review was undertaken to assess the presence of demographic, historic, and clinical findings commonly

TABLE 1. Traditional Indicators (Variables With Definition) Used to Construct Logistic Regression Model	
Continuous Variables	**Definition**
Age	Age in years at time of surgery
Gender	Gender
BMI	Body mass index at time of surgery
Dichotomous variables	
Smoking	Patient record indicates current smoker
Worker's compensation	Patient record indicates workmen's compensation injury
Litigation	Patient record indicates litigation in process regarding injury
Sensory deficit	Patient record indicates current dermatomal sensory deficit
Motor deficit	Patient record indicates current myotomal motor deficit
Tension sign	Patient record indicates current tension sign (straight leg raising or bowstring sign)
Reflex abnormality	Patient record indicates anatomically correlated absent reflex
Predominate leg pain	Patient record indicates leg pain predominates over back pain
Complications	Patient required return to OR for a complication during the study period

thought to influence surgical success (Table 1). These included age, sex, body mass index (BMI), and the following dichotomous variables: smoking, the presence of compensable injury or litigation, complications requiring return to surgery, leg pain predominate over back pain, the presence of sensory deficit, motor deficit, reflex changes or tension signs (straight leg raise or Bowstring sign).

Two multivariate logistic regression models were constructed for the sample. One incorporating the clinical factors and one incorporating clinical factors and PROMIS scores. In each model, ROC analysis was used to determine the AUC and the best cut-off value in each domain for predicting the achievement of MID. Likelihood ratios and pre- and post-test probabilities were calculated for each model. MEDCALC software (www.medcalc.org) was used to determine pre- and post-test probabilities and confidence intervals for likelihood ratios. All other analysis was performed using SPSS software (version 23, IBM Corp, Armonk, NY).

RESULTS

Population characteristics are listed in Table 2. The mean age of the study population was 40.1 ± 13.6 years. There were 40 women and 38 men. The mean preoperative evaluation was within 10 ± 7 days before surgery and the mean final postoperative evaluation was 88 ± 44 days after the procedure. The two-way analysis of variance (ANOVA) showed a significant interaction effect ($P < 0.01$) indicating PROMIS scale specific changes over time. Pairwise comparisons indicated that PROMIS PF t-scores improved by 9.7 ± 9.1 ($P < 0.001$), PROMIS PI t scores improved by -11.4 ± 9.8 ($P < 0.001$), and PROMIS D t-scores improved by -8.1 ± 10 ($P < 0.001$) (Table 2).

The AUCs resulting from ROC analysis of the preoperative PROMIS assessments and the two multivariate logistic regression models are presented in Table 3 and Figure 1. The accuracy of the PROMIS alone to predict the achievement of MID for PI (AUC = 0.68) and D (AUC = 0.76) are reasonable, for PF it is only slightly better than chance (AUC = 0.62). In the model incorporating only clinical factors, the AUCs suggest reasonably accurate prediction of achieving MID using PF, PI, and D with AUCs of 0.79, 0.72, and 0.65, respectively. In the model combining the

TABLE 2. Analysis of Population Used to Derive the Cut-Off Values at Baseline (Preoperative) PROs and to Validate the Postoperative Cut-Offs			
Study Population Characteristics			
	Population n = 78, Gender (40F) (38M)		
	Mean	**Standard Deviation**	**Range**
Age	40.1	13.6	16–68
Baseline (days before)	10.0	6.8	1–30
Follow-up (days after)	88.4	44.0	41–178
Baseline (pre-op)			
Physical function	34.3	7.5	15.4–51.2
Pain interference	66.5	7.8	15.4–51.2
Depression	54.4	13.1	31.9–81.8
Follow up (post-op)			
Physical function	44.1	8.7	23.2–67.3
Pain interference	55.1	9.8	32.0–80.1
Depression	46.3	11.0	31.8–75.9
Change baseline to follow-up			
Physical function	+9.7	9.1	−7.9–35.1 {$P < 0.001$}
Pain interference	−11.4	9.8	−36.6–6.1 {$P < 0.001$}
Depression	−8.1	10.0	−42.2–9.5 {$P < 0.001$}

TABLE 3. Receiver Operator Curve Analysis Applied to Derivation Sample (n = 78) Including the Denoted Covariates in Logistic Regression Model

	PF AUC (95% CI) [*P* Value]	PI AUC (95% CI) [*P* Value]	*D* AUC (95% CI) [*P* Value]
PROMIS alone (domain specific)	0.622 (0.477–0.767) [0.090]	0.676 (0.528–0.824) [0.019]	0.757 (0.649–0.866) [<0.001]
Traditional indicators (listed in Table 1)	0.786 (0.675–0.896) [<0.001]	0.721 (0.591–0.851) [0.003]	0.648 (0.527–0.769) [0.027]
PROMIS (PF, PI, *D*) and traditional indicators	0.867 (0.817–0.955) [<0.001]	0.842 (0.735–0.949) [<0.001]	0.834 (0.747–0.922) [<0.001]

preoperative PROMIS with clinical factors the ROC analysis yielded the largest AUCs resulting in PF, PI, and *D* with AUCs of 0.87, 0.84, and 0.83 (Figure 2).

The preoperative cut-off scores for achieving and not achieving MID, the likelihood ratios and the associated pre and post-test probabilities for PROMIS alone, for the "clinical" model and for the "combined" model are provided in Table 4. The post-test probabilities of achieving or failing to achieve MID were substantially higher for all three models indicating an increased accuracy in predicting outcome after incorporating the preoperative PROMIS scores in the analysis. The post-test probabilities of achieving MID were 94%, 95%, and 79% for PF, PI, and *D*, respectively. The post-test probabilities of not achieving MID in the combined model were 74%, 74%, and 90% for PF, PI, and *D*.

DISCUSSION

Predicting which patients will experience desired outcomes after spine surgery remains a challenge. PROMIS (CAT) tools have provided a validated, accurate, and efficient system to document a variety of health domains relevant to spine patients. This data can be graphically displayed and used in the clinic. New from this study is the finding that PRO measures represent a powerful new tool to augment surgical decision-making. This study suggests that combining preoperative PROMIS scores with clinical indicators strengthens the ability to predict which discectomy patients are likely to achieve a clinically important improvement in short-term follow-up.

This study shows that preoperative PROMIS scores alone are somewhat predictive of achievement of MID as measured by PROMIS in short term follow-up after surgery. Preoperative PROMIS scores for PF and PI demonstrated modest accuracy in predicting the ability to reach MID with AUCs of 0.62 and 0.68, respectively. PROMIS depression showed good accuracy with an AUC = 0.76. This predictive accuracy compares favorably with previous studies in the hip, knee, and shoulder arthroplasty populations where
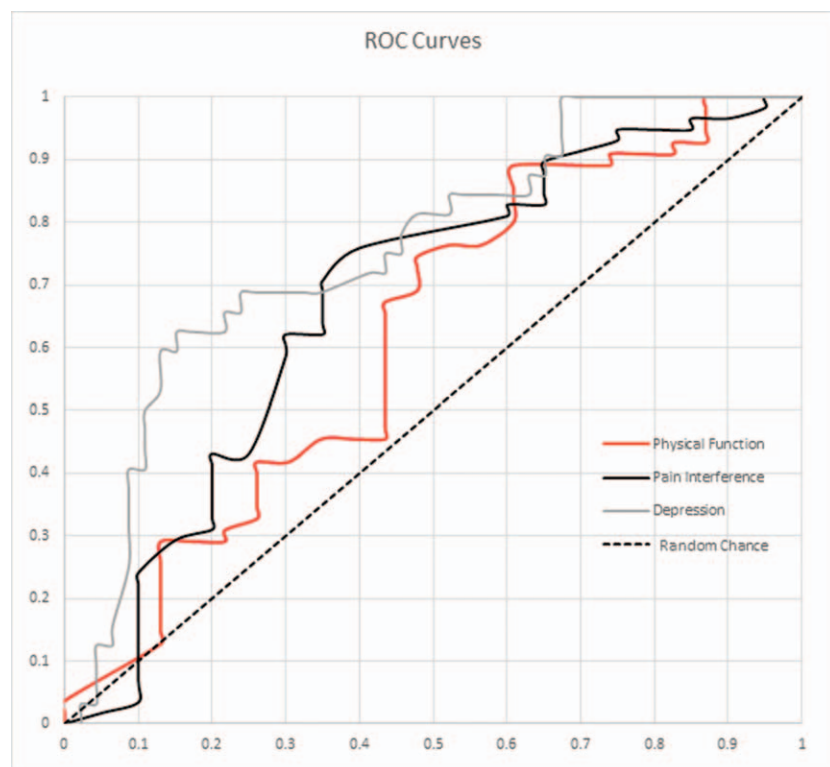

Figure 1. Receiver operating characteristic (ROC) curves depicting the accuracy of preoperative PROMIS PF, PI, and Depression for predicting success and failure of meeting MCID at final follow-up. PF indicates physical function; PI, pain interference; PROMIS, patient reported outcomes measurement information system.

Figure 2. Receiver operating characteristic (ROC) curves depicting the accuracy of preoperative traditional indicators and all three PROMIS domains for predicting success and failure of meeting MCID at final follow-up. PROMIS indicates patient reported outcomes measurement information system.
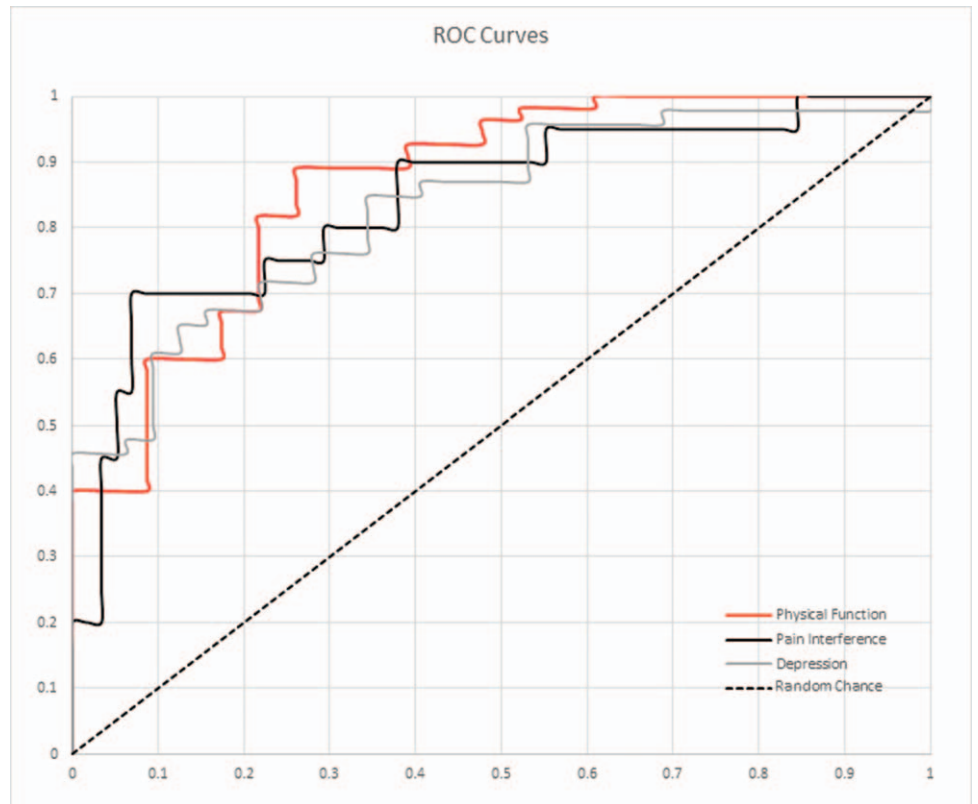
legacy PROs predicted reaching MID with AUCs ranging from 0.6 to 0.83.[33,35,38]

This study demonstrated that lumbar discectomy patients with low function, and high pain interference and depression had a better chance of meeting MID than those with high function and low pain interference and depression. An analysis of the cut-off scores shows that patients approximately two standard deviations or more "worse off" than the US population norm for PF and PI, and 1.5 standard deviations or more "worse off" for depression (*i.e.*, PF *t*-score <30, PI >70, and *D* >65) stand the best chance of reaching MID. Those who are approximately 1 standard deviation or less "worse off" for PF and PI than the US population norm, and those one-half standard deviation or less "worse off" for depression (*i.e.*, PF *t*-score >40, PI <60, and *D* <55) are less likely to reach MID after lumbar discectomy (Table 4). Similar findings using PROMIS scales occurred in ACL reconstruction, hip and knee arthroplasty, shoulder arthroplasty, and foot and ankle surgery.[33–35,37,38] By applying these cut-offs to preoperative decision-making the resulting post-test probabilities demonstrate the potential to decrease unwanted variability in patient outcome. For example, patients with preoperative PROMIS PF, PI, and *D* scores suggestive of not achieving an MID were accurate 55%, 55%, and 70% of the time, respectively. Patients with preoperative PROMIS PF, PI, and *D* scores suggestive of achieving an MID were accurate 100%, 88%, and 76% of the time, respectively (Table 4). In the logistic regression model combining preoperative PROMIS with traditional clinical indicators for our study population, the post-test

probabilities appear even more useful for clinical decision-making. Patients were predicted of achieving MID for PF, PI, and *D* with 94%, 95%, and 79% accuracy, respectively, and predicted of not achieving MID with 74%, 74%, and 90% accuracy for PF, PI, and *D* (Table 4). Although some adjustment to the identified cut off scores may occur when validated against or applied to another set of patients, the strength of the associations in the data is strong (AUCs 0.83–0.87). Therefore, the preoperative PROMIS cut-off scores we report are likely useful in decision-making regarding lumbar discectomy for many patients.

It is interesting that for this population, a ROC analysis of a logistic regression model constructed using clinical data showed that typical "clinical indicators" of potential surgical success are more accurately predictive than the PROMIS domains alone. Previous authors have reached mixed conclusions regarding the prognostic value of preoperative clinical factors as predictors of superior outcomes.[16,43] Utilizing SPORT, Koerner determined that patients with body mass index less than 30, no depression, no litigation, not injured at work, and no diabetes seemed to have improved outcomes.[18] Hebert identified a clinical history of leg pain greater than back pain, physical findings including sensory, motor or reflex deficits, and a positive straight leg raise predicted better outcomes.[25] Wilson *et al*[19] in a systematic review confirmed that the absence of work injury and leg pain greater than back pain were associated with positive outcomes; however, their study did not confirm preoperative motor deficit as a predictor of positive outcome. Our study shows that a logistic regression model

TABLE 4. Summary of Cutoff Values, Likelihood Ratios and Pre and Post-Test Probabilities for all Three Study Scenarios (PROMIS Alone, Traditional Clinical Indicators, and Clinical Indicators Combined With PROMIS) for those Achieving and Failing to Achieve MID

		MID	Achieving MID					NOT Achieving MID				
			Cutoff Achieved MID	Positive Likelihood Ratio (95% CI)	Pre-Test Probability of Achieving MID (%)	Post-Test Probability of Achieving MID (%)	Change in Pre-test Probability (%)	Cut off for Not Achieved MID	Positive Likelihood Ratio (95% CI)	Pre-Test Probability of Not Achieving MID (%)	Post-Test Probability of Not Achieving MID (%)	Change in Pre-Test Probability (%)
Cutoff values and likelihood ratios using (PROMIS-alone) cutoffs of those achieving and not achieving MID (using 90% specificity)	Physical Function	+3.75	23.3	*	71	100 (n=2)	29	41.8	2.9 (1.0–8.5)	29	55 (n=11)	26
	Pain Interference	−3.89	73.8	2.4 (0.6–9.7)	74	88 (n=16)	14	58.4	3.5 (1.2–10.2)	26	55 (n=11)	29
	Depression	−6.56	65.6	4.7 (1.7–13.0)	41	76 (n=17)	35	44.3	1.6 (1.0–2.8)	59	70 (n=37)	11
Values using multiple regression (traditional indicators) cut offs of those achieving and not achieving a MID (using 90% specificity)	Physical Function	+3.75	0.84	4.8 (1.2–18.8)	71	92 (n=25)	21	0.546	5.7 (2.3–14.4)	29	71 (n=17)	42
	Pain interference	−3.89	0.836	4.0 (1.0–15.3)	74	92 (n=25)	18	0.637	5.2 (2.0–13.7)	26	64 (n=14)	38
	Depression	−6.56	0.56	2.5 (0.8–7.9)	41	64 (n=11)	23	0.318	3.3 (1.0–10.4)	59	82 (n=17)	23
Values using multiple regression (traditional indicators and All three PROMIS domains) cut offs of those achieving and not achieving a MID (using 90% specificity)	Physical function	+3.75	0.817	6.9 (1.8–26.4)	71	94 (n=35)	23	0.498	6.7 (2.7–16.4)	29	74 (n=19)	45
	Pain interference	−3.89	0.798	6.2 (1.6–23.5)	74	95 (n=38)	21	0.595	8.1 (3.4–19.7)	26	74 (n=19)	48
	Depression	−6.56	0.614	5.4 (2.0–14.8)	41	79 (n=19)	38	0.331	6.5 (2.2–19.5)	59	90 (n=31)	31

MID indicates minimal important difference; PROMIS, patient reported outcomes measurement information system.

*No false positive results occurred in this subgroup, therefore likelihood ratio cannot be calculated.

incorporating these traditional clinical data demonstrated "good" accuracy for predicting MID with AUCs of 0.79, 0.72, and 0.65 for PF, PI, and D, respectively.[41] More importantly we show combining clinical data with preoperative PROMIS scores produce the most accurate predictions (*e.g.*, AUCs 0.83–0.87) (Table 3).

The fact that the study population was derived from an academic orthopedic clinic limits the generalizability of our results. However, chart review confirmed the presentation, clinical exam, imaging results, and operative procedure ensuring that the study population was representative of those patients electing surgical treatment for symptomatic herniated nucleus pulposus in our practice. Our patients had substantial physical function limitations and pain interference in comparison with the normal population. The average t-scores for the PROMIS PF and PI at baseline were 34.3 ± 7.5 and 66.5 ± 7.8, respectively. This indicates low function and high pain interference. The mean PROMIS depression t-score was 54.4 ± 13 suggesting that our population is modestly more depressed than the US population as a whole. An important future study would be to confirm our findings in a representative nation-wide population.

Patients had statistically significant improvement in all three PROMIS domains after surgery; however, the mean follow-up was only 12.6 weeks. Furthermore, we assessed preoperative scores based on PROMIS completed within 30 days of the surgery raising the issue of how close to the actual surgery a patient had completed the preoperative assessment. This study is based upon the actual clinical experience of patients seeking care in our clinic. Consequently, the preoperative assessments were collected at the visit where the patient was offered the option of discectomy. While short follow-up is a weakness, it results from the return as needed practice employed by surgeons beginning 6 weeks after discectomy. We believe it is a common practice to discharge lumbar discectomy patients who are doing well 6 weeks after surgery. Therefore, this study applies to patients who present for evaluation, and follow-up typically after discectomy. However, the results do not address whether the patients maintain MID, nor does it address the long-term efficacy of surgery for HNP. In our population, adding preoperative PROMIS assessments to the surgical decision-making process improved the accuracy of predicting the achievement of MID at early postoperative time points. However, further studies with longer follow-up periods would be of great value in not only documenting the true efficacy of surgical interventions, but also potentially identifying factors correlated with the development of chronic back complaints.

While our observation that patients with greater degrees of presurgical limitation in physical functioning and higher degrees of pain interference are more likely to experience postoperative improvement is somewhat intuitive, the observation that patients with greater degrees of depression are more likely to achieve a MID in depression than the less depressed is not. The present study does not explain this finding and invites future research. It is possible that a portion of the depression measured by PROMIS is reactive and arises from the pain and limitation of the HNP as opposed to intrinsic depression. Consequently, presumed relief of these limitations after surgery leads to a significant lessening of the depression. Further studies will be required to clarify this finding.

Another limitation is the small sample size. The calculation of cut-off scores and likelihood ratios in the "clinical" and "combined" models is somewhat abstract. Many of the patients achieved MID, and a larger sample would introduce additional variation needed to validate the multivariate models. Nevertheless, the reported accuracy is sufficiently high to suggest that the preoperative PROMIS cut-off scores are likely useful in decision-making regarding lumbar discectomy.

Questions can be raised concerning the choice of a distribution-based method for calculating MID. The calculation of MID can be accomplished via two methods: distribution-based and anchor-based. Anchor-based approaches account for patients' judgments of improvement and the distribution-based method does not; however, the anchor-based calculation must be uniquely determined for the condition or population being studied. No anchor-based MID determinations have been published for lumbar discectomy patients. Chen *et al*[44] recently used a triangulation technique to estimate MID for the PROMIS PI scale using data from three randomized controlled trials, two of which involved patients with low back pain. They estimated the MID for PI using distribution and anchor-based techniques in back pain patients to be between 2.0 and 4.0, a value very similar to that we utilized. Distribution-based MID values in our study were calculated with a method that has been shown to yield MID values for PF, PI, and D that are comparable to those previously calculated with an anchor-based method in both chronic back pain and cancer patients.[45,46] Until an anchor-based MID is reported for lumbar HNP patients, we believe that the distribution-based calculation of MID is the best available option.

Lumbar discectomy is undertaken to decompress lumbar nerve roots. The shape, size, morphology, and location of the herniated material can influence the patient's symptoms and potentially the outcome.[47] Billing records were utilized to identify patients for this study. In our practice, CPT 63030 includes HNP found in the central and posterolateral zones, but not other regions. Further detail regarding the HNPs was not recorded, and it is possible that there are subsets with unique anatomic abnormalities. Therefore, a subgroup analysis may show different results.

In conclusion, PROMIS physical function, pain interference, and depression scores obtained prior to lumbar discectomy were predictive of the patient's achievement of a minimal important difference after surgery. In our logistic regression model, traditional clinical data considered to augur success, was more predictive of MID than the PROMIS scores alone. However, the combination of preoperative PROMIS scores with traditional clinical data yielded a model which was the most accurate in predicting

achievement of MID. We routinely review the preoperative PROMIS results during surgical consultations. Patients with "reasonable" clinical presentations for potential discectomy surgery and whose preoperative PROMIS t-scores for physical function and pain interference are approximately two standard deviations worse off than the US population mean, and those whose PROMIS depression t-scores are 1.5 standard deviations worse-off (*i.e.*, PF <30, PI >70, D >65) are highly likely to achieve a MID. Those with PROMIS PF and PI t-scores one standard deviation, and D scores one-half standard deviation worse off than the US population norm (*i.e.*, PF >40, PI <60, D <55) are less likely to achieve a MID after discectomy This study provides surgeons an additional tool to use during preoperative consultations and in shared decision-making regarding lumbar discectomy surgery.

> ## Key Points

❑ PROMIS is a patient reported outcomes assessment tool that can be administered in a spine surgery practice.

❑ PROMIS physical function, pain interference, and depression scores obtained before a lumbar discectomy can be used to predict the probability of a patient reaching a minimal important difference after the surgery as determined by the PROMIS assessment.

❑ Patients with PROMIS PF and PI t-scores that are more than two standard deviations, and those with D t-scores more than 1.5 standard deviations worse than the US population mean are likely to achieve MID after discectomy.

❑ Patients whose PF and PI t-scores are less than one standard deviation, and whose D t-scores are within one-half standard deviation are unlikely to achieve a MID after lumbar discectomy.

❑ In this study, a logistic regression model combining preoperative PROMIS scores with clinical history and exam findings had the greatest predictive accuracy for the achievement of postoperative MID.

References

1. Awad JN, Moskovich R. Lumbar disc herniations: surgical versus nonsurgical treatment. *Clin Orthop Relat Res* 2006;443:183–97.
2. Jordan J, Konstantinou K, O'Dowd J. Herniated lumbar disc. *BMJ Clin Evid* 2011;2011:pii: 1118.
3. Truumees E. A history of lumbar disc herniation from Hippocrates to the 1990s. *Clin Orthop Relat Res* 2015;473:1885–95.
4. Schoenfeld AJ. Historical contributions from the Harvard system to adult spine surgery. *Spine (Phila Pa 1976)* 2011;36:E1477–84.
5. Best MJ, Buller LT, Eismont FJ. National trends in ambulatory surgery for intervertebral disc disorders and spinal stenosis: a 12-year analysis of the National Surveys of Ambulatory Surgery. *Spine (Phila Pa 1976)* 2015;40:1703–11.
6. Moore AJ, Chilton JD, Uttley D. Long-term results of microlumbar discectomy. *Br J Neurosurg* 1994;8:319–26.
7. Loupasis GA, Stamos K, Katonis PG, et al. Seven- to 20-year outcome of lumbar discectomy. *Spine (Phila Pa 1976)* 1999;24: 2313–7.
8. Atlas SJ, Keller RB, Wu YA, et al. Long-term outcomes of surgical and nonsurgical management of sciatica secondary to a lumbar disc herniation: 10 year results from the maine lumbar spine study. *Spine (Phila Pa 1976)* 2005;30:927–35.
9. van Tulder MW, Koes B, Seitsalo S, et al. Outcome of invasive treatment modalities on back pain and sciatica: an evidence-based review. *Eur Spine J* 2006;15 (suppl):S82–92.
10. Gibson JN, Waddell G. Surgical interventions for lumbar disc prolapse. *Cochrane Database Syst Rev* 2007;CD001350; Review. Update in: Cochrane Database Syst Rev. 2007;(2):CD001350.
11. Anderson PA, McCormick PC, Angevine PD. Randomized controlled trials of the treatment of lumbar disk herniation: 1983-2007. *J Am Acad Orthop Surg* 2008;16:566–73.
12. Jacobs WC, van Tulder M, Arts M, et al. Surgery versus conservative management of sciatica due to a lumbar herniated disc: a systematic review. *Eur Spine J* 2011;20:513–22.
13. Lurie JD, Tosteson TD, Tosteson AN, et al. Surgical versus nonoperative treatment for lumbar disc herniation: eight-year results for the spine patient outcomes research trial. *Spine (Phila Pa 1976)* 2014;39:3–16.
14. Weinstein JN, Lurie JD, Tosteson TD, et al. Surgical vs nonoperative treatment for lumbar disk herniation: the Spine Patient Outcomes Research Trial (SPORT) observational cohort. *JAMA* 2006;296:2451–9.
15. Lewis PJ, Weir BK, Broad RW, et al. Long-term prospective study of lumbosacral discectomy. *J Neurosurg* 1987;67:49–53.
16. Silverplats K, Lind B, Zoëga B, et al. Clinical factors of importance for outcome after lumbar disc herniation surgery: long-term follow-up. *Eur Spine J* 2010;19:1459–67.
17. Sharma H, Lee SW, Cole AA. The management of weakness caused by lumbar and lumbosacral nerve root compression. *J Bone Joint Surg Br* 2012;94:1442–7.
18. Koerner JD, Glaser J, Radcliff K. Which variables are associated with patient-reported outcomes after discectomy? Review of SPORT Disc Herniation Studies. *Clin Orthop Relat Res* 2015;473:2000–6.
19. Wilson CA, Roffey DM, Chow D, et al. A systematic review of preoperative predictors for postoperative clinical outcomes following lumbar discectomy. *Spine J* 2016;16:1413–22.
20. Azimi P, Benzel EC. Cut-off value for pain sensitivity questionnaire in predicting surgical success in patients with lumbar disc herniation. *PLoS One* 2016;11:e0160541.
21. Azimi P, Benzel EC, Montazeri A. Predictive score card in lumbar disc herniation: is it reflective of patient surgical success after discectomy? *PLoS One* 2016;11:e0154114.
22. White AP, Harrop J, Dettori JR. Can clinical and radiological findings predict surgery for lumbar disc herniation? A systematic literature review. *Evid Based Spine Care J* 2012;3:45–52.
23. Leonard DA, Schoenfeld AJ, Harris MB, et al. The prognostic value of preoperative participation in activities of daily living on postoperative outcomes following lumbar discectomy. *Clin Neurol Neurosurg* 2017;155:40–4.
24. Dewing CB, Provencher MT, Riffenburgh RH, et al. The outcomes of lumbar microdiscectomy in a young, active population: correlation by herniation type and level. *Spine (Phila Pa 1976)* 2008;33:33–8.
25. Hebert JJ, Fritz JM, Koppenhaver SL, et al. Predictors of clinical outcome following lumbar disc surgery: the value of historical, physical examination, and muscle function variables. *Eur Spine J* 2016;25:310–7.
26. Cella D, Yount S, Rothrock N, et al. The Patient-Reported Outcomes Measurement Information System (PROMIS): progress of an NIH Roadmap cooperative group during its first two years. *Med Care* 2007;45 (5 suppl 1):S3–11.
27. Garcia SF, Cella D, Clauser SB, 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; Review. Erratum in: J Clin Oncol. 2008 Feb 20;26(6):1018. Lad, Thomas [added].

28. Hung M, Nickisch F, Beals TC, et al. New paradigm for patient-reported outcomes assessment in foot & ankle research: computerized adaptive testing. *Foot Ankle Int* 2012;33:621−6.
29. Hung M, Stuart AR, Higgins TF, et al. 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.
30. Brodke DJ, Saltzman CL, Brodke DS. PROMIS for orthopaedic outcomes measurement. *J Am Acad Orthop Surg* 2016;24:744−9.
31. Papuga MO, Mesfin A, Molinari R, et al. Correlation of PROMIS physical function and pain CAT instruments with Oswestry Disability Index and Neck Disability Index in spine patients. *Spine (Phila Pa 1976)* 2016;41:1153−9.
32. Berliner JL, Brodke DJ, Chan V, et al. John Charnley award: preoperative patient-reported outcome measures predict clinically meaningful improvement in function after THA. *Clin Orthop Relat Res* 2016;474:321−9.
33. Berliner JL, Brodke DJ, Chan V, et al. Can preoperative patient-reported outcome measures be used to predict meaningful improvement in function after TKA? *Clin Orthop Relat Res* 2017;475:149−57.
34. Ho B, Houck JR, Flemister AS, et al. Preoperative PROMIS scores predict postoperative success in foot and ankle patients. *Foot Ankle Int* 2016;37:911−8.
35. Wong SE, Zhang AL, Berliner JL, et al. Preoperative patient-reported scores can predict postoperative outcomes after shoulder arthroplasty. *J Shoulder Elbow Surg* 2016;25:913−9.
36. Brodke DS, Goz V, Voss MW, et al. PROMIS® PF CAT Outperforms the ODI and SF-36 Physical Function Domain in Spine Patients. *Spine (Phila Pa 1976)* 2017;42:921−9.
37. Papuga MO, Beck CA, Kates SL, et al. Validation of GAITRite and PROMIS as high-throughput physical function outcome measures following ACL reconstruction. *J Orthop Res* 2014;32:793−801.
38. Berliner JL, Brodke DJ, Chan V, et al. John Charnley Award: preoperative patient-reported outcome measures predict clinically meaningful improvement in function after THA. *Clin Orthop Relat Res* 2016;474:321−9.
39. Wells G, Beaton D, Shea B, et al. Minimal clinically important differences: review of methods. *J Rheumatol* 2001;28:406−12.
40. Copay AG, Subach BR, Glassman SD, et al. Understanding the minimum clinically important difference: a review of concepts and methods. *Spine J* 2007;7:541−6.
41. Gortmaker SL. Applied Logistic-Regression - Hosmer, Dw, Lemeshow, S. *Contemp Sociol* 1994;23:159−1159.
42. Obuchowski NA. Determining sample size for ROC studies: what is reasonable for the expected difference in tests' ROC areas? *Acad Radiol* 2003;10:1327−8.
43. Haugen AJ, Brox JI, Grøvle L, et al. Prognostic factors for non-success in patients with sciatica and disc herniation. *BMC Musculoskelet Disord* 2012;13:183.
44. Chen CX, Kroenke K, Stump TE, et al. Estimating minimally important differences for the PROMIS pain interference scales: results from 3 randomized clinical trials. *Pain* 2018;159:775−82.
45. Amtmann D, Kim J, Chung H, et al. Minimally important differences for patient reported outcomes measurement information system pain interference for individuals with back pain. *J Pain Res* 2016;9:251−5.
46. Yost KJ, Eton DT, Garcia SF, et al. Minimally important differences were estimated for six Patient-Reported Outcomes Measurement Information System-Cancer scales in advanced-stage cancer patients. *J Clin Epidemiol* 2011;64:507−16.
47. Pearson AM, Blood EA, Frymoyer JW, et al. SPORT lumbar intervertebral disk herniation and back pain: does treatment, location, or morphology matter? *Spine (Phila Pa 1976)* 2008;33:428−35.