

2-2020

Role of Patient-Reported Outcome Measures on Predicting Outcome of Bunion Surgery

Ashlee MacDonald

Jeff Houck

Judith F. Baumhauer

Follow this and additional works at: https://digitalcommons.georgefox.edu/pt_fac



Part of the [Physical Therapy Commons](#)

Role of Patient-Reported Outcome Measures on Predicting Outcome of Bunion Surgery

Ashlee MacDonald, MD¹, Jeff Houck, PT, PhD²,
and Judith F. Baumhauer, MD, MPH¹ 

Abstract

Background: Prior studies have suggested preoperative patient-reported outcome scores could predict patients who would achieve a clinically meaningful improvement with hallux valgus surgery. Our goal was to determine bunionectomy-specific thresholds using Patient-Reported Outcomes Measurement Information System (PROMIS) values to predict patients who would or would not benefit from bunion surgery.

Methods: PROMIS physical function (PF), pain interference (PI), and depression assessments were prospectively collected. Forty-two patients were included in the study. Using preoperative and final follow-up visit scores, minimally clinically important differences (MCID), receiver operating characteristic (ROC) curves, and area under the curve (AUC) analyses were performed to determine if preoperative PROMIS scores predicted achieving MCID with 95% specificity or failing to achieve an MCID with 95% sensitivity.

Results: PROMIS PF demonstrated a significant AUC and likelihood ratio. The preoperative threshold score for failing to achieve MCID for PF was 49.6 with 95% sensitivity. The likelihood ratio was 0.14 (confidence interval, 0.02-0.94). The posttest probability of failure to achieve an MCID for PF was 94.1%. PI and depression AUCs were not significant, and thus thresholds were not determined.

Conclusion: We identified a PF threshold of 49.6, which was nearly 1 standard deviation higher than previously published. If a patient is hoping to improve PF, a patient with a preoperative *t* score >49.6 may not benefit from surgery. This study also suggests the need for additional research to delineate procedure-specific thresholds.

Level of Evidence: Level III, retrospective comparative series.

Keywords: hallux valgus, bunionectomy, patient-reported outcomes, PROMIS

Hallux valgus is a common condition of the foot with 4.4 million people seeking care annually, with over 200 000 undergoing surgical correction⁸ at an average cost of \$18 332 per surgery.²³ The decision of when to pursue operative intervention is multifactorial and includes patient factors such as pain, function, and an understanding of postoperative expectations. A prior report alerted surgeons that patients recall 10% or less of the surgical risks and postoperative expectations.²⁰ As a solution, some propose shared decision-making tools to ensure better comprehension of risks and expectations.^{21,22} However, these tools are usually specific to a procedure or disease and may lack input from patients.²²

A potential solution is to use patient-reported outcomes (PROs) for shared decision making preoperatively. PROs such as the Patient-Reported Outcomes Measurement Information System (PROMIS) are based on patients'

symptoms and are agnostic to disease. The PROMIS scales provide objective information derived directly from patients for shared decision making during preoperative counseling. Assessment of PRO scales often includes varied or short-term follow-up as this mirrors clinical delivery. These studies are distinct from therapeutic assessment of outcomes, which requires longer, more standardized follow-up. Initial

¹Department of Orthopaedics, University of Rochester, Rochester, NY, USA

²Department of Physical Therapy, George Fox University, Newberg, OR, USA

Corresponding Author:

Judith F. Baumhauer, MD, MPH, Department of Orthopaedics, University of Rochester, 601 Elmwood Avenue, Box 665, Rochester, NY 14642, USA.

Email: Judy_Baumhauer@urmc.rochester.edu

Table 1. Surgical Procedures and Postoperative Complications.

Characteristic	No.
Type of bunionectomy	
Distal chevron osteotomy	34
Lapidus	7
MTP arthrodesis	1
Additional procedures	
Second hammer toe corrections	9
Second and third hammer toe corrections	3
Akin osteotomy	2
Complications	
Removal of hardware	6
Superficial skin infection	1
Delayed wound healing	1
Persistent neuritis	1

Abbreviation: MTP, metatarsophalangeal.

assessments of the ability of PROMIS scales (pain interference [PI], depression, and physical function [PF]) to predict outcome after foot and ankle surgeries were positive.² However, a follow-up study performed in a separate geographic academic center confirmed this threshold value only for PROMIS PF.⁹ Although hallux valgus correction was identified as one of the most common surgeries in the previous studies,^{2,9} foot and ankle conditions have not been individually studied to determine if the previously published thresholds are generalizable or if each condition has its own unique set of preoperative PROMIS thresholds.

Therefore, we sought to assess the ability of preoperative PROMIS scores to predict postoperative improvement in hallux valgus surgery. The aims of this study were to evaluate the average change in PROMIS *t* scores from the pre- to postoperative follow-up after bunionectomy surgery and to determine if preoperative PROMIS PF, PT, and depression thresholds for bunionectomy surgery differed from the previously published values that encompassed all foot and ankle surgery.

Methods

PROMIS PF, PI, and depression scores were prospectively obtained from patients in an orthopedic foot and ankle clinic between February 2015 and November 2016. Patients 18 years and older who underwent various bunionectomy surgeries (Table 1) by a single surgeon were identified using *Current Procedural Terminology* codes. Exclusion criteria consisted of less than 2-month follow-up, multiple procedures unrelated to their bunionectomy during the follow-up period, and incomplete PROMIS assessment scores. A total of 65 patients who underwent bunionectomy during the study period were identified. After applying the exclusion criteria, 42 patients were included in the data analysis (Figure 1). The postoperative rehabilitation

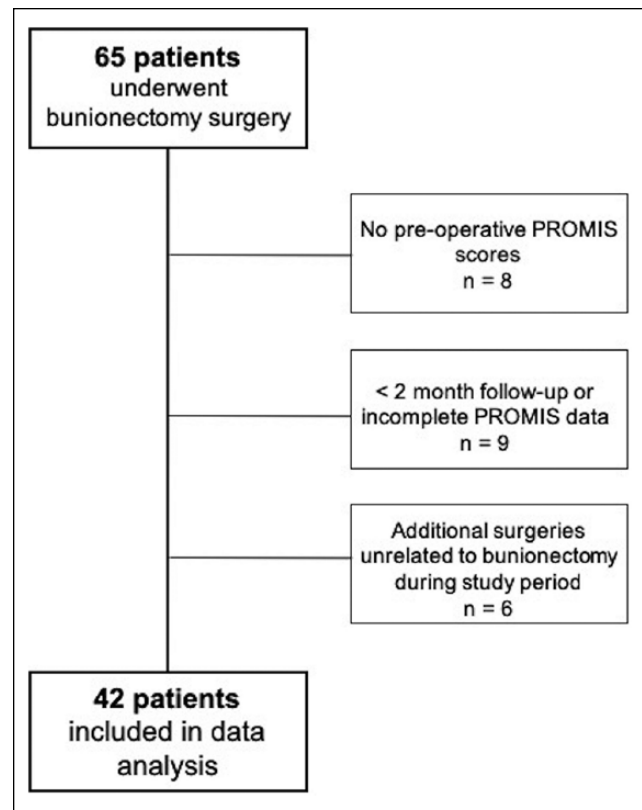


Figure 1. Patient selection for data analysis. PROMIS, Patient-Reported Outcomes Measurement Information System.

protocol depended on the type of bunionectomy performed (Table 2). The average follow-up during the study period was 18.3 weeks (8.7-48.3). Fourteen patients (33.3%) had at least 6 months of follow-up.

Outcome Measures

PROMIS PF, PI, and depression scales were administered during routine patient care. A strength of PROMIS scales is the integration of computer adaptive testing (CAT). The CAT approach selects the next most appropriate items for each scale based on a patient's previous answer, thus avoiding floor and ceiling effects.¹² This results in obtaining a score in 4 to 12 questions for each scale, minimizing the patient burden. For PROMIS PF, higher scores indicate greater perceived physical ability and therefore improvement. For PROMIS PI and depression scales, lower scores indicate less PI and depression and therefore improvement. For all 3 PROMIS scales, a *t* score of 50 is the average of the US population, and 10 points represents 1 standard deviation of the US population.

Statistical Analysis

A 2-way analysis of variance (ANOVA) was used to assess the average change in PROMIS scales from pre- to

Table 2. Postoperative Protocols.

Characteristic	Treatment
Chevron bunionectomy with and without hammer toe corrections	
Immediate postoperatively	Splint, NWB for 1 week
1 week postoperatively	Transition to WBAT in walking boot
5-6 weeks postoperatively	Pin removal if applicable, transition to normal shoe wear as able
First TMT arthrodesis	
Immediate postoperatively	Splint, NWB
1 week postoperatively	Transitioned to NWB in short leg cast for 6 weeks
6 weeks postoperatively	Transition to WBAT in walking boot
10 weeks postoperatively	Transition to normal shoe wear

Abbreviations: NWB, nonweightbearing; TMT, tarsometatarsal; WBAT, weightbearing as tolerated.

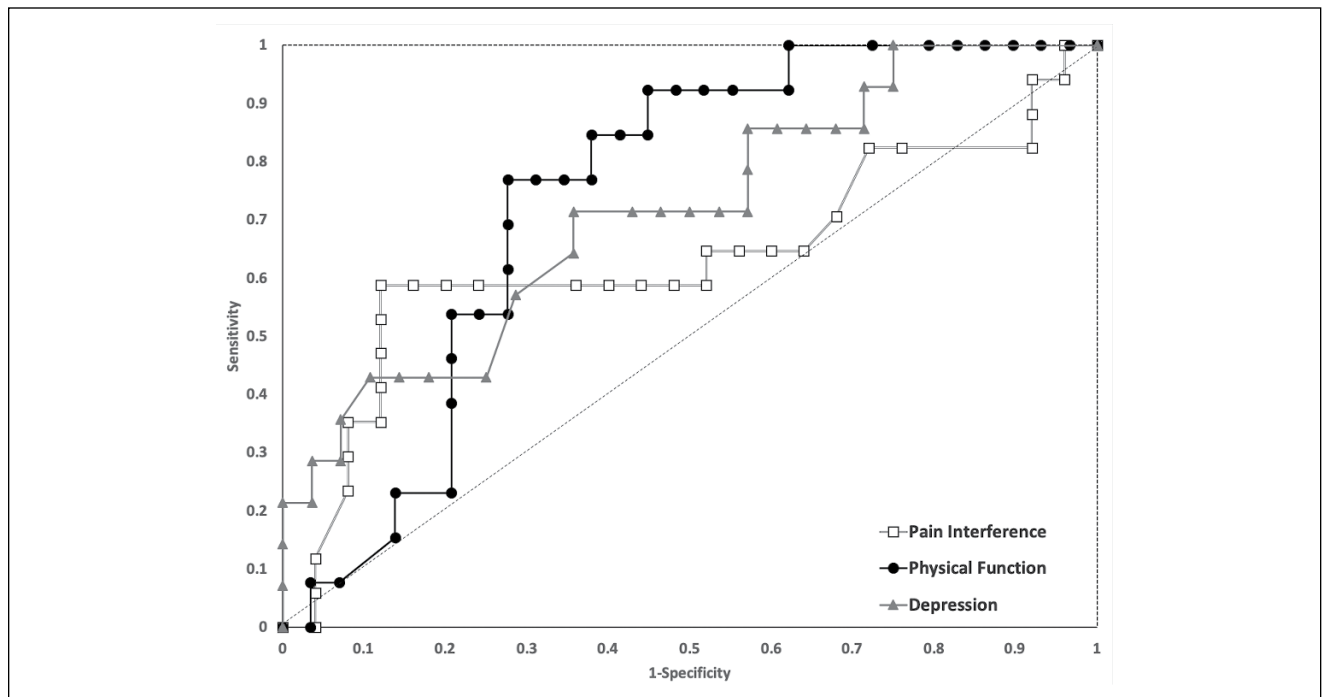


Figure 2. Receiver operating characteristic curves to assess the ability of preoperative Patient-Reported Outcomes Measurement Information System (PROMIS) scores to predict a minimally clinically important difference change for each PROMIS scale.

postoperative follow-up. One factor was PROMIS scales (PF, PI, and depression) and the second factor was time (preoperative/postoperative). Because improvement in PI and depression are indicated by lower scores and improvement in PF is indicated by a higher score, an interaction effect of the 2-way ANOVA is consistent with overall improvement across scales. To verify that the interaction resulted from improvement across all 3 PROMIS scales, pairwise comparisons were calculated to assess the change from pre- to postoperative scores on each scale.

Improvement was established by calculating the minimally clinically important differences (MCIDs) for each PROMIS scale. The distributive method estimates the MCID

as one-half standard deviation of the change score from pre- to postoperative follow-up and has been found to be a valid estimate of the MCID.¹⁵ Based on this approach, MCID thresholds were calculated as an increase of 3.9 or greater in PF, a decrease of 4.65 or greater in PI, and a decrease of 3.1 or greater in depression. Receiver operating characteristic (ROC) curves were used to assess the ability of preoperative PROMIS scores to predict an MCID change for each PROMIS scale (Figure 2). Area under the curve (AUC) was used to assess the ability of a PROMIS scale to discriminate an MCID change. Higher AUC values indicate greater ability to accurately discriminate individuals who achieved an MCID change. For PROMIS scales with significant AUCs,

Table 3. Patient Demographics and Patient-Reported Outcomes Measurement Information System Scores.

Characteristic	Mean	Standard Deviation	Mean Change (Preoperative to Postoperative)
Age	56.1	14.6	
Follow-up, wk	18.3	10.8	
Preoperative PF	48.5	6.9	
Preoperative PI	55.0	5.2	
Preoperative depression	46.4	8.2	
Postoperative PF	47.4	6.6	-1.1
Postoperative PI	52.4	8.6	-2.6
Postoperative depression	44.6	8.3	-1.8

Abbreviations: PF, physical function; PI, pain interference.

thresholds that resulted 95% specificity/sensitivity for determining patients who achieved an MCID were determined from the ROC curves. Likelihood ratios and posttest probabilities were also calculated for each preoperative PROMIS scale based on the new thresholds to determine the probability of correctly identifying those who would achieve or fail to achieve the MCID. The new threshold values specific to hallux valgus were then compared to the previous studies that included a wide variety of diagnoses.^{2,9}

To assess adequacy of the sample size, the AUC from the ROC curve was used. An AUC of 0.7 is considered acceptable or the minimum for clinical significance.¹³ The proportion of patients experiencing an MCID change was varied in 10% increments to determine significance of an AUC of 0.7.¹⁶ With a sample size of 42, the 95% confidence interval excluded 0.5 when the proportion of patients experiencing an MCID change was greater than 30%. All analysis was performed using SPSS version 25 (SPSS, Inc, an IBM Company, Chicago, IL).

Results

The demographics and mean preoperative and postoperative PROMIS scores are listed in Table 3. When stratified by follow-up less than 6 months and 6 months or greater, there were no differences in preoperative or postoperative PROMIS, although postoperative PI was nearly significant ($P = .051$, Table 4). Two-way ANOVA showed a significant effect for time where postoperative scores were lower on average than preoperative scores. A main effect occurred because all PROMIS scales were lower postoperatively compared to preoperatively. However, only depression was significantly lower, which represents an improvement in that domain (-2.5; 95% confidence interval [CI], -4.8 to -0.3). PROMIS PI (-2.8; 95% CI, -6.4 to 0.7) and PF (-1.1; 95% CI, -4.3 to 2.1)

Table 4. Preoperative and Postoperative PROMIS Scores Stratified by Follow-up.

Characteristic	<6 Months (n = 28), Mean (SD)	>6 Months (n = 14), Mean (SD)	P Value
Preoperative PROMIS scores			
PF	49.9 (6.5)	45.7 (6.9)	.06
PI	54.2 (5.1)	56.7 (5.2)	.15
Depression	46.4 (7.6)	46.5 (9.5)	.98
Postoperative PROMIS scores			
PF	48.0 (5.8)	46.2 (8.0)	.40
PI	50.6 (8.2)	56.1 (8.6)	.05
Depression	44.8 (8.6)	44.3 (7.9)	.88

Abbreviations: PF, physical function; PI, pain interference; PROMIS, Patient-Reported Outcomes Measurement Information System.

were minimally lower but not significantly different from preoperative scores. Lower PI scores would represent an improvement, whereas lower PF would represent worsening symptoms.

The ROC analysis yielded significant AUC for PF (0.74, $P = .014$) and depression (0.71, $P = .03$) but not for PI (0.63, $P = .14$). As a result, thresholds could not be determined for PI.

The threshold for achieving MCID for depression was 58.1 with 95% specificity. The likelihood ratio of 8.0 was not significant (95% CI, 0.98-65.0). The posttest probability of achieving an MCID for depression was 80%. The bunionectomy-specific threshold for failing to achieve MCID for depression was 39.8 with 95% sensitivity. The likelihood ratio of 0.25 was also not significant (95% CI, 0.03-1.81). The posttest probability of failing to achieve MCID for depression was 88.9%. Although the posttest probabilities (or probabilities that a patient will be correctly identified as achieving or failing to achieve MCID) were high, the likelihood ratios were not significant. This indicates that depression was not a major symptom contributing to outcomes following bunionectomy, and therefore the thresholds obtained would not be reliable predictors on postoperative outcomes.

The bunionectomy-specific preoperative threshold t score for achieving an MCID for PF was 39.6 with 95% specificity. The likelihood ratio was 2.23 (95% CI, 0.15-33.0). The posttest probability of achieving an MCID for PF was 70%. The preoperative threshold t score for failing to achieve an MCID for PF was 49.6 with 95% sensitivity (Figure 3). The likelihood ratio was 0.14 (95% CI, 0.02-0.94). The posttest probability of failing to achieve an MCID for PF was 94.1%. This was the only PROMIS domain that had a significant AUC and likelihood ratio, suggesting that this was the only domain that could reliably predict postoperative outcomes based on preoperative scores.

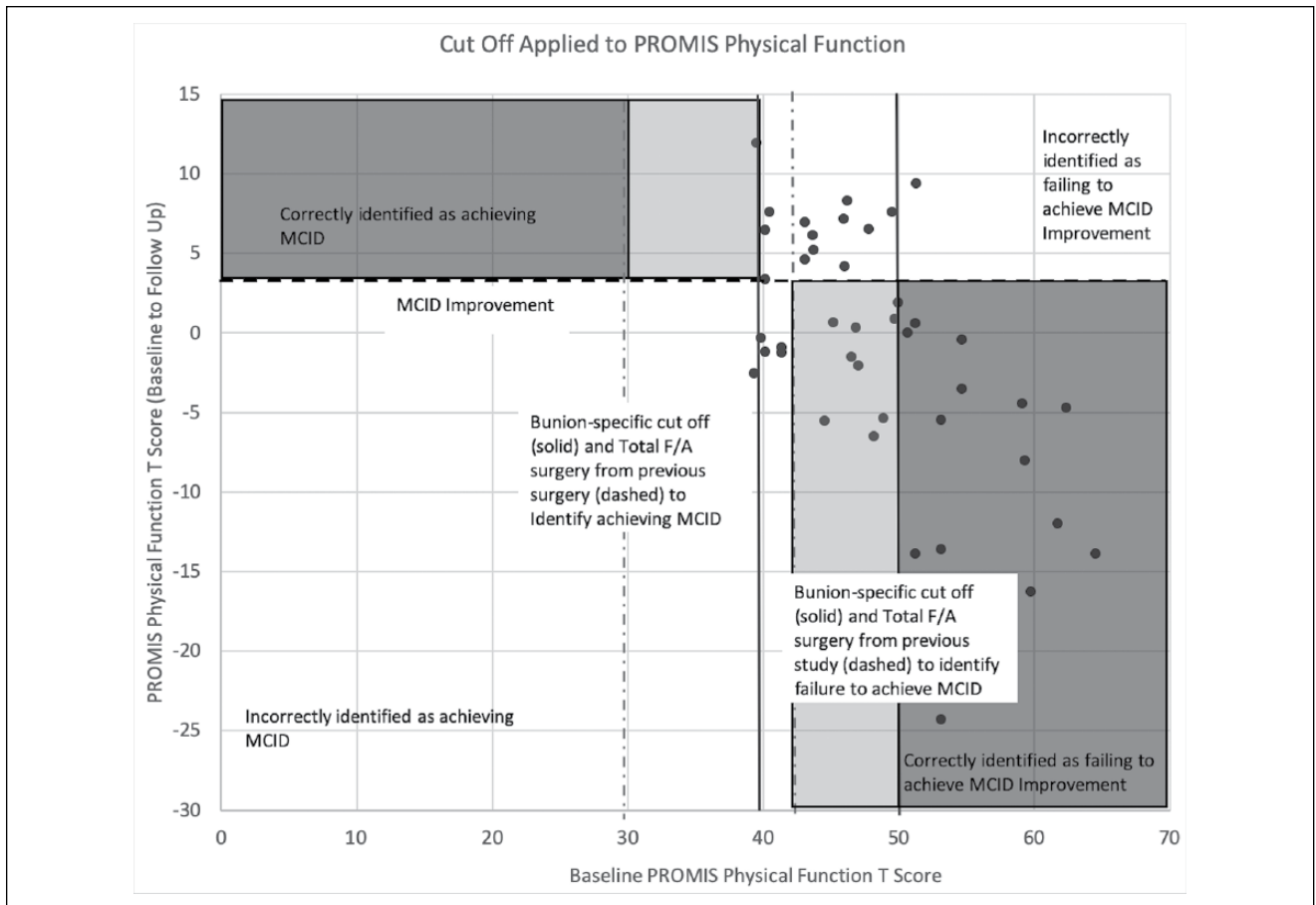


Figure 3. Correlation between the baseline Patient-Reported Outcomes Measurement Information System (PROMIS) physical function (PF) and change in PROMIS PF from baseline to follow-up in patients undergoing bunionectomy surgery. The total foot and ankle (F/A) surgery threshold to identify patients failing to achieve a minimally clinically important difference (MCID) improvement of >42 (dotted line) from a previous study, derived from a variety of foot and ankle surgeries, contrasts with the bunionectomy-specific threshold of >49.6 . In addition, the number of patients potentially influenced by this threshold is high. No patients met the bunionectomy-specific cutoff of 39.4 for achieving an MCID; therefore, its significance is questionable.

Discussion

Patient-reported outcome measures are becoming an integral component of health care practices. PROMIS scales specifically are emerging as powerful tools for preoperative counseling and shared decision making. PROMIS has been shown to be a reliable and efficient method of collecting PROs, making it an attractive alternative to legacy measures.^{3,10,14,17} In addition, PROMIS scores have been validated for foot and ankle outcome measures.^{2,11} Consistent with several studies,^{2,4-7,9,19} this study supports that preoperative PROs can be a useful tool for discussing postoperative treatment goals and expectations for domains such as PF. However, this study also suggests that specific diagnoses like hallux valgus may have distinct disease or procedure-specific thresholds as evidenced by the difference in our sample compared to prior studies evaluating all foot and ankle procedures.^{2,9}

The uniqueness of the sample may contribute to such differences in the thresholds. This sample suggests hallux valgus is characterized as nearly exclusively female (40/42), with relatively high physical function (48.5), moderate pain interference (55.0), and low depression (46.4) preoperatively (Table 3). Previous studies of foot and ankle surgery^{2,9} showed ~ 1 standard deviation (~ 10 points) worse scores for PF and PI, suggesting the preoperative status of patients in our sample was high compared to other foot and ankle surgeries. The bunionectomy-specific PF threshold score identified in this study (>49.6) was also nearly 1 standard deviation higher than previously reported for “ALL” foot and ankle surgeries (Table 5).^{2,9} A threshold of 49.6, or nearly the score for the “average” patient, also suggests that other factors not measured by PROMIS, such as shoe wear difficulties, skin issues, and pain, likely contribute to the decision to proceed with hallux valgus surgery.

Table 5. Preoperative Patient-Reported Outcomes Measurement Information System *t* Score Thresholds for Bunionectomy Compared to Prior Studies Evaluating All Foot and Ankle Surgeries.

Characteristic	Achieving MCID		Failing to Achieve MCID	
	Scale	<i>t</i> Score	Scale	<i>t</i> Score
Bunionectomy specific	Physical function	—	Physical function	>49.6
	Pain interference	—	Pain interference	—
	Depression	—	Depression	—
Ho et al ⁹	Physical function	<29.7	Physical function	>42
	Pain interference	>67.2	Pain interference	<55
	Depression	—	Depression	<41.5
Anderson et al ²	Physical function	<28	Physical function	>41.6
	Pain interference	—	Pain interference	—
	Depression	Not tested	Depression	Not tested

Abbreviations: MCID, minimally clinically important difference; —, values were not statistically significant, and therefore thresholds were unable to be determined.

Taking these considerations into account, the data of this study suggest that higher health status prior to orthopedic procedures makes preoperative decisions relative to surgery more relevant. Anderson et al¹ showed that while foot and ankle surgeries are largely successful (88% of patients), patients who considered surgery unsuccessful showed worse PROMIS PF, PI, and depression scales. The understanding that patients with high PRO scores preoperatively have less room to make significant improvements and vice versa is important and could represent a common statistical phenomenon called regression to the mean. Future data sets with larger samples may try to model this effect similar to application of other PRO data to preoperative orthopedic decisions.¹⁸ Nonetheless, PROMIS and PROs in general can be powerful tools and useful adjuncts to preoperative shared decision making. Additional studies may provide more refined predictive models in the future.

Our study has limitations. First, our minimum follow-up was only 2 months. Although 14 patients (33.3%) had at least 6 months of follow-up, 19 patients (45.2%) had less than a 3-month follow-up, all of which were allowed to follow up on an as-needed basis at that time point. The undefined postoperative follow-up limits the validity of the inference of the analysis. However, when comparing those with less than 6 months of follow-up to those with 6 months or greater follow-up, there were no differences in average postoperative PF or depression *t* scores. In those with at least six months of follow-up, there was a trend toward worse PI scores that nearly reached statistical significance. With a longer and more homogeneous follow-up period, it is possible that patients with lower PF scores may experience further improvement over time, which could ultimately alter the threshold *t* scores. It is also possible this would allow a threshold of PI *t* scores to be obtained to help delineate the differences seen in patients with less than and greater than 6 months

of follow-up. With these significant limitations, it is important to realize this study is best viewed as an assessment of the outcomes instrument and not necessarily an assessment of clinical outcome. Second, the cohort in our study represents patients from a single surgeon. Therefore, the outcomes and threshold scores may not be representative of patients treated by all foot and ankle surgeons. Last, all bunionectomy surgeries were included and therefore any differences between the type of surgery performed or the postoperative protocols were not evaluated given the small sample size, particularly of those who underwent a first tarsometatarsal arthrodesis. Despite our limitations, we were able to demonstrate that preoperative PROs can be predictors of postoperative outcomes. Our study also suggests that not all foot and ankle procedures are treated equal, and for improved preoperative counseling on postoperative expectations, additional studies are needed to further evaluate procedure-specific thresholds.

Conclusion

Patient-reported outcomes can be powerful predictors of postoperative outcomes in foot and ankle surgery. The preoperative bunionectomy-specific PROMIS PF threshold was nearly 1 standard deviation higher than previously published thresholds for all foot and ankle surgeries. This suggests the need for further studies to delineate procedure-specific thresholds for enhanced preoperative patient counseling and optimized shared decision making.

Author Contributions

Ashlee MacDonald, MD: data collection, data analysis, manuscript writing and editing; Jeff Houck, PT, PhD: data analysis, manuscript writing and editing; Judith F. Baumhauer, MD, MPH: study design, manuscript editing.


Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article. ICMJE forms for all authors are available online.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Judith F. Baumhauer, MD, MPH,  <https://orcid.org/0000-0003-2142-7778>

References

1. Anderson MR, Baumhauer JF, DiGiovanni BF, et al. Determining success or failure after foot and ankle surgery using Patient Acceptable Symptom State (PASS) and Patient Reported Outcome Information System (PROMIS). *Foot Ankle Int.* 2018;39(8):894-902.
2. Anderson MR, Houck JR, Saltzman CL, et al. Validation and generalizability of preoperative PROMIS scores to predict postoperative success in foot and ankle patients. *Foot Ankle Int.* 2018;39(7):763-770.
3. Baumhauer JF. Patient-reported outcomes—are they living up to their potential? *N Engl J Med.* 2017;377(1):6-9.
4. Berliner JL, Brodke DJ, Chan V, SooHoo NF, Bozic KJ. Can preoperative patient-reported outcome measures be used to predict meaningful improvement in function after TKA? *Clin Orthop Relat Res.* 2017;475(1):149-157.
5. Bernstein DN, Houck JR, Gonzalez RM, et al. Preoperative PROMIS scores predict postoperative PROMIS score improvement for patients undergoing hand surgery [published online August 3, 2018]. *Hand (N Y).*
6. Chen RE, Papuga MO, Nicandri GT, Miller RJ, Voloshin I. Preoperative Patient-Reported Outcomes Measurement Information System (PROMIS) scores predict postoperative outcome in total shoulder arthroplasty patients. *J Shoulder Elbow Surg.* 2019;28(3):547-554.
7. Chen RE, Papuga MO, Voloshin I, et al. Preoperative PROMIS scores predict postoperative outcomes after primary ACL reconstruction. *Orthop J Sports Med.* 2018;6(5):2325967118771286.
8. Coughlin MJ, Thompson FM. The high price of high-fashion footwear. *Instr Course Lect.* 1995;44:371-377.
9. Ho B, Houck JR, Flemister AS, et al. Preoperative PROMIS scores predict postoperative success in foot and ankle patients. *Foot Ankle Int.* 2016;37(9):911-918.
10. Hung M, Baumhauer JF, Brodsky JW, et al. Psychometric comparison of the PROMIS physical function CAT with the FAAM and FFI for measuring patient-reported outcomes. *Foot Ankle Int.* 2014;35(6):592-599.
11. Hung M, Baumhauer JF, Latt LD, et al. Validation of PROMIS Physical Function computerized adaptive tests for orthopaedic foot and ankle outcome research. *Clin Orthop Relat Res.* 2013;471(11):3466-3474.
12. Hung M, Franklin JD, Hon SD, et al. Time for a paradigm shift with computerized adaptive testing of general physical function outcomes measurements. *Foot Ankle Int.* 2014;35(1):1-7.
13. Mandrekar JN. Receiver operating characteristic curve in diagnostic test assessment. *J Thorac Oncol.* 2010;5(9):1315-1316.
14. Nixon DC, McCormick JJ, Johnson JE, Klein SE. PROMIS pain interference and physical function scores correlate with the Foot and Ankle Ability Measure (FAAM) in patients with hallux valgus. *Clin Orthop Relat Res.* 2017;475(11):2775-2780.
15. Norman GR, Sloan JA, Wyrwich KW. Interpretation of changes in health-related quality of life: the remarkable universality of half a standard deviation. *Med Care.* 2003;41(5):582-592.
16. Obuchowski NA. Determining sample size for ROC studies: what is reasonable for the expected difference in tests' ROC areas? *Acad Radiol.* 2003;10(11):1327-1328.
17. 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(6):793-801.
18. Riddle DL, Stratford PW, Perera RA. The incident tibiofemoral osteoarthritis with rapid progression phenotype: development and validation of a prognostic prediction rule. *Osteoarthritis Cartilage.* 2016;24(12):2100-2107.
19. Rubery PT, Houck J, Mesfin A, Molinari R, Papuga MO. Preoperative PROMIS scores assist in predicting early postoperative success in lumbar discectomy. *Spine (Phila Pa 1976).* 2019;44(5):325-333.
20. Shurnas PS, Coughlin MJ. Recall of the risks of forefoot surgery after informed consent. *Foot Ankle Int.* 2003;24(12):904-908.
21. Steffens NMNM. Engaging patients, health care professionals, and community members to improve preoperative decision making for older adults facing high-risk surgery. *JAMA Surg.* 2016;151(10):938-945.
22. Toupin-April K, Barton J, Fraenkel L, et al. Toward the development of a core set of outcome domains to assess shared decision-making interventions in rheumatology: results from an OMERACT Delphi survey and consensus meeting. *J Rheumatol.* 2017;44(10):1544-1550.
23. Willey JC, Reuter LS, Belatti DA, Phisitkul P, Amendola N. Availability of consumer prices for bunion surgery. *Foot Ankle Int.* 2014;35(12):1309-1315.