5-23-2018

Determining Success or Failure After Foot and Ankle Surgery Using Patient Acceptable Symptom State (PASS) and Patient Reported Outcome Information System (PROMIS)

Michael R. Anderson

Judith F. Baumhauer

Benedict F. DiGiovanni

Sam Flemister

John P. Ketz

See next page for additional authors

Authors

Michael R. Anderson, Judith F. Baumhauer, Benedict F. DiGiovanni, Sam Flemister, John P. Ketz, Irvin Oh, and Jeff Houck

Determining Success or Failure After Foot and Ankle Surgery Using Patient Acceptable Symptom State (PASS) and Patient Reported Outcome Information System (PROMIS)

Michael R. Anderson, DO[1,2], Judith F. Baumhauer, MD, MPH[3], Benedict F. DiGiovanni, MD[3], Sam Flemister, MD[3], John P. Ketz, MD[3], Irvin Oh, MD[3], and Jeffrey R. Houck, PhD[4]

Abstract

Background: As the role of generic patient-reported outcomes (PROs) expands, important questions remain about their interpretation. In particular, how the Patient Reported Outcome Measurement Instrumentation System (PROMIS) *t* score values correlate with the patients' perception of success or failure (S/F) of their surgery is unknown. The purposes of this study were to characterize the association of PROMIS *t* scores, the patients' perception of their symptoms (patient acceptable symptom state [PASS]), and determination of S/F after surgery.

Methods: This retrospective cohort study contacted patients after the 4 most common foot and ankle surgeries at a tertiary academic medical center (n = 88). Patient outcome as determined by phone interviews included PASS and patients' judgment of whether their surgery was a S/F. Assessment also included PROMIS physical function (PF), pain interference (PI), and depression (D) scales. The association between S/F and PASS outcomes was evaluated by chi-square analysis. A 2-way analysis of variance (ANOVA) evaluated the ability of PROMIS to discriminate PASS and/or S/F outcomes. Receiver operator curve (ROC) analysis was used to evaluate the ability of pre- (n = 63) and postoperative (n = 88) PROMIS scores to predict patient outcomes (S/F and PASS). Finally, the proportion of individuals classified by the identified thresholds were evaluated using chi-square analysis.

Results: There was a strong association between PASS and S/F after surgery (chi-square <0.01). Two-way ANOVA demonstrated that PROMIS *t* scores discriminate whether patients experienced positive or negative outcome for PASS (*P* < .001) and S/F (*P* < .001). The ROC analysis showed significant accuracy (area under the curve > 0.7) for postoperative but not preoperative PROMIS *t* scores in determining patient outcome for both PASS and S/F. The proportion of patients classified by applying the ROC analysis thresholds using PROMIS varied from 43.0% to 58.8 % for PASS and S/F.

Conclusions: Patients who found their symptoms and activity at a satisfactory level (ie, PASS yes) also considered their surgery a success. However, patients who did not consider their symptoms and activity at a satisfactory level did not consistently consider their surgery a failure. PROMIS *t* scores for physical function and pain demonstrated the ability to discriminate and accurately predict patient outcome after foot and ankle surgery for 43.0% to 58.8% of participants. These data improve the clinical utility of PROMIS scales by suggesting thresholds for positive and negative patient outcomes independent of other factors.

Level of Evidence: II, prospective comparative series.

Keywords: PROMIS, PASS scores, PRO, patient-reported outcomes

The use of patient-reported outcomes (PROs) is rapidly increasing in foot and ankle surgery as health care evolves to value-based care models.[1,11,20] Generic health metrics like the Patient Reported Outcome Measurement Instrumentation System (PROMIS) are appealing given their applicability across many medical conditions, allowing a common measurement metric to be followed throughout a medical system.[2,3] In support of this approach, recent head-to-head

psychometric comparisons of generic PROMIS scales outperformed disease-specific legacy measures, leading some authors to call for a paradigm shift in orthopedics to scales like PROMIS.[9,10,12] Proof in concept has been published, documenting the widespread adoption of computer adaptive testing (CAT) PROMIS physical function (PF), pain interference (PI), and depression (D) scales across a medical system with an 80% patient completion rate.[1] Although the validity of PROMIS scales for foot and ankle surgery has been demonstrated and high patient completion rates are possible, further study is needed to better understand how PROMIS scales can be successfully applied to clinical decision making.[8]

Current research has used a change in PROMIS t scores over an episode of care to determine a positive outcome rather than concrete benchmarks of success.[8,18] Although clinically meaningful change is important to determine when patients are improving, it does not mark when patients will feel at an acceptable level of symptoms and activity or when they are likely to consider their surgery a success. No studies to date have evaluated the association of PROMIS scales with measures of patient-reported success after foot and ankle orthopedic surgery. Knowing at what point PROMIS PF, PI, and D scales are likely to indicate a positive (ie, successful surgery) and negative (ie, failed surgery) patient outcome after orthopedic surgery would be valuable for clinical decision making.

The patient acceptable symptom state (PASS) is a validated, single question that asks patients if their current levels of symptoms and activity are satisfactory.[13,14,16,24] The PASS question captures patients' ability to adapt to their current health state. Although a strong association between a PASS rating and a patient's judgment of success after surgery is expected, they may capture different concepts.[23] Studies note that patients frequently define success based on explicit expectations discussed with the surgeon, as well as implicit assumptions not discussed with the surgeon.[4,17,22] Concordance of PASS "yes" and "success" is expected when patient symptoms are "livable" and they consider that the surgery meets expectations. Concordance of PASS "no" and "failure" is expected when symptoms and functional limitations persist. However, discordance may occur when patients are PASS yes but surgery does not meet expectations or when patients are PASS no and expectations are established that symptom relief or improvement in functional limitations is not the goal of surgery.[4,9,19,22] Clearly, understanding the association between PASS and success/failure (S/F) is likely informative to presurgical decisions. Further, associations of CAT PROMIS PF, PI, and D with patient outcomes (ie, PASS and/or S/F) after foot and ankle surgery may provide additional guidance for the surgeon and patient by establishing thresholds for goal setting.

The purpose of this study was fourfold. The first was to evaluate the association between definitions of patient outcome after orthopedic surgery (PASS and S/F). We hypothesized that PASS yes would strongly correlate with the patient's perception of surgical success but that the correlation of PASS no and surgical failure would be weaker due the influence of patient expectations of success. Second, we hoped to determine whether PROMIS t scores for PF, PI, and D discriminated groups of patients with a positive (PASS yes or success) as compared with a negative (PASS no or failure) outcome after foot and ankle orthopedic surgery. We expected large differences in PROMIS scales in patients with a positive as compared with a negative outcome. Third, we wanted to determine if PROMIS accurately predicts patients' perceptions of outcome after foot and ankle orthopedic surgery (assessed by PASS and S/F) based on their preoperative and/or follow-up scores. Based on a previous study,[8] it was hypothesized that preoperative PROMIS scores may predict patient outcomes (PASS or S/F). Also, we hypothesized that a specific degree of recovery indicated by PROMIS PF, PI, and D scales at follow-up may classify patients with sufficient accuracy (95% sensitivity/specificity) to determine positive or negative surgical outcome (PASS or S/F). Fourth, we wanted to evaluate the proportions of patients accurately classified versus those who are unable to be classified using the accuracy criteria (95% sensitivity/specificity). Higher proportions of patients accurately classified could demonstrate greater potential clinical impact.

Methods

Participants

Institutional review board approval was obtained to call surgical patients older than the age of 18 identified by billing

[1]Orthopedic Foot and Ankle Fellow, University of Rochester Medical Center, Rochester, NY, USA
[2]Summit Orthopedics, Woodbury, MN, USA
[3]Department of Orthopaedic Surgery, University of Rochester Medical Center, Rochester, NY, USA
[4]George Fox University, Newberg, OR, USA

Corresponding Author:
Michael R. Anderson, DO, Orthopedic Foot and Ankle Fellow, University of Rochester Medical Center, 601 Elmwood Avenue, Box 655, Rochester, NY 14642, USA.
Email: 1mranderson@gmail.com

Success	Which of the following applies to you? _____ I consider my surgery a success _____ I consider my surgery a failure
PASS	Taking into account all the activity you have during your daily life, your level of pain, and also your functional impairment, do you consider that the current state of your foot and ankle is satisfactory? _____ Yes _____ N0

Abbreviations: PASS, patient acceptable symptom state.

Figure 1. Questions asked to patients during the call-back interview.

data from a busy foot and ankle academic surgical practice from March 2015 to January 2017. The 4 most common CPT codes (CPT 28899, subtalar arthrodesis; CPT 28296, bunion correction; CPT 27698, secondary reconstruction of ankle collateral ligaments [Brostrom-type procedure]; CPT 27691, deep tendon transfer) were selected to generate the group of patients to be contacted. This provided a preoperative to postoperative follow-up time frame from 7 months to 2 years. These CPT codes yielded 539 patients within the study time frame.

Previous published studies have documented the validity of verbal administration of PROMIS scales via telephone call. Hence, a team of callers were trained to perform phone interviews.[21] Callers attempted to contact each listed patient 3 times. Ultimately, 117 of 539 (21.7%) patients were reached by phone: 2 declined to participate, and 115 completed the PROMIS PF, PI, and D scales and S/F and PASS questions. The age of contacted patients (mean, 55.4 years; range, 18-93 years) and that of not contacted patients 53.0 years (range, 18-80 years) were similar. Gender breakdown was also similar (contacted, 34.2% male; not contacted, 30.2% male). Minimum follow-up was set at 6 months postoperatively; of the 115 participants reached, 91 had follow-up time points that met criteria. Three of the 91 participants had bilateral problems and were excluded. This left a total of 88 participants included in the overall analysis. Of these 88 patients, preoperative PROMIS (PF, PI, and D) data were available on a subset of 63 participants.

Outcome Measures

PROMIS PF, PI, and D scales were used to assess patients. The strength of PROMIS lies in the CAT approach, whereby appropriately difficult items are selected based on the participant's previous answer, avoiding floor and ceiling effects.[5] Furthermore, administration is efficient, achieving a score in 4 to 12 questions in a mean time of 1 minute for each scale.[5] For PROMIS PF, higher scores indicate greater physical ability; for PROMIS PI and D scales, lower scores indicate less pain interference and depression. For all 3 CAT PROMIS scales, a T score of 50 is the mean of the US population, and 10 points represents 1 standard deviation.[15]

To determine patient outcome, participants were asked 2 dichotomous questions. The first question was the patient's perception of whether his or her surgery was a success or failure (Figure 1). The second question, PASS, asks patients to judge whether their current symptoms and activity level are satisfactory (Figure 1).[13] The PASS question has been used in other studies assessing musculoskeletal conditions.[13,14,16,24]

Statistical Analysis

The first analysis evaluated whether the patient-reported PASS and S/F judgments were similar using a chi-square analysis.

The second set of analyses compared the PROMIS scales in patients with a positive as compared with a negative patient outcome using analysis of variance (ANOVA). Two analyses, one using PASS and one using S/F as the grouping variable (ie, fixed factor) were applied to the PROMIS scales. Each analysis used an ANOVA (mixed 2-way) where PROMIS scales (PF, PI, and D) were considered the repeated factor. Using this approach, a significant interaction effect was consistent with the hypothesis that all 3 PROMIS scales discriminate patient outcome (ie, PASS or S/F). Similarly, to evaluate pre- to postoperative change in PROMIS scores, 2-way repeated-measure ANOVA (2 factors: time × PROMIS scales) was used followed by pairwise comparisons for each PROMIS scale. A power analysis for this comparison suggests that a difference of t scores of 4, pre- to postoperative, with a sample of 49 could be detected (alpha = .05, power = 0.80, standard deviation = 10). Our sample of 63 exceeds this projection.

The third analysis used receiver operator curve (ROC) analyses to determine if PROMIS accurately predicted patients' perception of outcome (assessed by PASS and S/F) based on their preoperative and/or follow-up scores. The area under the curve (AUC) of PROMIS (preoperative and follow-up) to predict patient outcome (PASS and success) was assessed to determine the significance. An AUC between 0.7 and 0.8 was considered accurate and between 0.8 to 0.9 clinically useful.[7] For these analyses, if the AUC was significant ($P < .05$ and $>.7$), then thresholds were

determined for PROMIS by selecting the values closest to 95% specificity/sensitivity.

The fourth analysis determined the potential clinical impact by examining proportions of patients classified using the PROMIS threshold values. The calculated PROMIS threshold values enabled patients to be categorized into 1 of 3 groups: (1) positive patient-designated outcome (PASS yes or success), (2) negative patient-designated outcome (PASS no or failure), and (3) ambiguous categorization (PROMIS score placed them between the 95% sensitivity/specificity threshold values for a positive/negative patient outcome). The proportions of patients classified using each PROMIS scale were evaluated using 3 × 2 tables where the rows are the PROMIS categorization for that scale and the columns are the known patient outcome (PASS yes/no or S/F). This resulted in six 3 × 2 tables that were each evaluated using a chi-square test.

Results

Patient demographics, PASS (yes/no), and S/F are listed in Table 1. Overall, 75.0% of patients designated themselves as PASS yes, and 85.3% identified their surgery as a success at follow-up. Of the 4 procedures, 18.2% were subtalar fusions, 17.0% were bunion corrections, 33.0% were repair of collateral ligaments, and 31.8% were tendon transfers. For specific procedures, the mean preoperative scores on PROMIS PF ranged from 38.2 ± 9.0 to 40.1 ± 3.9, PROMIS PI from 58.5 ± 5.2 to 61.6 ± 5.4, and PROMIS D from 54.9 ± 12.0 to 47.3 ± 8.4. Of the patients with complete preoperative and follow-up PROMIS scores (n = 63), the improvement for PROMIS PF was 6.5 ± 9.7, PROMIS PI 8.1 ± 10.9, and PROMIS D 6.3 ± 9.2.

A chi-square analysis demonstrated a significant correlation between PASS status and S/F (chi-square = 33.1, $P <$ .001) with an observed agreement of 85.5%. The high agreement occurred because 64 of 66 (71.1%) participants who answered PASS yes also judged their surgery a success. There was also agreement between a proportion of participants who were PASS no (13/22, 14.4%) who judged their surgery a failure. A proportion of participants reported PASS no yet considered their surgery a success (11/22, 12.2%), which made up the majority of discordant answers between the PASS and S/F question. Only 2 of 66 participants answered PASS yes and considered their surgery a failure.

The analyses showed that both definitions of patient outcome were discriminated by PROMIS. The 2-way ANOVA for PROMIS and PASS showed a significant interaction effect ($P < .001$) (Figure 2A). Pairwise comparisons showed a significant difference between PASS yes and PASS no for PROMIS PF of 9.6 (95% confidence interval [CI], 5.5 to 13.8), for PROMIS PI of −8.4 (95% CI, −3.9 to −12.8), and for PROMIS D of −7.9 (95% CI, −3.6 to −12.2). Similarly, the 2-way ANOVA for PROMIS and success exhibited a

Table 1. Sample Characteristics.

Parameter	No.	Mean ± SD	Range
Age, y	88	53.9 ± 15.6	18-80
Gender			
Female	48/90 (54.5%)		
Male	40/90 (45.5%)		
Follow-up, mo	88	17.0 ± 5.9	7.1-28.8
Procedure			
SA	16/88 (18.2%)		
BC	15/88 (17.0%)		
CL	29/88 (33.0%)		
TT	28/88 (31.8%)		
Outcomes			
PASS (yes %)	66/88 (75.0%)		
Success	75/88 (85.3%)		
PROMIS PF	88	46.3 ± 9.5	27.2-73.3
PROMIS PI	88	51.1 ± 10.1	38.7-74.1
PROMIS D	88	44.9 ± 9.4	34.2-69.5

Abbreviations: BC, bunion correction; CL, repair of collateral ligament (Brostrom-Gould-type procedure); PASS, patient acceptable symptom state; PROMIS, Patient Reported Outcome Measurement Instrumentation System; SA, subtalar arthrodesis; TT, deep tendon transfer.

significant interaction effect ($P < .001$) (Figure 2B). Pairwise comparisons showed a significant difference between success and failure for PROMIS PF of 10.6 (95% CI, 5.5 to 15.7), for PROMIS PI of −10.2 (95% CI, −4.9 to −15.5), and for PROMIS D of −7.6 (95% CI, −2.3 to −12.8).

The ROC analysis indicated that PROMIS scales at follow-up, but not at the preoperative time point, were predictive of PASS status (Table 2). The AUC for PROMIS PF was significant and higher than 0.8 at follow-up. The AUCs for PROMIS PI and PROMIS D were also significant at 0.75 and 0.74, respectively (Figure 3A). All AUCs for preoperative PROMIS scales were not significant and below 0.7 (ranging from 0.6 to 0.63). For the follow-up PROMIS scales, the thresholds for PASS yes were >50 for PF (95.8% specificity), <46.9 for PI (95.8% specificity), and <44.0 for D (91.3% specificity). Similarly, for the follow-up PROMIS scales, the thresholds for PASS no were <33.5 for PF (95.7% sensitivity), >65.6 for PI (95.7% sensitivity), and >60.3 for D (95.6% sensitivity).

The ROC analysis showed that PROMIS domains at follow-up, but not at the preoperative time point, were predictive of S/F (Table 3). The AUC for PROMIS PF was significant and 0.8 or higher at the follow-up time point. The AUCs for PROMIS PI and PROMIS D were also significant at 0.78 and 0.74, respectively. All AUCs for preoperative PROMIS scales were not significant, except for PROMIS D (Figure 3B), with the AUCs below 0.7 (ranging from 0.49 to 0.68). For the follow-up PROMIS scales, the thresholds for success were >45.3 for PF (92.3% specificity), <47.3 for PI (92.3% specificity), and <45.8 for D
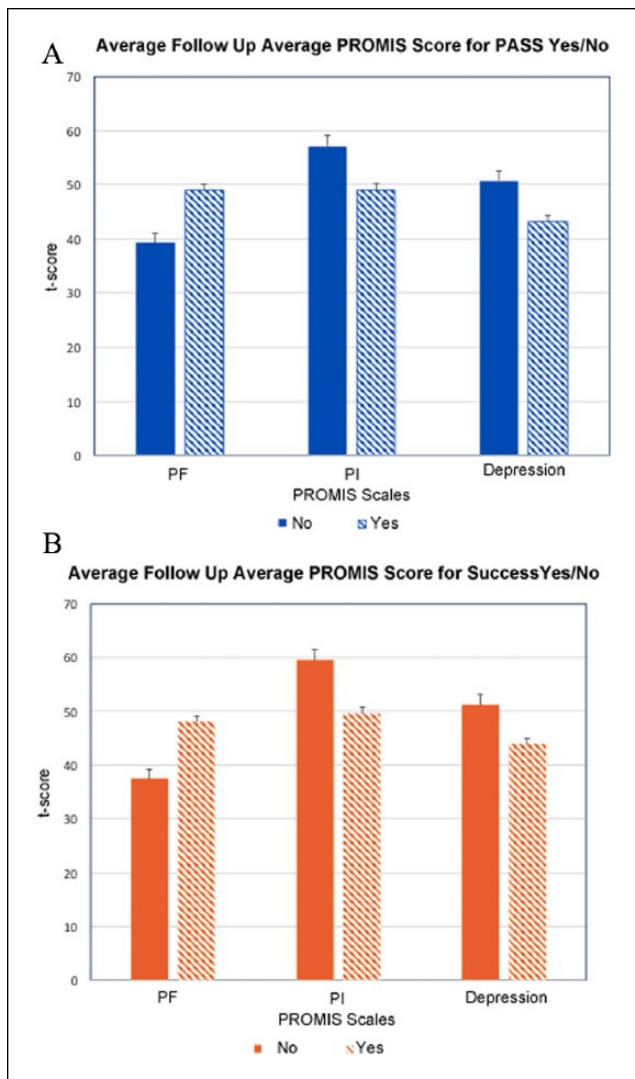
Figure 2. Mean values for postoperative Patient Reported Outcome Instrumentation System (PROMIS) scales for patients who were patient acceptable symptom state (PASS) yes versus PASS no (A) and for patients who considered their surgery a success versus a failure (B).

Table 2. Receiver Operator Curve Analysis for Patient Acceptable Symptom State.

Parameter	AUC	Significance	95% CI
Preoperative PROMIS			
PF	0.60 (0.08)	.21	0.45-0.74
PI	0.60 (0.08)	.19	0.44-0.76
D	0.63 (0.08)	.10	0.48-0.78
Postoperative PROMIS			
PF	0.80 (0.05)	<.001	0.70-0.90
PI	0.75 (0.05)	<.001	0.64-0.86
D	0.75 (0.06)	<.001	0.64-0.86
Parameter	PASS Yes	PASS No	Sensitivity	Specificity
PF	>50.0	<33.5	95.7%	95.8%
PI	<46.9	>65.6	95.7%	95.8%
Depression	<44.0	>60.3	95.6%	91.3%

Abbreviations: AUC, area under the curve; CI, confidence interval; D, depression; PASS, patient acceptable symptom state; PF, physical function; PI, pain interference; PROMIS, Patient Reported Outcome Instrumentation System.

Discussion

To our knowledge, this is the first study to evaluate the PASS concept as it relates to PROMIS scores. It is also the first study to evaluate the patient outcomes PASS and success in foot and ankle surgery. By using PASS scores, this study provides further context for the interpretation of PROMIS when considering patient status during postoperative follow up.[1,2,20]

There are several important findings in this analysis of patient outcome after foot and ankle orthopedic surgery. First, there is a strong correlation between PASS and the patient's perception of success after surgery. Second, a large majority of patients experience a positive patient outcome after common procedures in foot and ankle surgery (PASS yes or success). Third, PROMIS discriminates between aggregate positive and negative patient outcomes (PASS and S/F). Finally, preoperative PROMIS scales were unable to predict positive and negative patient outcome as defined by PASS and success at follow-up. However, follow-up PROMIS scores were effective at identifying patients with positive and negative patient outcome designation, suggesting ongoing tracking of PROMIS may have benefit.

There was strong agreement between positive classifications of PASS and success at follow-up. This association confirms that the PASS question is measuring the patient's perception of a successful outcome. Patients overwhelmingly considered their surgery a success (83.3%), and this matched well with a PASS yes classification (73.3%). However, some patients who were PASS no considered their surgery a success. While 13.4% of patients categorized themselves as PASS no and also rated their surgery as a failure, a small subset of patients (12.2%) gave a discordant rating of PASS no and rated their surgery as a success. This
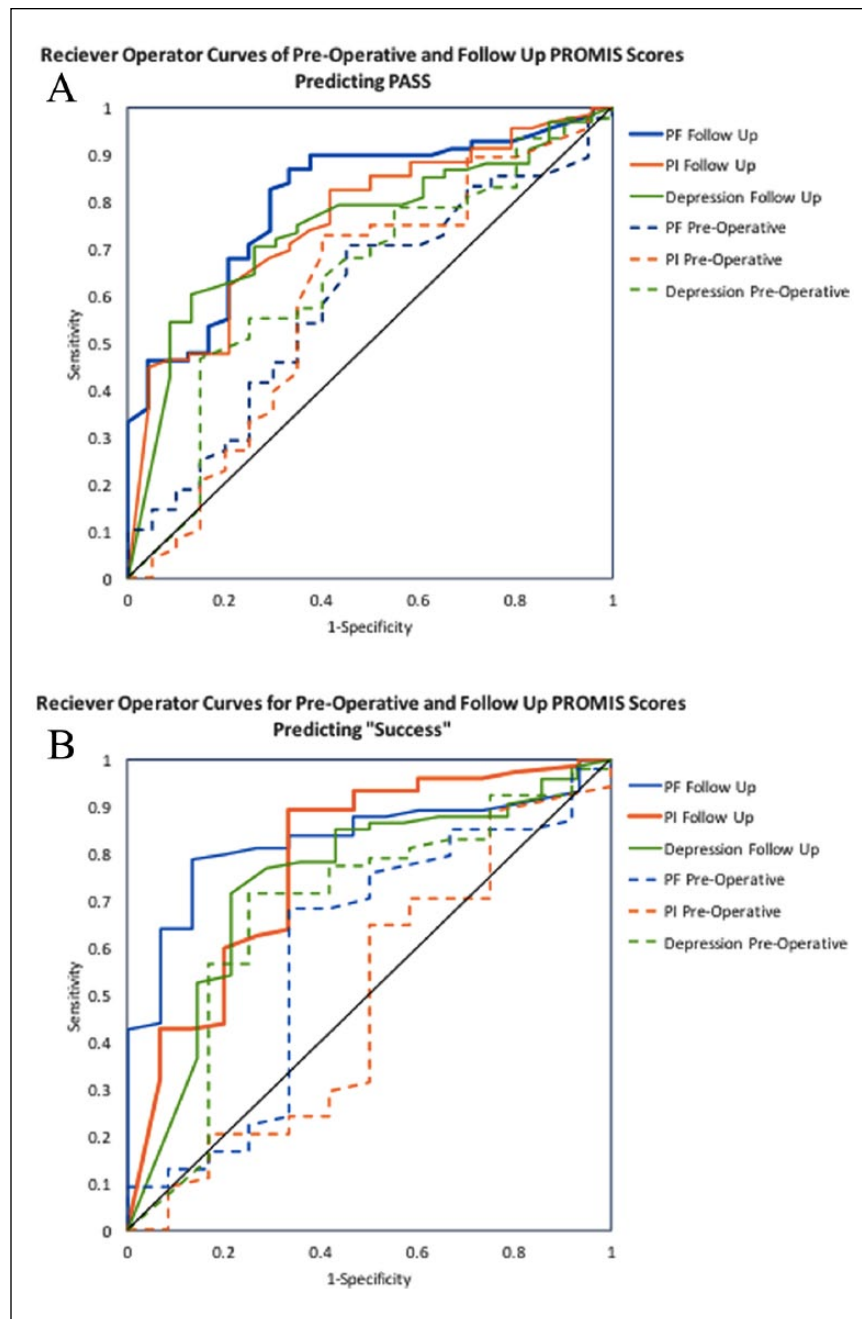
(85.7% specificity). Similarly, for the follow-up PROMIS scales, the thresholds for failure were <33.2 for PF (96.0% sensitivity), >64.5 for PI (94.7% sensitivity), and >60.3 for D (95.6% sensitivity).

Chi-square analyses of 3 × 2 tables for the ability of PROMIS to categorize participants into groups that are expected to have experienced a positive or negative patient outcome as designated by PASS or success scores were all significant (Table 4). Thresholds applied to PROMIS scales correctly placed 38% to 44% of patients into a PASS yes or no category (Table 4A). Similarly, thresholds applied to PROMIS scales correctly placed 43% to 57% of patients into an S/F category (Table 4B).

Figure 3. Receiver operator curves for predicting patient acceptable symptom state (PASS) (A) and success (B) from preoperative and follow-up data. Abbreviations: PF, physical function; PI, pain interference.

fits the hypothesis that success, unlike PASS, which focuses on current symptoms, may depend on expectations established between the surgeon and patient. Preoperative recognition of patients who are unlikely to achieve PASS yes may allow surgeons to guide patient expectations.[17,22]

These data affirm the ability of PROMIS scales to determine differences in aggregate scores between patients with a positive and negative outcome (PASS or S/F) after foot

and ankle surgery. Although previous studies showed aggregate differences over time, this is the first study to show aggregate differences of positive as compared with negative patient outcome at follow-up (Figure 2).[8,11] The differences between patients with a negative outcome and positive outcome were large, ranging from 7.3 to 10.6 for all PROMIS scales, irrespective of whether patient outcome was defined by PASS or S/F. This large difference explains how

Table 3. Receiver Operator Curve Analysis and Thresholds for Success.

Parameter	AUC	Significance	95% CI	
Preoperative PROMIS				
PF	0.59 (0.10)	.32	0.40-0.79	
PI	0.49 (0.10)	.90	0.29-0.69	
D	0.68 (0.09)	.05	0.50-0.86	
Postoperative PROMIS				
PF	0.83 (0.05)	<.001	0.73-0.92	
PI	0.78 (0.07)	.001	0.63-0.92	
D	0.75 (0.08)	.003	0.61-0.90	
Parameter	PASS Yes	PASS No	Sensitivity	Specificity
PF	>45.3	<33.2	96.0%	92.3 %
PI	<47.3	>64.5	94.7%	92.3 %
D	<45.8	>60.3	95.6%	85.7 %

Abbreviations: AUC, area under the curve; CI, confidence interval; D, depression; PASS, patient acceptable symptom state; PF, physical function; PI, pain interference; PROMIS, Patient Reported Outcome Instrumentation System.

Table 4. Proportions of Patients Classified by Thresholds.
A: Patient Acceptable Symptom State.

Parameter	PASS Yes	PASS No	Total	Chi-Square, P Value
PROMIS PI				
PASS no	3.4%	5.7 %	9.1%	
Ambiguous	35.2%	20.5%	55.7%	
PASS yes	34.1%	1.1%	35.2%	16.3, <.001
PROMIS PF				
PASS no	3.4%	3.4%	6.8%	
Ambiguous	35.2%	20.5%	55.7%	
PASS yes	34.1%	3.4%	37.5%	9.1, .01
PROMIS D (n = 88)				
PASS no	3.5%	3.5%	7.0%	
Ambiguous	29.1%	20.9%	50.0%	
PASS yes	40.7%	2.3%	43.0%	15.3, <.001

B: Success.

Parameter	Success	Failure	Total	Chi-Square, P Value
PROMIS PI				
Success no	3.4%	6.8 %	10.2%	
Ambiguous	43.2%	9.1%	52.3%	
Success yes	36.4%	1.1%	37.5%	20.3, <.001
PROMIS PF				
Success no	1.1%	3.4%	4.5%	
Ambiguous	26.1%	14.8%	40.9%	
Success yes	53.4%	1.1%	54.5%	17.0, <.001
PROMIS D (n = 88)				
Success no	2.3%	4.7%	7.0%	
Ambiguous	32.6%	10.5%	43.0%	
Success yes	46.5%	3.5%	50.0%	5.8, .056

Abbreviations: D, depression; PASS, patient acceptable symptom state; PF, physical function; PI, pain interference; PROMIS, Patient Reported Outcome Instrumentation System.

significance was achieved, with a relatively small sample of patients experiencing a negative outcome (ie, PASS no [n = 23] and failure [n = 15]). Furthermore, these aggregate differences suggest patients with a negative outcome are experiencing important limitations in physical function, as well as high levels of pain and depression.[15] These findings reinforce the emphasis on preventing unnecessary foot and ankle surgeries.[8]

Irrespective of how positive patient outcome was defined, patients were likely to report a positive outcome if they were at or near (<0.5 standard deviation) the average of the US population (*t* score = 50). Specifically, the threshold for PASS yes for PROMIS PF was >50, while PROMIS PI was <46.9. Similarly, the threshold for success for PROMIS PF was >45.3, while PROMIS PI was <47.3. Both PROMIS PI and PF fall within one-half standard deviation of 50 for both success and PASS yes, signifying that patients who view themselves as "average" are likely to be satisfied with their outcomes. These new benchmarks are useful for goal setting clinically and when discussing recovery with patients. In addition, the identified thresholds may be helpful in determining patients who may not need ongoing face-to-face follow-up.

In contrast, a definitive negative patient outcome was thresholds greater than 1 standard deviation lower on PROMIS scales. Thresholds for a negative patient outcome as defined by PASS no were PF <33.5, PI >65.6, and D >60.3. Similarly, thresholds for a negative outcome as defined by patients rating their surgery as a failure were PF <33.2, PI >64.5, and D >60.3. Thresholds for a negative outcome across all 3 domains deviate by more than 1 standard deviation from the average of the US population, reinforcing the aggregate findings that some patients experience important limitations in function and high levels of pain and depression after foot and ankle surgery. These thresholds for negative patient outcomes should assist providers in knowing when recovery is below expectations.

The ability of specific PROMIS thresholds to predict patient outcomes supports the use of these generic scales for clinical decision making. The advantage of using these PROMIS scales is that they capture key health domains (function, pain, and depression) across medical conditions. When PROMIS scores deviate from the identified thresholds, providers are alerted to address patients' needs specific to physical function, pain, and depression. The thresholds identified suggest generic goals that lead to definitive (95% accuracy) determination of who will likely experience a positive or negative outcome. Although a significant proportion of patients are successfully classified based only on their PROMIS rating, not surprisingly, other factors also influence patient outcome.

Application to clinical practice may also be enhanced by understanding what other factors contribute to patient outcomes aside from generic PRO assessment. The proportion of patients classified (negative or positive outcome) using the calculated thresholds for PROMIS varied from 38% to

57% (Table 4). For the remainder of the patients, other factors must be considered. One potential factor is preoperative severity. Preoperative PROMIS scores were available for 63 patients; however, ROC analysis demonstrated that preoperative PROMIS scores were not associated with patient-designated outcome (Figures 3 and 4). This contrasts with other studies that determined preoperative PROMIS scores were predictive of minimal clinically important difference (MCID).[6,8] This difference between MCID and patient outcome requires further exploration due to differences in follow-up and sample characteristics. Other factors may also differentiate patient outcome, such as socioeconomic status and level of education.[24] Nevertheless, what is surprising is that PROMIS assessment alone accounted for 43.0% to 58.8% of patients definitively (95% accuracy) classified as reporting a positive or negative outcome independent of other likely important factors. Future studies might consider a combination of socioeconomic, clinical examination, and PROMIS assessment to enhance preoperative assessment.

There are several limitations of this study that are important to consider. A low percentage of patients were contacted within the target population. At least 3 attempts were made to contact each patient. Most patients simply did not answer the phone. Among those patients who did answer, 98% completed the survey. Although demographic analysis of patients who did and did not answer shows that the 2 groups are similar in terms of age and gender, other factors may make this group distinct. A second important limitation is the low number of patients reporting a negative outcome (ie, PASS no and success no). Due to the low number of patients reporting a negative outcome, there is risk that the findings are not generalizable to other populations. The follow-up length is also likely important, with a mean follow-up of 17 months, of which 86.4% of the participants were contacted at greater than 9 months postsurgery. The current t score values are an initial estimate of thresholds for patient outcomes that should be verified in a separate sample.[8] This study looked at PROMIS t scores, and there are other patient-specific data that might be helpful in further discriminating the variability (radiographic, physical tests, comorbidities, etc). Finally, there is considerable heterogeneity among and within the various procedures evaluated. For example, multiple hallux valgus techniques were represented, including chevron osteotomy, modified Mau osteotomy, and Lapidus. In addition, deep tendon transfer is frequently undertaken for a multitude of purposes, including drop foot, adult-acquired flat foot, and Achilles pathology. The results reported herein are potentially specific to the foot and ankle procedures evaluated. There may be important differences across procedures that are masked in this sample where patient data were collapsed across the 4 different CPT codes.

Conclusions

Patient-designated outcomes for PASS and success were similar (concordance) for patients who have a positive outcome; however, these definitions were distinct for patients with a negative patient outcome. Patients who found their symptoms and activity at a satisfactory level (ie, PASS yes) were also likely to consider their surgery a success. However, patients who did not consider their symptoms and activity at a satisfactory level were not likely to consistently consider their surgery a failure. PROMIS t scores for physical function and pain demonstrated the ability to discriminate and accurately predict patient outcome after foot and ankle surgery for 43.0% to 58.8% of participants. These data improve the clinical utility of PROMIS scales by suggesting thresholds for positive and negative patient outcomes independent of other factors.

References

1. Baumhauer JF. Patient-reported outcomes: are they living up to their potential? *N Engl J Med*. 2017;377(1):6-9.
2. Baumhauer JF, Bozic KJ. Value-based healthcare: patient-reported outcomes in clinical decision making. *Clin Orthop Relat Res*. 2016;474(6):1375-1378.
3. Cella D, Riley W, Stone A, 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(11):1179-1194.
4. Cody EA, Mancuso CA, Burket JC, et al. Patient factors associated with higher expectations from foot and ankle surgery. *Foot Ankle Int*. 2017;38(5):472-478.
5. Cook KF, O'Malley KJ, Roddey TS. Dynamic assessment of health outcomes: time to let the CAT out of the bag? *Health Serv Res*. 2005;40(5 pt 2):1694-1711.
6. Escobar A, Riddle DL. Concordance between important change and acceptable symptom state following knee arthroplasty: the role of baseline scores. *Osteoarthritis Cartilage*. 2014;22(8):1107-1110.
7. Gortmaker S, Hosmer D, Lemeshow S. Applied logistic regression. *Contemp Sociol*. 1994;23:159.
8. 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.
9. 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.

10. 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(8):621-626.

11. Hunt KJ, Alexander I, Baumhauer J, et al. The Orthopaedic Foot and Ankle Outcomes Research (OFAR) network: feasibility of a multicenter network for patient outcomes assessment in foot and ankle. *Foot Ankle Int*. 2014;35(9):847-854.

12. Koltsov JCB, Greenfield ST, Soukup D, Do HT, Ellis SJ. Validation of Patient-Reported Outcomes Measurement Information System computerized adaptive tests against the Foot and Ankle Outcome Score for 6 common foot and ankle pathologies. *Foot Ankle Int*. 2017;38(8):870-878.

13. Kvien TK, Heiberg T, Hagen KB. Minimal clinically important improvement/difference (MCII/MCID) and patient acceptable symptom state (PASS): what do these concepts mean? *Ann Rheum Dis*. 2007;66(suppl 3):iii40-1.

14. Levy DM, Kuhns BD, Chahal J, et al. Hip arthroscopy outcomes with respect to patient acceptable symptomatic state and minimal clinically important difference. *Arthroscopy*. 2016;32(9):1877-1886.

15. Liu H, Cella D, Gershon R, et al. Representativeness of the Patient-Reported Outcomes Measurement Information System Internet panel. *J Clin Epidemiol*. 2010;63(11):1169-1178.

16. Muller B, Yabroudi MA, Lynch A, et al. Defining thresholds for the patient acceptable symptom state for the IKDC subjective knee form and KOOS for patients who underwent ACL reconstruction. *Am J Sports Med*. 2016;44(11):2820-2826.

17. Noble PC, Fuller-Lafreniere S, Meftah M, Dwyer MK. Challenges in outcome measurement: discrepancies between patient and provider definitions of success. *Clin Orthop Relat Res*. 2013;471(11):3437-3445.

18. 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.

19. 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.

20. Porter ME, Lee TH. From volume to value in health care: the work begins. *JAMA*. 2016;316(10):1047-1048.

21. Quach CW, Langer MM, Chen RC, et al. Reliability and validity of PROMIS measures administered by telephone interview in a longitudinal localized prostate cancer study. *Qual Life Res*. 2016;25(11):2811-2823.

22. Steffens NM, Tucholka JL, Nabozny MJ, et al. 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.

23. Tubach F, Ravaud P, Baron G, et al. Evaluation of clinically relevant states in patient reported outcomes in knee and hip osteoarthritis: the patient acceptable symptom state. *Ann Rheum Dis*. 2005;64(1):34-37.

24. Wright AA, Hensley CP, Gilbertson J, Leland JM III, Jackson S. Defining patient acceptable symptom state thresholds for commonly used patient reported outcomes measures in general orthopedic practice. *Man Ther*. 2015;20(6):814-819.