2018

Ability of Patient-Reported Outcomes to Characterize Patient Acceptable Symptom State (PASS) After Attending a Primary Care Physical Therapist and Medical Doctor Collaborative Service: A Cross-Sectional Study

Jeff Houck
George Fox University, jhouck@georgefox.edu

Daniel Kang
George Fox University, dkang@georgefox.edu

Tyler Cuddeford
George Fox University, tcuddeford@georgefox.edu

Sarah Rahkola

Ability of Patient-Reported Outcomes to Characterize Patient Acceptable Symptom State (PASS) After Attending a Primary Care Physical Therapist and Medical Doctor Collaborative Service: A Cross-Sectional Study

Jeff Houck, PT, PhD,[a] Daniel Kang, PT, DPT,[a] Tyler Cuddeford, PT, PhD,[a] Sarah Rahkola, MD[b]

From the [a]George Fox University, Newberg, OR; and [b]Providence Medical Group, Newberg, OR.

Abstract

Objectives: To determine if the Patient-Reported Outcome Measurement Information System (PROMIS) physical function, pain interference, self-efficacy, and global rating of normal function (GRNF) scales are able to accurately characterize a patient's acceptable symptom state (PASS).

Design: A cross-sectional analysis, using receiver operator curves and chi-square analysis to explore criteria to determine thresholds (80% and 95% sensitivity/specificity) for PASS that are applicable to PROMIS and GRNF scales.

Setting: Phone survey after primary care.

Participants: Patients (N = 94) attending primary care for musculoskeletal problems.

Interventions: Not applicable.

Main Outcomes Measures: Accuracy and proportion of patients classified as PASS Yes or No.

Results: Receiver operator curve analysis showed significant area under the curve (AUC) values for each PROMIS scale (AUC > .72) and the GRNF rating (AUC = .74). Identified PROMIS thresholds suggested PASS was achieved when scores were at or slightly worse than the US population average. A score of ≥7 and ≤4 characterized patients that were PASS Yes and No, respectively, on the GRNF rating. A moderate (80%) specificity/sensitivity criteria yielded 72.3%-73.5% accuracy for a majority of participants (>69.9%).

Conclusion: This analysis suggests the PROMIS and GRNF scales are able to characterize PASS status with moderate accuracy (~70%) for a large portion of patients (~70%). New to this study is the association of self-efficacy with PASS status. PROMIS scales at or slightly worse than the US population average characterized PASS status.

Archives of Physical Medicine and Rehabilitation 2019;100:60-6

There is enthusiasm for new generic patient-reported outcomes (PRO), such as the Patient-Reported Outcome Measurement Information System (PROMIS), to assess patient status of musculoskeletal problems in primary care.[1-5] The generic PROMIS scales are not disease specific, making them applicable across diagnoses.[1-5] Also, use of computer adaptive testing (CAT) reduces floor and ceiling effects while maintaining low patient burden (ie, average completion time is <1min per scale).[5-7] However, selection of specific domains of health for targeted groups of patients and the interpretation of PROMIS scales are still evolving.

Ideally, health domains for PRO assessment should match the patient's needs and inform provider decisions. The PROMIS physical function (PF) and pain interference (PI) scales have demonstrated validity for tracking patients with musculoskeletal problems.[4,5,8,9] Self-efficacy (SE), the concept of the ability to act on one's own behalf, is a positive behavioral health attribute

Presented to the American Physical Therapy Association, Combined Sections Meeting, February 23, 2018, New Orleans, LA.

Supported by Yamhill County Coordinated Care Organization - Build Forward Grant and Innovation Grant Providence Health and Services.

associated with health outcomes.[10] Because individuals ultimately will self-manage their musculoskeletal problems,[11] understanding a patients' confidence in managing their symptoms may influence provider decisions of how much and what kind of support is needed. The PROMIS SE of symptom management scale may serve this purpose, complementing the PROMIS PF and PI scales in a primary care setting.

Another PRO that is potentially useful in primary care is associated with the patient's internal reference of whether their joint or body region is healthy. Global ratings of normal function (GRNF) are included in current scales to capture a patient's rating of their joint relative to healthy joints.[12-14] The GRNF scale contrasts with the PROMIS scales by asking the patient to judge whether their joint or body region is normative rather than measuring a specific health domain. Previous data document that differences on how a question is framed, relative to function, can lead to distinct differences in patient responses.[15] Patients may place more importance on a sense of normative functioning of a joint rather than their abilities relative to a specific health domain (ie, PF, PI, SE).

Key challenges to using the new PROMIS and GRNF scales are knowing which scale to use and how to best apply them to patient decisions.[15,16] Patient acceptable symptom state (PASS) is a single yes or no question that asks patients to judge whether they are satisfied with their current level of symptoms and activity.[17-22] This validated question is useful to establish what score might be important for patients to achieve on the PROMIS and GRNF scales to be satisfied with their care.[23] Previous studies of disease-specific scales suggest a PASS level is associated with normative values or slightly worse.[19,22] Similar PASS thresholds are not known for PROMIS or GRNF. PROMIS PF and PI are expected to predict PASS status because their constructs are consistent with measuring activity and symptoms. It is unclear whether PROMIS scales unrelated to symptoms or activity, such as PROMIS SE, will also predict a patient's PASS state.[18,19] The ability of PROMIS SE to predict PASS status would support its application to patient care. Identifying the threshold for PASS Yes and No on the PROMIS and GRNF scales would help providers applying these scales to practice understand when patients reach an acceptably low symptom and high activity state or when symptoms and activity are unacceptably high and low, respectively.[18,19,21,22]

The first purpose of this study was to determine if selected PRO scales (PROMIS PF, PI, SE, and GRNF) scales are able to accurately classify an individual's PASS status shortly (<45 days) after collaborative physical therapist (PT) and medical doctor (MD) primary care service. The short-term follow-up (<45 days) was emphasized because this corresponds to when patients may decide to seek further care if recovery is slow. The second purpose was to improve interpretation of each scale by evaluating 2 thresholds providers may use to determine PASS status. Although a single threshold per scale is optimal, it was anticipated that participants at the midrange of the scales would not classify accurately.[8,10,24] Therefore, 2 criteria based on high (95%) and moderate (80%) sensitivity/specificity were compared to determine if focusing on scores that were either higher or lower than the midrange might accurately classify participants. The hypothesis was that high sensitivity/specificity (95%) would result in higher accuracy for a lower proportion of participants. A priori it was unclear if a moderate sensitivity or specificity (80%) might achieve adequate accuracy while identifying a higher proportion of participants.

List of abbreviations:

AUC	**area under the curve**
GRNF	**global rating of normal function**
PASS	**patient acceptable symptom state**
PF	**physical function**
PI	**pain interference**
PROMIS	**Patient-Reported Outcome Information System**
SE	**self-efficacy**

Methods

Overview

This cross-sectional study surveyed patients by phone after their primary care visit for a musculoskeletal problem. The survey included PASS, GRNF, and PROMIS outcome scales. The analysis focuses on the accuracy and proportion of patients classified for PASS using the GRNF and PROMIS scales.

Participants

Patients who attended a rural primary care service from October 2016 to December 2016 were surveyed 7-42 days (average, 13.8±7.3 days) after their primary care encounter (table 1). As part of their initial intake patients agreed to receive a call back to assess the service they received. This agreement included a signed consent that complied with a protocol approved by George Fox University. The goal was to call patients within 30 days of their clinic visit. There were no other inclusion and/or exclusion criteria applied.

The collaborative PT and MD primary care service consisted of evaluation and treatment during the primary care MD visit. The PT service included consultation for (1) diagnosis, (2) need for imaging, and/or (3) referral to another provider. In addition, nearly all patients received minimal treatment including manual therapy (~30%), exercise (>90%), and education (>90%). Time with patients varied from a few minutes to 45 minutes depending on clinical presentation and the patient's needs.

Measures

The CAT PROMIS PF, PI, and SE scales were used to assess patients at follow-up. All CAT PROMIS scales were scored using the web-based assessment center (https://assessmentcenter.net/) website. The CAT PROMIS PF version 1.2 asks patients to rate their level of difficulty performing functional activities. The CAT approach selects appropriately difficult activities, avoiding floor and ceiling effects, in less time than traditional tests.[3,6,24] Similarly, the CAT PROMIS PI scale asks patients to rate the degree to which pain interferes with aspects of their life. For both scales a T score of 50 is the average of the US population, and 10 points represents 1 SD. For PROMIS PF higher scores indicate improved PF and for the PI scale lower scores indicate less PI. The CAT PROMIS SE asks patients to rate their confidence in their ability to manage symptoms. This PROMIS SE scale demonstrated concurrent validity with established scales in patients with a variety of medical conditions.[10] Although the CAT PROMIS SE also determines a T score (higher scores indicate higher SE), the

Table 1 Comparison of demographic and patient-reported outcomes between PASS Yes and PASS No participants

	PASS No (n=49)	PASS Yes (n=45)	*P* Value
Demographic Variables			
Age (y)	61.7 (15.8)	65.4 (15.3)	.25*
Sex (% female)	67.3	66.7	.56†
Height (cm)	168.1 (10.2)	167.2 (9.2)	.74*
Weight (kg)	84.9 (25.4)	86.4 (19.3)	.68*
Body mass index (kg/cm^2)	30.9 (8.2)	31.8 (6.7)	.55*
PROMIS PI	60.1 (6.1)	54.8 (7.1)	<.01‡
PROMIS PF	39.6 (6.5)	47.0 (10.5)	<.01‡
PROMIS SE: Symptom Management	43.6 (4.5)	50.0 (10.5)	<.01‡
GRNF (n=93)	4.9 (1.8)	6.8 (2.0)	<.001*

* Independent *t* test.
† Chi-square test of proportions.
‡ Pairwise comparison (2-way analysis of variance).

reference population consists of patients managing a spectrum of chronic health issues.[10] A T score of 50 represents the average of people managing chronic health conditions, and 10 points is 1 SD. A GRNF rating asks patients to rate their joint and/or body region relative to healthy joints and/or body regions. A global rating is used widely in psychology to capture a broad judgement from the patient regarding various attributes.[14] Similarly, they are used in rehabilitation scales to capture healthy function.[12,25] In this study, participants were asked to rate their joint (eg, knee) or body region (eg, lower back) relative to healthy function. The caller altered the joint or body region (table 2) for the primary or treated problem determined from the medical record. For patients with multiple problems (24%) the patient answered relative to their primary problem.

"How would you rate the function of your _____ [*Fill in joint problem*] on a scale of 0 to 10 with 10 being normal, excellent function, and 0 being the inability to perform any of your usual daily activities, which may include sports?"

The PASS question was derived from previous studies that sought to define when patients reached a point of symptoms and activity that they judged satisfactory.[17] A common wording to define a PASS state was, "Taking into account all the activities you do during your daily life, your level of pain, and also your function, do you consider your current state satisfactory?"[17] A PASS Yes state is consistent with low levels of pain and moderate levels of function on other PRO scales that approximate normative values or slightly worse than normative in patients with musculoskeletal problems.[19,22]

Table 2 The mean ± SD for each PROMIS health domain by body region

Body Region (Percentage of Sample)	PE	PF	SE
Spine (back/neck) (35.3%)	57.6±7.7	43.1±7.8	45.8±6.5
Shoulder (14.1%)	57.7±6.3	42.1±11.3	44.6±5.9
Knee (14.1%)	57.7±7.0	42.3±9.2	47.7±8.9
Hip (10.6%)	53.9±8.1	49.4±15.2	49.8±9.9
Other (1.1%)*	64.2	45.7	42.4
>1 body region (24.7%)	58.3±5.6	43.5±8.1	48.2±6.9

* For this category n=1 patient, so no SDs are given.

Data analysis

For the first purpose, receiver operator curves were used to determine accuracy (ie, area under the curve [AUC]) for each PRO scale to predict PASS. An AUC >0.7 is considered reasonably accurate and, therefore, was adopted as a minimum.[26] For sample size, a prevalence of 30% PASS Yes responses was used. A sample of 90 participants at a 95% CI enables the detection of an AUC as low as .57. To evaluate if a single threshold was useful, the shortest distance to no errors (sensitivity/specificity=1) on the receiver operator curve was assessed. A clear minimum indicates an optimal threshold that balances errors (ie, relatively equal false positives and/or negatives). When there is no clear minimum (ie, flat region) (fig 1), this indicates consistent errors over this range of the scale. No clear minimum suggests there is no optimal single threshold, but rather that using high and low thresholds (ie, outside the flat region) may demonstrate the potential of the scale for a proportion of participants outside the region where errors are higher (ie, flat region). Because the midrange of the PROMIS

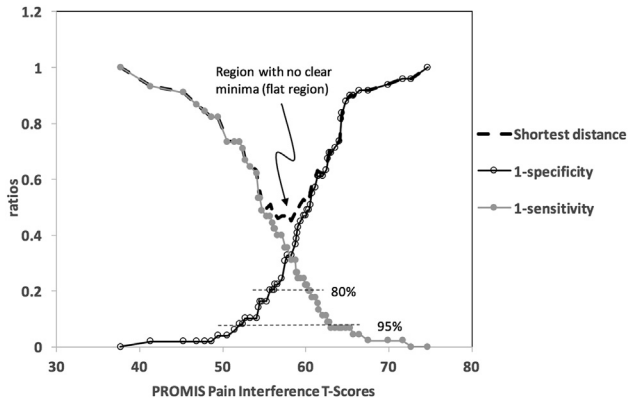

Fig 1 The ratio of false positives (1-specificity) (gray filled circles) and false negatives (1-sensitivity) (black open circles) for determining a participant's PASS from their PROMIS PI scale is shown. The shortest distance to no errors (black hashed) is flat at a ratio of 0.45. The 80% and 95% sensitivity/specificity criteria for determining thresholds for PASS are shown. The region that shows poor ability to discriminate a participant's PASS state is illustrated by the flat region with no clear minima.

Table 3 Results of receiver operator curve analysis

	Area Under the Curve	*P* Value
Scale/Question		
PROMIS PI	.73 (.63-.83)	<.001
PROMIS PF	.72 (.62-.83)	<.001
PROMIS SE: Symptom Management	.76 (.66-.86)	<.001
GRNF (n=93)	.74 (.64-.84)	<.001

scales corresponded to a flat region, thresholds closest to 80%/95% sensitivity/specificity were identified and used to classify participants. Participants that showed better scores than the 80%/95% specificity thresholds were classified as PASS Yes. Participants that showed worse scores than the 80%/95% sensitivity thresholds were classified as PASS No. Participants between the 80%/95% sensitivity/specificity thresholds were identified as unclassified.

To assess the clinical use of the thresholds (95%/80% sensitivity/specificity) the accuracy and proportion of patients classified were evaluated using chi-square analysis. For each PRO scale 3×2 tables were assessed for each classification. The rows of each 3×2 table were patients classified by PRO as PASS Yes, PASS No, and PASS unclassified, and the columns were the participants PASS Yes and No responses. This resulted in 8 3×2 tables, 1 for each PRO scale using the 2 different thresholds. The accuracy and proportions of patients classified were used to understand clinical impact of using the reported PASS thresholds in practice. An alpha level of .05 and SPSS v23 was used for all analyses.

Logistic regression was used to assess the potential for age, sex, body mass index (BMI), and length of follow-up to confound the prediction of PASS status. For each classification of PASS (n=8) PRO variables were converted to a nominal variable using the derived thresholds (PASS Yes=1, 0 and PASS No=1, 0). Logistic regression models for each PRO were then completed (n=16) with the nominal variable PRO score, age, sex, BMI, and length of follow-up entered in each model. If the PRO nominal variable was significant, this suggested the PRO is an independent predictor of PASS.

Results

There were approximately equal numbers of PASS Yes (n=45) and PASS No (n=49) participants (see table 1). The spectrum of musculoskeletal problems included most body regions (see table 2). A total of 169 patients were called; 112 were reached by phone, and 94 completed surveys. One patient did not complete the GRNF rating. Only 18 patients declined to participate when reached, yielding an 83.9% (94/112) participation rate, accounting for 55.6% (94/169) of all calls attempted. The total participation of all patients seen during this time period was 22.2% (94/405).

The AUC for all of the PROMIS scales and GRNF were statistically significant and higher than 0.7 (table 3). The highest AUC was for PROMIS SE (.76; $P<.01$), and the lowest AUC was for PROMIS PF (.72; $P<.01$). The closest thresholds to 95% specificity/sensitivity (PASS Yes/PASS No, respectively) were ≤51.4/63.2≥ for PROMIS PI, ≥50.6/≤33.7 for PROMIS PF, and ≥50.80/≤41.0 for PROMIS SE (table 4). The closest thresholds to 80% specificity/sensitivity (PASS Yes/PASS No, respectively) were ≤56.2/60.4≥ for PROMIS PI, ≥45.4/≤39.8 for PROMIS PF, and ≥47.6/≤43.4 for PROMIS SE (see table 4). The closest threshold to 95% specificity/sensitivity (PASS Yes/PASS No, respectively) below 100% for GRNF was 8≥/≤3. The closest threshold to 80% specificity/sensitivity (PASS Yes/PASS No, respectively) for GRNF was 7≥/≤4.

The 8 chi-square analyses for all PROs were significant (table 5 and 6). The combined accuracy for classifying participants as PASS Yes or No using the 95% specificity/sensitivity criteria ranged from 75%-79.4%. The proportion of participants classified as PASS Yes or No ranged from 25.6%-42.0% (see table 5). The combined accuracy for classifying participants as PASS Yes or No using the 80% specificity/sensitivity criteria ranged from 72.3% to 73.5% (see table 6). The proportion of participants classified as PASS Yes or No ranged from 69.9% to 75.5%.

Logistic regression analysis showed that all PROs evaluated remained significant ($P<.05$) predictors of PASS irrespective of possible confounding factors included in the model.

Discussion

The key findings are that the PROMIS and GRNF scales are able to characterize PASS status with moderate accuracy (~70%) for a large portion of patients (~70%). Also, PROMIS SE, which is not associated with physical abilities or symptoms and GRNF (which references normative function), predicts PASS status as well as PROMIS PF and PI (see table 3). The thresholds identified characterize PASS Yes status as at or slightly worse than the US population. However, the accuracy and proportion of patients successfully classified into a PASS category may influence how these scales are used. The 80% specificity/sensitivity criteria identified thresholds that achieved 72.3%-73.5% accuracy for

Table 4 Thresholds established using the 95% specificity/sensitivity and 80% specificity/sensitivity

	Higher Thresholds (95%)				Moderate Thresholds (80%)			
	Specificity / Sensitivity (%)	PASS No	Not Classified	PASS Yes	Specificity/ Sensitivity (%)	PASS No	Not Classified	PASS Yes
Scale/Question								
PROMIS PI	93.3/93.9	≥63.2	51.5-63.1	≤51.4	80.0/79.6	≥60.4	56.3-60.3	≤56.2
PROMIS PF	93.3/93.9	≤33.7	33.8-50.5	≥50.6	79.6/80.0	≤39.8	39.9-45.3	≥45.4
PROMIS SE: Symptom Management	89/93.9	≤40.9	41.0-50.8	≥50.9	79.6/82.2	≤43.4	43.5-47.5	≥47.6
GRNF (n=93)	88/95.5	≤3	4-7	≥8	79.6/81.8	≤4	5-6	≥7

Table 5 Proportions of patients classified at a PASS Yes and PASS No by threshold derived from the nearest values 95% specificity/sensitivity

PROMIS		PASS Yes, No. (%)	PASS No, No. (%)	Total, No. (%)	Accuracy (%)	*P* Value*
PI						
PASS No	≥65.0	3 (3.2)	6 (6.4)	9 (9.6)	66.7	ND
Unclassified	51.5-64.9	30 (31.9)	40 (42.6)	70 (74.5)	NA	ND
PASS Yes	≤51.4	12 (12.8)	3 (3.2)	15 (16.0)	80.0	.02
Total		45 (47.9)	49 (52.1)	94 (100.0)	NA	ND
PF						
PASS No	≤33.8	3 (3.2)	8 (8.5)	11 (11.7)	72.3	ND
Unclassified	33.9-50.4	26 (27.7)	38 (40.4)	64 (68.1)	NA	ND
PASS Yes	≥50.5	16 (17.0)	3 (3.2)	21 (20.2)	76.2	<.01
Total		45 (47.9)	49 (52.1)	94 (100.0)	NA	ND
SE						
PASS No	≤41.0	5 (5.3)	14 (14.9)	19 (20.2)	73.6	ND
Unclassified	41.1-51.0	24 (25.6)	32 (34.0)	46 (59.6)	NA	ND
PASS Yes	≥51.1	16 (17)	3 (3.2)	19 (20.2)	84.2	<.01
Total		45 (47.9)	49 (52.1)	94 (100.0)	NA	ND
Other						
GRNF (n=93)						
PASS No	≤3	2 (2.2)	11 (11.8)	13 (14.0)	84.6	ND
Unclassified	4-7	22 (23.7)	32 (34.4)	54 (58.1)	NA	ND
PASS Yes	≥7	20 (21.5)	6 (6.5)	26 (28.0)	76.9	<.01
Total		44 (47.3)	49 (52.7)	93 (100.0)	NA	ND

Abbreviations: NA, not applicable; ND, no data.
* *P* value from chi-square analysis.

69.9%-75.0% of participants. There was little advantage to applying the higher 95% specificity/sensitivity threshold criteria. Clinically, this is encouraging for the use of patient-derived data (ie, PROMIS scales) to be used to track outcomes; however, how to identify all patients' PASS status accurately was unresolved in this study.

Table 6 Proportions of patients classified at a PASS Yes and PASS No by threshold derived from the nearest values 80% specificity/sensitivity

PROMIS		PASS Yes, No. (%)	PASS No, No. (%)	Total, No. (%)	Accuracy (%)	*P* Value*
PI						
PASS No	≥60.4	9 (9.6)	25 (26.6)	34 (36.2)	73.5	ND
Unclassified	56.3-60.3	10 (10.6)	14 (14.9)	24 (25.5)	NA	ND
PASS Yes	≤56.2	26 (27.7)	10 (10.6)	36 (38.3)	72.2	<.01
Total		45 (47.9)	49 (52.1)	94 (100.0)	NA	ND
PF						
PASS No	≤39.8	9 (9.6)	29 (30.9)	38 (40.4)	76.3	ND
Unclassified	39.9-45.3	13 (13.8)	10 (10.6)	23 (24.5)	NA	ND
PASS Yes	≥45.4	23 (24.5)	10 (10.6)	33 (35.1)	69.7	<.01
Total		45 (47.9)	49 (52.1)	94 (100.0)	NA	ND
SE						
PASS No	≤43.4	8 (8.5)	25 (26.6)	33 (35.1)	75.8	ND
Unclassified	43.5-47.5	12 (12.8)	14 (14.9)	26 (27.7)	NA	ND
PASS Yes	≥47.6	25 (26.6)	10 (10.6)	35 (37.2)	71.4	<.01
Total		45 (47.9)	49 (52.1)	94 (100.0)	NA	ND
Other						
GRNF (n=93)						
PASS No	≤4	8 (8.6)	20 (21.5)	28 (30.1)	71.4	ND
Unclassified	5-6	9 (9.7)	19 (20.4)	28 (30.1)	NA	ND
PASS Yes	≥7	27 (21.5)	10 (6.5)	37 (39.8)	72.9	<.01
Total		44 (47.3)	49 (52.7)	93 (100.0)	NA	ND

Abbreviations: NA, not applicable; ND, no data.
* *P* value from chi-square analysis.

A challenge with documenting outcomes in primary care is the variety of diagnoses (see table 2). Despite a wide range of diagnoses, the sample was nearly evenly split between PASS Yes and No participants (see table 1). On average, patients that were PASS No (see table 1) were experiencing low PF (39.6) and high PI (60.1). PROMIS SE was also low (43.6), suggesting patients had low confidence in their ability to manage their symptoms. Overall, patients rated themselves as less than 5 out of 10 on the GRNF. In contrast, average PROMIS scores were near 50 for PASS Yes (within 5 points of 50), and GRNF was 7 out of 10. This suggests that on average a PASS Yes score is near the US population average (PROMIS scales) or 70% of normal (GRNF).

Consistent with previous studies, characterization of PASS Yes was near the average of the US population or slightly worse (see table 4).[19,22] PASS Yes status (95% specificity/sensitivity) was characterized by PI, PF, and SE scores within 1.4 points of a T score of 50. The moderate criteria (80% specificity/sensitivity) identified worse scores on the PROMIS scales by up to 5.1 points and 1 point on the GRNF scale. This suggests patients judge themselves as PASS Yes when they reach average or slightly worse on specific health domains (PF, PI) and relative to "normal" (ie, GRNF). New to this study is that people with average or slightly lower SE are more likely to rate themselves as PASS Yes compared with people with chronic conditions. This alerts providers to the influence of SE as well as PF and PI to achieve a positive patient outcome. The corresponding rating relative to healthy (GRNF) was 7-8 out of 10 or above. Therefore, if therapeutically reasonable, clinicians may seek to set goals to achieve PF, PI, and SE at or slightly worse than average and a GRNF of 7 out 10 or above.

In contrast to a positive patient outcome, a PASS No categorization was associated with worse health domain scores on PROMIS and GRNF (see table 4). Using the 95% specificity/sensitivity criteria led to a high PI (>65.0), low PF (<33.8), and low SE (<41.0) to characterize PASS No status. The 80% specificity/sensitivity criteria led to less severe scores (see table 4). Nevertheless, both criteria indicate significant problems with function (ie, PF), high pain (ie, PI), and low SE as characterizing PASS No. Similarly, a PASS No threshold for GRNF was 30%-40% (3-4/10) of normal. Patients meeting these thresholds likely require treatment or ongoing monitoring targeting a particular health domain.

The proportions of patients accurately classified suggests the potential impact of using these thresholds clinically (see table 5 and 6). The 95% specificity/sensitivity criteria was explored to determine the ability of PRO scales to contribute to more definitive clinical decisions.[8] However, the 95% specificity/sensitivity criteria accuracy (75.0%-79.4%) was similar to the accuracy (72.3%-75.5%) using the 80% specificity/sensitivity criteria. The minimal decrease in accuracy using the 80% sensitivity/specificity thresholds was offset by a much larger proportion of patients classified (>69.9%), making these thresholds more useful clinically (see table 6). Although higher accuracy is desirable, the 72.3%-75.5% accuracy is impressive given the diversity of the sample and that PASS status was based solely on a PRO (ie, PROMIS scales or GRNF). Typically, clinicians might use several different factors (patient interview and clinical exam) to determine PASS status.[20]

This study suggests that providers may gain a better impression of patient status by including SE and GRNF as well as PF and PI health domains typically assessed through interview. Whether accuracy for all patients could be improved by combining scales (ie, multivariate analysis) remains unresolved. New to this study, the PASS No thresholds provided suggest that when severity is significant enough patients likely benefit from a change in treatment or referral. Further studies are necessary to determine whether patient-derived general health scales like PROMIS empowers providers to confidently act on patient needs.

Study limitations

There are several limitations to this study. Although the selected scales are generic, the sample diagnoses are diverse (see table 2). Prediction of PASS may vary by diagnosis, which is not addressed here. Also, multiple health domains, some distinct from the construct of PASS, may influence a patient's PASS status. However, this analysis considered each health domain as a single predictor of PASS. Multivariate models may find more optimal combinations of health domains and other factors as important predictors of PASS.[18,20] Although the logistic regression analysis shows that potentially confounding factors of age, sex, BMI, and length of follow-up did not influence PRO predictions in this sample, larger more varied samples are necessary to assess whether these variables are significant predictors of PASS.[18,20] Finally, to validate that the identified thresholds generalize to other samples, the classification proposed here should be applied to a separate set of patients.[8]

Conclusions

This analysis shows that PROMIS and GRNF scales were able to characterize PASS status with moderate accuracy (~70%) for a large portion of patients (~70%). New to this study is the association of SE and GRNF with PASS status. PASS was characterized as at or slightly worse than the US population average. The association of the PROMIS and GRNF scales support their use as possible outcomes to document patient status after primary care for musculoskeletal problems.

Keywords

Orthopedics; Pain interference; Patient reported outcome measure; Physical function; self-efficacy; Rehabilitation

Corresponding author

Jeff Houck, PT, PhD, Physical Therapy Department, George Fox University, 505 Villa Rd, Newberg, OR 97132. *E-mail address:* jhouck@georgefox.edu.

Acknowledgments

We thank the collaborative MD/PT primary care service provided by Yamhill County Coordinated Care Organization Innovation Grant and Providence Health and Services. We acknowledge the outstanding patient service provided by Li-Zandre Van Eeden, DPT (a treating therapist of the service), and Catherine Vandehaar and Andrea Mulligan, who made all the follow-up phone calls to collect the patient-reported outcomes.

References

1. Broderick JE, DeWitt EM, Rothrock N, Crane PK, Forrest CB. Advances in patient-reported outcomes: the NIH PROMIS(®) Measures. EGEMS (Wash DC) 2013;1:1015.
2. Hong I, Velozo CA, Li CY, Romero S, Gruber-Baldini AL, Shulman LM. Assessment of the psychometrics of a PROMIS item bank: self-efficacy for managing daily activities. Qual Life Res 2016;25:2203-14.
3. Hung M, Baumhauer JF, Latt LD, Saltzman CL, Soohoo NF, Hunt KJ. Validation of PROMIS® physical function computerized adaptive tests for orthopaedic foot and ankle outcome research. Clin Orthop Relat Res 2013;471:3466-74.
4. Rose M, Bjorner JB, Gandek B, Bruce B, Fries JF, Ware JE. The PROMIS Physical Function item bank was calibrated to a standardized metric and shown to improve measurement efficiency. J Clin Epidemiol 2014;67:516-26.
5. Cook KF, Jensen SE, Schalet BD, et al. PROMIS measures of pain, fatigue, negative effect, physical function and social function demonstrate clinical validity across a range of chronic conditions. J Clin Epidemiol 2016;73:89-102.
6. 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;35:592-9.
7. 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;63:1179-94.
8. Ho B, Houck JR, Flemister AS, et al. Preoperative PROMIS scores predict postoperative success in foot and ankle patients. Foot Ankle Int 2016;37:911-8.
9. Papuga MO, Beck CA, Kates SL, Schwarz EM, Maloney MD. Validation of GAITRite and PROMIS as high-throughput physical function outcome measures following ACL reconstruction. J Orthop Res 2014;32:793-801.
10. Gruber-Baldini AL, Velozo C, Romero S, Shulman LM. Validation of the PROMIS® measures of self-efficacy for managing chronic conditions. Qual Life Res 2017;26:1915-24.
11. Frogner BK, Harwood K, Pines J, Andrilla H, Schwartz M, Washington TG. Does unrestricted direct access to physical therapy reduce utilization and health spending? Health Care Cost Institute and National Academy for State Health Policy State Health Grant Program. Washington, DC: Health Care Cost Institute, January 2016.
12. Irrgang JJ, Anderson AF, Boland AL, et al. Development and validation of the international knee documentation committee subjective knee form. Am J Sports Med 2001;29:600-13.
13. Martin RLR, Irrgang JJJ, Burdett RG, Conti SF, Van Swearingen JM. Evidence of validity for the Foot and Ankle Ability Measure (FAAM). Foot Ankle Int 2005;26:968-83.
14. Aas IHM. Guidelines for rating Global Assessment of Functioning (GAF). Ann Gen Psychiatry 2011;10:2.
15. Noble PC, Dwyer M, Brekke A. Commonalities, differences, and challenges with patient-derived outcome measurement tools: function/activity scales. Clin Orthop and Relat Res 2013;471:3457-65.
16. 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;471:3437-45.
17. Kvien TK, Heiberg T, Hagen B. 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):40-1.
18. Salaffi F, Carotti M, Gutierrez M, Carlo MD, Angelis RD. Patient acceptable symptom state in self-report questionnaires and composite clinical disease index for assessing rheumatoid arthritis activity: identification of cut-off points for routine care. Biomed Res Int 2015;2015:930756.
19. Tashjian RZ, Deloach J, Porucznik CA, Powell AP. Minimal clinically important differences (MCID) and patient acceptable symptomatic state (PASS) for visual analog scales (VAS) measuring pain in patients treated for rotator cuff disease. J Shoulder Elbow Surg 2009;18:927-32.
20. Wright AA, Hensley CP, Gilbertson J, Leland JM, Jackson S. Defining patient acceptable symptom state thresholds for commonly used patient reported outcomes measures in general orthopedic practice. Man Ther 2015;20:814-9.
21. Levy DM, Kuhns BD, Chahal J, Philippon MJ, Kelly BT, Nho SJ. Hip arthroscopy outcomes with respect to patient acceptable symptomatic state and minimal clinically important difference. Arthroscopy 2016;32:1877-86.
22. 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;1:9-11.
23. Andersen N-bDV, Kent P, Hjort J, Christiansen DH. Clinical course and prognosis of musculoskeletal pain in patients referred for physiotherapy: does pain site matter? BMC Musculoskelet Disord 2017:1-11.
24. Hung M, Franklin JD, Hon SD, Cheng C, Conrad J, Saltzman CL. Time for a paradigm shift with computerized adaptive testing of general physical function outcomes measurements. Foot Ankle Int 2014;35:1-7.
25. Martin RL, Irrgang JJ. A survey of self-reported outcome instruments for the foot and ankle. Journal Orthop Sports Phys Ther 2007;37:72-84.
26. Gortmaker S, Hosmer D, Lemeshow S. Applied logistic regression. Contemp Sociol 1994;23:159.