RESEARCH ARTICLE

Do patients judge success of treatment and patient acceptable symptom state based on current self-reported health status?

Authors

Ryan P. Jacobson, DPT, PCS. (Corresponding author)
George Fox University, Newberg, Oregon, United States of America
rjacobson@georgefox.edu

Daniel Kang, DPT
George Fox University, Newberg, Oregon, United States of America
dkang@georgefox.edu

Jeff Houck, PhD, PT
George Fox University, Newberg, Oregon, United States of America
jhouck@georgefox.edu

Abstract

Background: Value-based care models call for better interpretation of patient-reported outcomes. Patients may reference health status differently when appraising if an intervention was successful versus if their current state is acceptable. The purpose of this study was to determine the association between success of treatment (SOT), patient acceptable symptom state (PASS), and PROMIS measure T-scores, following a single primary care physical therapy encounter.

Methods: Cross-sectional study. Ninety-two patients with musculoskeletal complaints were administered standard SOT and PASS questions, and PROMIS Physical Function, Pain Interference, and Self-Efficacy for Managing Symptoms measures. Association of PASS and SOT was determined using rank biserial correlation and chi-squared analysis. Accuracy of PROMIS T-scores to discriminate PASS and SOT was determined through receiver-operator curve analysis and likelihood ratios.

Results: There was significant association between PASS and SOT ($r=.393$, $p<0.001$; $X^2=15.7$, $p=0.001$). The three PROMIS measures discriminated PASS with AUCs of 0.73 to 0.88 (accuracy 67.3% to 82.6%), Self-Efficacy being the strongest discriminator. Only Pain Interference T-scores discriminated SOT with AUC >0.70 (accuracy 76.1%).

Conclusion: PASS was more strongly associated with health status than SOT. Patients make a meaningful distinction between these two questions. Accurate clinician interpretation of PASS, SOT, and PROMIS T-scores can allow more targeted goal setting and treatment decision making.

Keywords: patient reported outcome measures, PROMIS, primary care, physical therapy, musculoskeletal disease

1. Introduction

Patient-reported outcome (PRO) measures are being more widely adopted in standardized ways across health systems.[1] A key reason is that value-based care models call for better definitions of outcomes that include measures of self-reported health status from patients.[2,3] Although PRO measures are only one aspect of value-based care outcomes assessment, PRO data is a critical part of this assessment for healthcare decision making.[4] Interpreting this data from patients accurately is important for guiding clinical decision making toward meaningful health outcomes.

The World Health Organization defines *health* as "a state of complete physical, mental, and social well-being, and not merely the absence of disease" (https://www.who.int/about/who-we-are/frequently-asked-questions). *Self-reported health status*, then, encompasses the patient's perception of how well they are functioning physically, mentally, and socially, as well as their current symptom burden.[5-7] It has been proposed in the literature that when patients judge the outcome of a healthcare intervention, current health status is an important factor tied to well-being.[8-11] Various single-item PRO questions have been used to appraise patient perception of outcome. A *success of treatment* (SOT) question can be asked to capture patients' current point-in-time perception of success in reference to a recent healthcare intervention.[12,13] This is distinct from global rating of change scales which are meant to specifically quantify patient-perceived improvement or deterioration, and which have limitations around patients' ability to accurately recall previous health status and to quantify their own degree of change.[14] Meanwhile, the *patient acceptable symptom state* (PASS) question is used to demarcate the level of function and symptoms beyond which a patient considers their state acceptable.[12,15,16]. Both the SOT and the PASS questions have potential to be useful tools in clinical practice for quick interpretation of post-treatment PROs.[17] However, it is unclear to what degree patients reference health status—that is their own physical, mental, and social functioning, along with symptom burden—in answering each of these questions.

When judging SOT and/or PASS, patients may consider different factors beyond just health status, influencing their responses when asked to self-appraise the treatment outcome.[12,18] Indeed, it has been suggested that patient appraisal of SOT might be referenced to a more broad range of factors beyond just health status.[12,19] An example would be the potential influence of expectations of care established between the clinician and patient during their episode of care.[20,21] For example, patients' pre-surgery expectations have been associated with patient appraisal of outcome following orthopedic surgery.[22] Excluding factors like expectations, it is unclear if patients reference self-reported health status differently when answering SOT versus PASS questions. One study showed a strong association between PASS "Yes" status and patient judgement that their surgery was a success. However, this same study showed that PASS "No" status was a poor discriminator of patients who considered their surgery a failure. Judging their treatment a success, these patients may have emphasized expectations established with their provider prior to surgery over health status.[12] Whatever the cause, if there is a disconnect between SOT and PASS also occurring after conservative care, then this is important when considering how to define value in new models of care. If SOT is capturing the patient's judgement specific to the treatment (e.g. emphasizing expectations), then this outcome might be more appropriate to evaluate a service or episode of care. If the PASS question is more associated with health status, then this outcome would be more appropriate to reflect how well they are doing related to their current functioning and/or symptom burden.

To determine the degree to which SOT and PASS are each associated with specific measures of self-reported health status, person-centered measures specific to biopsychosocial health are ideal. The Patient-Reported Outcomes Measurement Information System (PROMIS) computer adaptive measures provide a means to assess specific areas of self-reported health status. The PROMIS includes publicly available person-centered measures of physical, mental, and social health for clinical research and patient care. Clinically, in contrast with more common disease-specific scales, PROMIS measures are person-centered and typically referenced to the general US population. Disease-specific scales capture responses with respect to a specific condition, for example low back pain or knee osteoarthritis, and typically do not reference an average for the US population.[23-26] Additionally, PROMIS measures were developed using modern measurement methods including item-response theory. Measures developed using this methodology have many advantages, including the basis to be administered via computer adaptive testing so that patient response burden may be minimized (<1 minute per measure).[23,27]

Two specific health domains commonly used to define health status in patients with musculoskeletal problems are pain interference and physical function.[24,26] If patient testing burden were minimized for these, it would then be possible clinically to also assess psychological variables such as self-efficacy.[28] Self-efficacy is viewed as a key variable associated with behavior change across medical conditions, including musculoskeletal,[29] cancer,[30] and heart disease.[31] While some approaches focus on selected psychological variables based on the fear avoidance model,[32] more general positive psychological variables like self-efficacy are integrated into social cognitive theory and applied to many different health behaviors including musculoskeletal conditions.[29,33] Therefore, together the PROMIS measures for Physical Function (PF), Pain Interference (PI), and Self-Efficacy for Managing Symptoms (SE) represent a potentially useful suite of PROs for assessing health status in patients with musculoskeletal conditions, one that might also discriminate two constructs (i.e. SOT and PASS) that are important to providers when defining patient outcome.

Some studies have investigated the association of health status with SOT and PASS in musculoskeletal patients; however, few studies were found that assessed both constructs simultaneously.[34,35] One study demonstrated that PROMIS PF, PI, and SE measures discriminated PASS status with >70% accuracy in primary care at the initial assessment.[17] Another study showed an association between PASS and a disease-specific measure at discharge from physical therapy.[36] These studies support the hypothesis that PASS and health status are associated, but SOT was not assessed in these studies. An analysis of SOT and PASS together was not found specific to physical therapy. However, one study that analyzed both SOT and PASS after orthopedic surgery demonstrated that PROMIS PF, PI and Depression scores were moderately associated with general patient-appraisal of treatment success.[12] The SOT question asks patients to judge how successful their recent treatment was,[13] potentially inviting a broad range of factors to be considered other than health status. Thus, it is plausible when a clinician queries a patient on success of treatment that the patient's response may be positive or negative, unrelated to a change in health status. Because the PASS question asks patients to judge whether their daily activity, pain, and functional impairment are at an acceptable level,[12,16] it is hypothesized to more likely align with self-reported health status as measured using PROMIS. Clinically, this distinction could be important when interpreting patient responses. Similar to surgeons, when physical therapists

discuss progress with patients, if they desire an assessment of health status, then a question like PASS should be prioritized over questions focused on treatment success. Based on previous studies outside of physical therapy care,[12,16] patients may answer differently relative to current health status when asked SOT and PASS single-item questions. As providers are required to make judgements about the value of health care they deliver, and anticipate the potential need for future utilization, understanding how queries of SOT versus questions like PASS may elicit different responses from patients is useful.

The purposes of this study were: 1) to determine the degree of association between PASS and SOT; and 2) to determine to what extent PROMIS PF, PI, and SE measures are able to discriminate both PASS and SOT status at 45-60 days after a primary care encounter for a musculoskeletal problem. The null hypotheses tested are as follows. Hypothesis 1: That PASS and SOT responses will show no significant correlation. Hypothesis 2: That there will be no chance agreement between the response categories of PASS and those of the SOT question. Hypothesis 3: That PASS will not be discriminated by each measure of health status (i.e. PROMIS PF, PI, and SE). Hypothesis 4: That SOT will not be discriminated by each measure of health status (i.e. PROMIS PF, PI, and SE). Testing these hypotheses will improve understanding of how patient self-reported health status is associated with these two related but potentially distinct questions (i.e. PASS and SOT). Additionally, clinicians might be able to better judge the relative meaning of their patients' responses, interpreting how health status impacts PASS and SOT.

2. Methods

2.1 Study Design

This observational, cross-sectional study collected data over the phone as follow-up from a patient's encounter in primary care. Patients received care for musculoskeletal diagnoses in a collaborative physician/PT primary care service in a rural hospital-based clinic. The treatments provided in the one-time visit included: 1) education (97.5%), 2) exercise (97.5%), 3) hands-on treatment (e.g. manual therapy) (28.9%), and/or 4) recommendation for further physical therapy from an outpatient provider (34.2%).[37] Note that this service provides more physical therapy care than is typically available during a primary car encounter.[38] Patients signed written informed consent to participate in the phone survey in compliance with an approved protocol by the Institutional Review Board at George Fox University, Newberg, Oregon, USA. All patient information was fully anonymized. The SOT, PASS, and PROMIS data were collected at 1-7 days and 45-60 days after the primary care encounter. Only the 45-60 day sample data is reported here, the 1-7 day data being reported elsewhere.[17] The 45-60 day follow up time was based on: 1) prognosis for some common musculoskeletal problems,[39,40] and 2) when follow-up from primary care may occur based on clinical experiences in primary care.[37] The data for this secondary analysis was collected from visits occurring March through August 2017. The inclusion criteria were non-specific. All patients over 18 years old who participated in this collaborative care service during the specified time frame were invited to participate; however, not all patients who received the service were called (see calling details below). There were no other inclusion or exclusion criteria.

2.2 Patient-Reported Outcome Measures

Patients were administered the SOT question, PASS question, and PROMIS PF, PI, and SE measures consistent with previous studies.[17,37] PASS is a validated single-item question meant to demarcate the level of function and symptoms beyond which a patient considers their state acceptable. Patients in this study responded Yes or No to the anchoring

question, "Taking into account all the activities you have during your daily life, your level of pain, and also your functional impairment, do you consider that your current state is satisfactory?"[12,15,16,36] The SOT question is meant to capture patient judgment about a recent treatment for a given condition and/or symptom.[12] Used previously as a validating question for a patient satisfaction questionnaire, patients in the present study were asked, "How successful was the treatment for your [primary complaint]?" and given the choices "Not Helped," "Improved," "Partly Cured," or "Cured."[13] Unlike the PASS question, the context of this question is treatment success, rather than satisfaction with current activities and symptoms.

The three PROMIS measures used to define health status were administered via computer adaptive testing using the HealthMeasures iPad app (Glinberg & Associates, Inc). All PROMIS measure items use a 5-point Likert scale response set, representing increasing degrees of the trait being measured (i.e. PF, PI, SE).[27,41,42] The Physical Function v1.2 measure comprises questions around functioning in mobility, use of arms and body, and capability in instrumental activities of daily living, with higher T-scores representing better functioning.[43] The Pain Interference v1.0 measure assesses the extent to which pain hinders daily life, with lower T-scores representing less pain interference.[44] Both the PF v1.2 and PI v1.0 measures were calibrated and validated on the general US population, with a T-score mean of 50, standard deviation of 10.[43,44] The Self-Efficacy for Managing Symptoms v1.0 measure assesses confidence in controlling symptoms in work, play, sleep, and relationships, with higher T-scores representing better self-efficacy. The reference population for the SE v1.0 measure is patients with chronic conditions, again with a T-score of mean of 50, standard deviation of 10.[45]

All PRO measures were administered over the phone by paid research staff not otherwise associated with the delivery of clinical care. Phone administration of PROMIS has been validated in a previous study.[46] All calls were directly with patients (no proxies). Callers received the following training/support to obtain valid responses: 1) a standardized phone script was provided with initial practice on mock calls, 2) the initial 5-10 patient calls were conducted under supervision, and 3) intermittent feedback was provided when the standardized approach was difficult to apply (e.g. if patients found it difficult to respond within scripting). The intent of the training was to minimize the influence of the caller on patient responses. Research staff developed a calling schedule during their funded research time. When a patient's scheduled call back time fell outside of the calling windows (1-7 days, 45-60 days) no attempts were made to contact them. This resulted in only a portion of the sample being called.

2.3 Data Analysis

Descriptive statistics for the group of patients were used to describe sample characteristics. To address hypothesis 1, the association of the PASS and SOT responses was calculated using rank biserial correlation—correlations below 0.4 were considered weak and likely of questionable clinical significance.[47] Hypothesis 2 was addressed with analysis of a 4x2 cross-tabs table (rows=4 SOT responses; columns=2 PASS responses) to assess probability of chance agreement using the chi-squared statistic.

For the remaining hypotheses, receiver operator characteristic (ROC) curves were used to determine area under the curve (AUC) for each PROMIS measure's ability to discriminate PASS (hypothesis 3) and SOT (hypothesis 4). The alpha level was adjusted using a Bonferroni correction, dividing the

standard 0.05 by number of comparisons (6) to derive an adjusted alpha level of p≤.008 for significance.[48] An acceptable AUC value was considered 0.70-0.79 with "excellent" values at ≥0.80, which has been described.[49] For this analysis, SOT responses were dichotomized into two categories: "Not Helped" and "Improved, Partly Cured, or Cured" (ImpPCoC) Thus, the ImpPCoC category represented any amount of perceived treatment success. This dichotomization resulted from analysis which showed that of the four possible SOT responses, this approach yielded the highest possible AUC values for the SOT question. Also, this dichotomization (i.e. "Not Helped" vs ImpPCoC) is in alignment with other studies where 4- or 5-point SOT response sets were dichotomized and reported.[9,13] Next, each of the six ROC curves was analyzed using the closest-to-(0,1) criterion,[50] in order to determine T-score thresholds based on optimal sensitivity/specificity values (i.e. T-score cut-off at which sensitivity and specificity are together maximized). These optimal threshold values are the T-score value where better health status is associated with a positive patient response to the PASS or SOT questions. This standardized approach allows for direct comparison across the six thresholds determined for PASS and SOT. Total accuracy of each PROMIS T-score threshold to discriminate PASS and SOT was then calculated through a 2x2 cross-tabs table, with T-scores worse than the optimal threshold representing a negative test. Finally, to control further for prevalence, likelihood ratios with 95% confidence intervals were calculated from the cross-tabs tables. This combined approach resulted in robust side-by-side comparison of AUC and accuracy values, as well as likelihood ratios with confidence intervals, for discriminating PASS and SOT using each of the three PROMIS measures.

Sample size was evaluated by systematically varying the ratio of negative to positive responses on SOT and on PASS (in 5% increments, starting at 10%), using analysis for paired AUC data [51]. This analysis suggested that an AUC as low as 0.57 would be detected with a 25% ratio, at an alpha of .05, for a sample of 92 participants. Because a 0.70 AUC was considered clinically meaningful, this sample size was determined to be adequate [17]. All descriptive and inferential statistical analyses were performed using SPSS (v25, IBM Corporation).

3. Results

Ninety-two primary care physical therapy patient records yielded follow-up testing on all PRO measures. Age ranged from 20-90 years old, with an average age of 56.8 years (16.7), 63.0% female, and an average BMI of 31.0 (7.3). Additional sample characteristics are described in Table 1. A majority of patients reported their status as PASS Yes (56.5%) and that following treatment they were ImpPCoC (75.0%). PROMIS PF T-scores ranged from 28.8 to 63.5, PI T-scores from 38.7 to 76.4, and SE T-scores from 28.2 to 68.7.

Table 1. Sample characteristics (n=92) and patient-reported outcomes scores at 45-60 day follow-up.

Characteristic	Mean (SD) / Count (%)
Age (y)	56.8 (16.7)
Female	58 (63.0%)
Height (cm)	168.0 (10.2)
Weight (kg)	87.7 (23.2)
BMI	31.0 (7.3)
Length of Follow-up (d)	52.8 (4.6)
Primary musculoskeletal complaint	
low back pain	20 (21.7%)
neck / thoracic pain	21 (22.8%)
lower extremity pain	29 (31.5%
upper extremity pain	18 (19.6%)
other	4 (4.3%)
PASS	
PASS No	40 (43.5%)
PASS Yes	52 (56.5%)
SOT	
not helped (NH)	23 (25.0%)
improved (Imp)	42 (45.7%)
partly cured or cured (PCoC)	27 (29.3%)
PROMIS T-scores	
Physical Function (PF)	45.3 (7.9)
Pain Interference (PI)	55.7 (9.3)
Self-Efficacy (SE)	47.0 (7.5)

BMI=body mass index, PASS=patient acceptable symptom state, SOT=success of treatment, PROMIS=Patient-Reported Outcomes Measurement Information System

There was a significant association between PASS and SOT categories (r=.393, p<0.001; hypothesis 1). Further the 4x2 table (Table 2) to assess chance agreement between PASS and SOT (X^2=15.7, p=0.001) revealed good agreement between the Not Helped response on the SOT question and PASS No (17/23 = 73.9%), as well as between Partly Cured and PASS Yes (70%), and Cured and PASS Yes (100%). However, the level of agreement between Improved and PASS Yes was relatively lower at 59.5% (hypothesis 2).

Table 2. Association of PASS and SOT responses at 45-60 day follow-up.

SOT	PASS No	PASS Yes	Total
Not Helped	17	6	23
Improved	17	25	42
Partly Cured	6	14	20
Cured	0	7	7
Total	40	52	92

PASS=patient acceptable symptom state, SOT=success of treatment

The ROC curve analysis (Figure 1) showed significant AUC values for PROMIS measures with both PASS and SOT, as detailed in Table 3. All three AUC values discriminating PASS were significant at the p<.001 level (hypothesis 3). All three AUC values were also ≥0.70, with the PF AUC at 0.73, PI at 0.87, and SE at 0.88. The optimal T-score thresholds for these three measures ranged between 4.4 to 5.0 points worse than the reference mean (about ½ standard deviation), with T-scores better than these thresholds (e.g. PI<55.0) discriminating PASS Yes. Based on these threshold cut-offs, total accuracy for discriminating PASS Yes and PASS No were PF=67.3%, PI=80.4%, and SE=82.6%, with likelihood ratios (95% confidence interval) of 2.07 (1.27-3.36), 6.00 (2.60-13.83), and 3.54 (2.05-6.11), respectively. In contrast, only two of three AUC values discriminating SOT were significant at the adjusted alpha level—PI at p<.001 and SE at p=.006 (hypothesis 4). The only AUC ≥0.70 was PI at 0.77, with SE approaching an acceptable level at 0.69. The optimal T-score thresholds for these two measures ranged between 4.0 and 9.4 points worse than the US mean, with T-scores better than these thresholds (e.g. PI< 59.4) discriminating SOT ImpPCoC. Total accuracy based on these thresholds were PI=76.1% and SE=70.1%, with likelihood ratios (95% confidence interval) of 2.94 (1.46-5.93) and 2.67 (1.32-5.40), respectively.

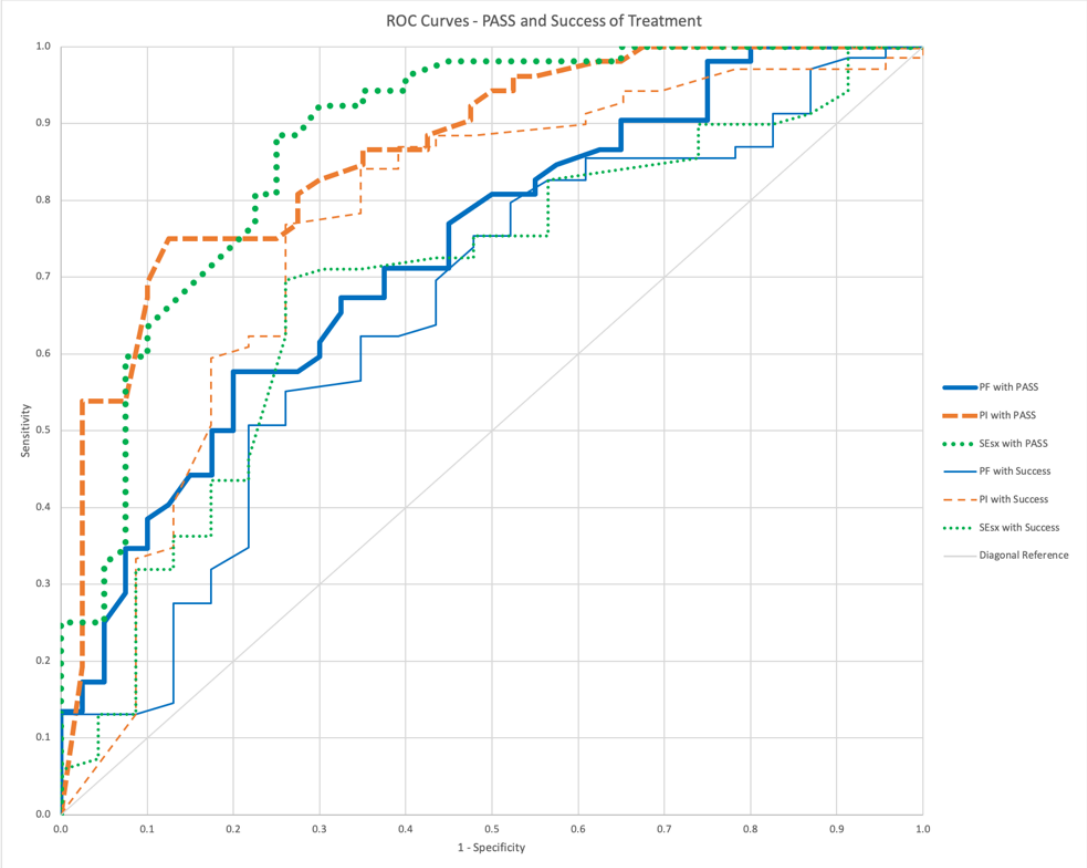

Figure 1. Receiver-operator curves for patient acceptable symptom state (PASS) and success of treatment responses, as discriminated by T-scores for PROMIS Physical Function (PF), Pain Interference (PI), and Self-Efficacy for Managing Symptoms (SE).

Table 3. Area under the curve, likelihood ratios, and related values for each PROMIS measure's ability to predict PASS and SOT responses at 45-60 day follow-up.

Scale Statistic	PASS	SOT
PROMIS Physical Function (PF)		
AUC (95% CI)	0.73 (0.63-0.84)**	0.66 (0.53-0.79)
optimal T-score threshold	45.0	44.1
sensitivity	67.3	62.3
specificity	67.5	65.2
likelihood ratio (95% CI)	2.07 (1.27-3.36)	1.79 (0.99-3.23)
accuracy	67.3%	63.0%
PROMIS Pain Interference (PI)		
AUC (95% CI)	0.87 (0.79-0.94)**	0.77 (0.64-0.89)**
optimal T-score threshold	55.0	59.4
sensitivity	75.0	76.8
specificity	87.5	73.9
likelihood ratio (95% CI)	6.00 (2.60-13.83)	2.94 (1.46-5.93)
accuracy	80.4%	76.1%
PROMIS Self-Efficacy (SE)		
AUC (95% CI)	0.88 (0.80-0.95)**	0.69 (0.57-0.82)*
optimal T-score threshold	45.6	46.0
sensitivity	88.5	69.6
specificity	75.0	73.9
likelihood ratio (95% CI)	3.54 (2.05-6.11)	2.67 (1.32-5.40)
accuracy	82.6%	70.7%

*p<.01
**p<.001
PROMIS=Patient-Reported Outcomes Measurement Information System, PASS=patient acceptable symptom state, SOT=success of treatment, AUC=area under the curve, CI=confidence interval

4. Discussion

The findings of this analysis show that PASS and SOT are weakly associated in patients after care for non-surgical musculoskeletal conditions (hypotheses 1 and 2). PASS Yes and PASS No were significantly discriminated by all tested areas of patient current health status—PROMIS PF, PI, and SE (hypothesis 3). In contrast, SOT ImpPCoC and SOT Not Helped were significantly discriminated by PROMIS PI and SE only (hypothesis 4). Specific AUC, accuracy, and likelihood ratio values showed that PASS status was strongly (AUC ≥0.80) associated with PROMIS PI and SE, such that PI T-scores below 55 and SE T-scores above 45.6 strongly discriminated PASS Yes. In contrast, SOT was associated moderately (AUC 0.70-0.79) only with PROMIS PI T-scores, such that PI T-scores below 59.4 moderately discriminated SOT ImpPCoC. (Table 3) These combined findings suggest that outcomes focused on SOT may be more singularly associated with pain and less able to describe overall patient health status as compared to PASS.

The characteristics of the sample were typical primary care patients seeking care in a rural hospital-based clinic for musculoskeletal problems. Patients were 20-90 years of age with an average BMI of 31.0 (7.3) and with

44.5% of patients having a primary complaint associated with the spine (Table 1). These data are similar in age, gender and proportion of patients with a primary complaint associated with the spine to the larger sample this data was taken from (Kang et al., 2018). Although the treatment delivered was a single visit, the majority of patients reported some SOT (Improved (45.7%) or Partly Cured or Cured (29.3%)) and PASS Yes status (56.5%).

It is not surprising that many patients reported some level of success on the SOT question (i.e. ImpPCoC) without reaching PASS Yes status (Table 2). There was good agreement between Not Helped, Partly Cured, and Cured status and the corresponding PASS category (>70%). However, the Improved category showed little agreement, with this data roughly equivocal if patients also identified with a PASS Yes response. Overall the data could be interpreted as supporting PASS and SOT as being convergent. However, the association was relatively low (r=.39) suggesting the two single-item questions may be related but distinct assessments. The ROC analysis helped distinguish the association of SOT and PASS with health status. The ROC analysis and likelihood ratios suggest that patients reference both their level of functional limitation/perceived symptoms (i.e. PF/PI) and their confidence in managing symptoms (i.e. SE) differently when answering the PASS and SOT questions. Although this study did not assess change data, the lack of association between "improved" and PASS categories suggest that improvement was less important than reaching an acceptable level of symptoms as defined by PASS. This study supports that measures of current health status (PROMIS PF, PI, SE) discriminate PASS more accurately than they do SOT when post test status is considered. All three PROMIS measures had AUC values of greater than 0.70 for discriminating PASS, with SE and PI values approaching 0.9. Other studies have found similar or higher accuracy of disease-specific measures and of PROMIS measures to discriminate PASS.[17,36] The AUC, likelihood ratios, and accuracy values of health status (i.e. PROMIS measures) consistently discriminated PASS. However, for the same measures of health status only PI showed an AUC above 0.70 for SOT. Hence, PASS was supported as having a strong, meaningful association with patients' own perceived health status.

Differences in the AUC values, accuracies, and T-score thresholds give further insight into what patients emphasize when determining PASS and SOT based on post treatment data. For PASS, AUC and overall accuracy percentages were high for SE (0.88, 82.6%) and PI (0.87, 80.4%), compared to values for PF (0.73, 67.3%). This suggests that self-efficacy of managing symptoms in addition to pain interference and physical function influences whether patients decide whether or not their current status is acceptable. Additionally, thresholds on PROMIS measures indicating PASS were all roughly 5.0 T-score points from the US mean of 50, about ½ standard deviation worse. Meanwhile, SOT was only discriminated at an acceptable level by PI (AUC=0.77, overall accuracy 76.1%). The PI threshold for SOT was a T-score of 59.4, nearly one standard deviation worse than the US mean. Per general T-score interpretation guidelines from the HealthMeasures PROMIS website (http://www.healthmeasures.net/score-and-interpret/interpret-scores/promis/promis-score-cut-points), this PI threshold for SOT approaches a level at which clinicians could interpret this as "moderate" severity. In contrast, the PF, PI, and SE thresholds for PASS at approximately ½ standard deviation demarcate the point of "mild" severity. These findings suggest patients were willing to report at least some degree of success (i.e. SOT ImpPCoC) at a greater pain interference severity than patients would for reporting their status as acceptable (i.e. PASS Yes).

It appears that patients with non-surgical musculoskeletal conditions answer the PASS and SOT questions differently, reflecting a real discrimination on their part between their own current health status and success. Published variations of the SOT question have used 4-, 5-, or 7-point Likert rating scales, but all ask the patient to self-appraise treatment success following intervention.[9,13,18,19] Multiple studies involving a variety of patient populations (those with arthritis, chronic back pain, ankylosing spondylitis, femoral-acetabular impingement) have highlighted how patients' self-appraisal of a prior treatment does not equate to an acceptable level of function and symptoms (i.e. PASS).[9,35,52] This study adds to the discussion of how to establish health status by suggesting that clinicians who are interested in patients' health status would be better off focusing on acceptable/unacceptable symptoms. The PASS question at post assessment appears to reflect health status as defined by PROMIS measures. In comparison, the SOT question primarily aligns with the PROMIS pain interference measure, a limited assessment of health status. If more patients had experienced gains in health status then the SOT question may have more closely aligned with PASS. Therefore, this study's results best apply when patient outcomes are mixed.

Clinically this data supports the use of health status (as measured using PROMIS) to set goals and guide treatment decision making toward achieving PASS status. Achieving PASS benchmarks or better has been shown to be a more desirable outcome than judgements about success of treatment for patients with various conditions.[9,11,52] Past studies suggest that SOT might be used to identify whether a service or episode of care is meeting expectations,[12,19-21] and differences between perceptions associated with SOT and those related to health status are important to clinicians. In this study, only PI lowered to less than "moderate" severity (~1 standard deviation worse) was associated with SOT. When evaluating a given treatment's success, patients with musculoskeletal complaints may be more focused on perceptions of pain than other health status factors. It is apparent that both PASS and SOT have clinical value for understanding patient health outcomes. However, these two single-item PRO questions are not the same constructs related to patients achieving satisfactory health status. Clinicians may consider goal setting in specific domains of patient health status (i.e. PF, PI, and/or SE) based on the PROMIS T-score thresholds provided here and elsewhere,[17] allowing for more targeted treatment decision making for their patients with musculoskeletal conditions.

This study has several limitations. First, this study only sampled a portion of all the patients participating in this service and therefore should be considered a convenience sample which may not generalize to other samples and settings. Second, the study evaluated outcomes from a one-time collaborative physician/PT intervention for patients with primary musculoskeletal complaints, and therefore the degree of improvement was not evaluated as a factor influencing SOT and PASS scores. However, note that many patients in the present study might not have otherwise sought physical therapy following a physician-only visit.[38] Third, while the SOT question used in this study has a Not Helped response choice, no data around self-reported treatment failure was collected, as has been reported recently in sports medicine patients with knee injury.[10] Fourth, while patient-reported self-efficacy was measured, this study did not investigate the potential impact of patient beliefs and expectations as mediators of PASS and/or SOT responses. Indeed, while PASS associations here are strong, there are patients who report worse (i.e. below threshold) PROMIS T-scores and yet see themselves as PASS Yes. This has been demonstrated in various PRO measure

scores for patients with lower income status and/or diagnosed depression,[36] as well as for patients with rheumatoid arthritis of longer duration and/or with lower baseline scores.[34,35] Hence, optimal PASS Yes T-score thresholds reported here should be applied alongside other clinical findings. Finally, there was no control for patients who sought out additional treatment after the one-time physician/PT intervention but before the 45-60 day follow-up, nor was there consideration of differences in outcomes reporting amongst various patient characteristic groupings. Diagnosis-specific patient preferences may occur that are not distinguished here.

5. Conclusion

This data advances understanding of the association of PASS and SOT, encouraging clinicians to understand the distinction between these two constructs of outcome with their patients. The PASS question was more closely linked to patient-reported physical function, pain interference, and self-efficacy than SOT. A PROMIS PF, PI, or SE T-score around ½ standard deviation worse than the US mean best discriminates PASS status. A PROMIS PI T-score approximately one standard deviation worse reflects "moderate" severity and discriminates SOT response. Because of the strong ability in particular of PROMIS PI and SE measures to discriminate PASS, it may be a more meaningful patient current health status outcome than SOT. Clinicians seeking to impact health status—independent of whether or not treatment is judged as "successful"—might base goal setting and treatment decision making on their patients' PASS status and those PROMIS T-scores that are worse than ½ standard deviation.

6. Disclosures

The authors report no conflicts of interest. There was no funding for this study.

7. Acknowledgements

We would like to acknowledge the contributions of Cecelia Gutierrez, Andrea Kuebbing, Amador Marcano, Kevin Morikawa, Catherine Vandehaar for their work in recruitment and data collection as part of their doctoral studies. We would also like to thank the participants who graciously completed the outcome measures reported here.

References

1. Papuga MO, Dasilva C, McIntyre A, Mitten D, Kates S, Baumhauer JF. Large-scale clinical implementation of PROMIS computer adaptive testing with direct incorporation into the electronic medical record. *Health Syst.* 2018;7(1):1-12.
2. Baumhauer JF. Patient-reported outcomes—Are they living up to their potential? *N Engl J Med.* 2017;377(1):6-9.
3. Porter ME. What is value in health care? *N Engl J Med.* 2010;363(26):2477-2481.
4. Field J, Holmes MM, Newell D. PROMs data: can it be used to make decisions for individual patients? A narrative review. *Patient Relat Outcome Meas.* Jul 2019;10:233-241.
5. Karimi M, Brazier J. Health, health-related quality of life, and quality of life: What is the difference? *Pharmacoecon.* 2016;34(7):645-649.
6. Ngamaba KH, Panagioti M, Armitage CJ. How strongly related are health status and subjective well-being? Systematic review and meta-analysis. *Eur J Public Health.* 2017;27(5):879-885.
7. Rumsfeld JS, Alexander KP, Goff DC, Jr., et al. Cardiovascular health: The importance of measuring patient-reported health status: A scientific statement from the American Heart Association. *Circulation.* 2013;127(22):2233-2249.
8. Batbaatar E, Dorjdagva J, Luvsannyam A, Savino MM, Amenta P. Determinants of patient satisfaction: A systematic review. *Perspect Public Health.* 2017;137(2):89-101.
9. Impellizzeri FM, Mannion AF, Naal FD, Hersche O, Leunig M. The early outcome of surgical treatment for femoroacetabular impingement: Success depends on how you measure it. *Osteoarthr Cartil.* 2012;20(7):638-645.
10. Roos EM, Boyle E, Frobell RB, Lohmander LS, Ingelsrud LH. It is good to feel better, but better to feel good: Whether a patient finds treatment 'successful' or not depends on the questions researchers ask. *Br J Sports Med.* 2019;53(23):1474-1478.
11. Strand V, Boers M, Idzerda L, et al. It's good to feel better but it's better to feel good and even better to feel good as soon as possible for as long as possible: Response criteria and the importance of change at OMERACT 10. *J Rheumatol.* 2011;38(8):1720-1727.
12. 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.
13. Hawthorne G, Sansoni J, Hayes L, Marosszeky N, Sansoni E. Measuring patient satisfaction with health care treatment using the Short Assessment of Patient Satisfaction measure delivered superior and robust satisfaction estimates. *J Clin Epidemiol.* 2014;67(5):527-537.
14. Kamper SJ, Maher CG, Mackay G. Global rating of change scales: a review of strengths and weaknesses and considerations for design. *J Man Manip Ther.* 2009;17(3):163-170.

15. 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):40-41.
16. 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.
17. Houck J, Kang D, Cuddeford T, Rahkola S. 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. *Arch Phys Med Rehabil.* 2019;100(1):60-66.
18. Onida S, Shalhoub J, Moore HM, Head KS, Lane TRA, Davies AH. Factors impacting on patient perception of procedural success and satisfaction following treatment for varicose veins. *Br J Surg.* 2016;103(4):382-390.
19. Hossack T, Woo H. Validation of a patient reported outcome questionnaire for assessing success of endoscopic prostatectomy. *Prostate Int.* 2014;2(4):182-187.
20. Bernhardsson S, Larsson M, Johansson K, Oberg B. "In the physio we trust": A qualitative study on patients' preferences for physiotherapy. *Physiother Theory Pract.* 2017;33(7):535-549.
21. Steffens NM, Tucholka JL, Nabozny MJ, Schmick AE, Brasel KJ, Schwarze ML. Engaging patients, health care professionals, and community members to improve preoperative decision making for older adults facing high-risk surgery. *JAMA Surgery.* 2016;151(10):938-945.
22. Cody EA, Mancuso CA, Burket JC, Marinescu A, MacMahon A, Ellis SJ. Patient factors associated with higher expectations from foot and ankle surgery. *Foot Ankle Int.* 2017;38(5):472-478.
23. Brodke DS, Goz V, Voss MW, Lawrence BD, Spiker WR, Hung M. PROMIS PF CAT outperforms the ODI and SF-36 physical function domain in spine patients. *Spine.* 2017;42(12):921-929.
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):1-7.
25. 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.
26. Papuga MO, Mesfin A, Molinari R, Rubery PT. Correlation of PROMIS physical function and pain CAT instruments with Oswestry Disability Index and Neck Disability Index in spine patients. *Spine.* 2016;41(14):1153-1159.
27. Amtmann D, Cook KF, Johnson KL, Cella D. The PROMIS initiative: Involvement of rehabilitation stakeholders in development and examples of applications in rehabilitation research. *Arch Phys Med Rehabil.* 2011;92(10 Suppl):S12-S19.
28. Jacobson R, Philbrook L, Kang D, Cuddeford T, Houck J. Does multidimensional health assessment using PROMIS scales enhance clinical

decision-making for patients with orthopedic problems? A case series. *Orthop Phys Ther Pract.* 2018;30(4):528-536.
29. Costa Lda C, Maher CG, McAuley JH, Hancock MJ, Smeets RJ. Self-efficacy is more important than fear of movement in mediating the relationship between pain and disability in chronic low back pain. *Eur J Pain.* 2011;15(2):213-219.
30. Twomey R, Martin T, Temesi J, Culos-Reed SN, Millet GY. Tailored exercise interventions to reduce fatigue in cancer survivors: study protocol of a randomized controlled trial. *BMC Cancer.* 2018;18(1):757.
31. Yeh GY, Mu L, Davis RB, Wayne PM. Correlates of exercise self-efficacy in a randomized trial of mind-body exercise in patients with chronic heart failure. *J Cardiopulm Rehabil Prev.* 2016;36(3):186-194.
32. George SZ, Valencia C, Beneciuk JM. A psychometric investigation of fear-avoidance model measures in patients with chronic low back pain. *J Orthop Sports Phys Ther.* 2010;40(4):197-205.
33. Bandura A, O'Leary A, Taylor CB, Gauthier J, Gossard D. Perceived self-efficacy and pain control: opioid and nonopioid mechanisms. *J Pers Soc Psychol.* 1987;53(3):563-571.
34. Heiberg TT, Kvien TK, Mowinckel P, Aletaha D, Smolen JS, Hagen KB. Identification of disease activity and health status cut-off points for the symptom state acceptable to patients with rheumatoid arthritis. *Ann Rheum Dis.* 2008;67(7):967-971.
35. Tubach F, Ravaud P, Martin-Mola E, et al. Minimum clinically important improvement and patient acceptable symptom state in pain and function in rheumatoid arthritis, ankylosing spondylitis, chronic back pain, hand osteoarthritis, and hip and knee osteoarthritis: Results from a prospective multinational study. *Arthritis Care Res.* 2012;64(11):1699-1707.
36. 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(6):814-819.
37. Kang D, Rahkola S, Vandehaar C, et al. A study of outcomes following collaborative medical doctor/physical therapist primary care service for musculoskeletal problems. *Orthop Phys Ther Pract.* 2018;30(4):510-517.
38. Frogner BK, Harwood K, Andrilla CHA, Schwartz M, Pines JM. Physical therapy as the first point of care to treat low back pain: An instrumental variables approach to estimate impact on opioid prescription, health care utilization, and costs. *Health Serv Res.* 2018;53(6):4629-4646.
39. da C Menezes Costa L, Maher CG, Hancock MJ, McAuley JH, Herbert RD, Costa LO. The prognosis of acute and persistent low-back pain: a meta-analysis. *CMAJ.* 2012;184(11):E613-624.
40. Thompson JY, Byrne C, Williams MA, Keene DJ, Schlussel MM, Lamb SE. Prognostic factors for recovery following acute lateral ankle ligament sprain: a systematic review. *BMC Musculoskel Dis.* 2017;18(1):421.
41. Ader DN. Developing the Patient-Reported Outcomes Measurement Information System (PROMIS). *Med Care.* 2007;45(5):S1-2.
42. Cella D, Yount S, Rothrock N, et al. The Patient-Reported Outcomes Measurement Information System

(PROMIS): Progress of an NIH roadmap cooperative group during its first two years. *Med Care.* 2007;45(5):S3-S11.

43. 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 Clinical Epidemiol.* 2014;67(5):516-526.
44. Amtmann D, Cook KF, Jensen MP, et al. Development of a PROMIS item bank to measure pain interference. *Pain.* 2010;150(1):173-182.
45. 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(7):1915-1924.
46. 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.
47. Bernstein DN, Kelly M, Houck JR, et al. PROMIS Pain Interference is superior vs Numeric Pain Rating Scale for pain assessment in foot and ankle patients. *Foot Ankle Int.* 2019;40(2):139-144.
48. Haynes W. Bonferroni Correction. In: Dubitzky W, Wolkenhauer O, Cho K-H, Yokota H, eds. *Encyclopedia of Systems Biology*. New York, NY: Springer New York; 2013:154.
49. Mandrekar JN. Receiver operating characteristic curve in diagnostic test assessment. *J Thorac Oncol.* 2010;5(9):1315-1316.
50. Perkins NJ, Schisterman EF. The inconsistency of "optimal" cutpoints obtained using two criteria based on the receiver operating characteristic curve. *Am J Epidemiol.* 2006;163(7):670-675.
51. 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.
52. Tubach F, Dougados M, Falissard B, Baron G, Logeart I, Ravaud P. Feeling good rather than feeling better matters more to patients. *Arthritis Rheum.* 2006;55(4):526-530.

