Reliability and Validity of Patient-reported Outcomes Measurement Information System Short Forms in Ankylosing Spondylitis

Mark C. Hwang, Alexis Ogdie, Abin Puravath, and John D. Reveille

ABSTRACT. ***Objective.*** To assess the reliability and validity in ankylosing spondylitis (AS) of selected Patient Reported Outcomes Measurement Information System (PROMIS) Short Forms (SF) developed by the US National Institutes of Health. The analysis was done across core sets and patient-identified domains of the Assessment of Spondyloarthritis international Society.

Methods. Participants in the Prospective Study of Outcomes in Ankylosing Spondylitis (PSOAS), an ongoing, prospective longitudinal observational study, completed 6 PROMIS SF assessing global health, depression, fatigue, physical function, pain intensity, and pain interference during their PSOAS visits from September 2017 to January 2019. Test-retest reliability and internal consistency were assessed using intraclass correlation coefficients and Cronbach's alpha coefficient, respectively. PROMIS SF were compared to legacy measures collected. Construct validity was evaluated through examination of score distributions and floor effects, and through examination of the Spearman correlation coefficients between PROMIS measures and existing legacy AS measures. Discriminant validity was tested across Ankylosing Spondylitis Disease Activity Score (ASDAS) groups.

Results. Participants (n = 119) were mostly male (69%), white (81%), and with a mean (SD) age of 51 (± 15) years. Legacy measures demonstrated floor effects that were not present in PROMIS SF. Good test-retest reliability ($r > 0.8$) and excellent internal consistency ($\alpha > 0.9$) was noted in the PROMIS SF. The 6 PROMIS SF correlated moderately to strongly [ρ 0.68 (Depression) to –0.87 (Physical Function)] with appropriate legacy measures. PROMIS scores measures worsened significantly ($p < 0.05$) with higher ASDAS groups.

Conclusion. This study supports the reliability and construct validity of PROMIS SF to assess AS symptoms from a single-center sample of patients with AS. Further research is needed to test responsiveness, feasibility/resource burden, and different cultural/societal contexts for patients with AS. (J Rheumatol First Release June 15 2020; doi:10.3899/jrheum.190201)

Key Index Terms:
ANKYLOSING SPONDYLITIS ARTHRITIS IMPACT PATIENT REPORTED OUTCOMES MEASUREMENT INSTRUMENT PSYCHOMETRICS SURVEYS

Patient-reported outcomes (PRO) are an important component of rheumatologic care and research. They have increased patient participation and yielded valuable information on treatment efficacy and quality of life that is pertinent to the management of patients with these complex, chronic diseases[1,2]. Subsets of PRO have also been established as core outcome domains for many rheumatic diseases to evaluate therapeutic efficacy[3].

Ankylosing spondylitis (AS) is a disease characterized by inflammatory back pain and radiographic disease of the axial spine, with an estimated prevalence of 0.2% to 0.5% in the US population[4]. Clinicians have widely adopted the use of PRO as important tools in AS management. In fact, PRO comprise the largest share of the primary outcomes in randomized controlled trials in AS[5]. The Assessment of Spondyloarthritis international Society/Outcome Measures

From the Division of Rheumatology and Clinical Immunogenetics, Department of Internal Medicine, John P. and Kathrine G. McGovern Medical School, University of Texas Health Science Center at Houston (UTHealth), Houston, Texas; the Division of Rheumatology, Department of Medicine, Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania, USA.

This project was supported by the US National Institutes of Health (NIH) PO1-AR052915. Dr. Ogdie is supported by NIH K23 AR063764, R01 AR072363, and the Rheumatology Research Foundation. Dr. Reveille is supported by the Spondyloarthritis Association of America.

M.C. Hwang, MD, Division of Rheumatology and Clinical Immunogenetics, Department of Internal Medicine, John P. and Kathrine G. McGovern Medical School, UTHealth; A. Ogdie, MD, MSCE, Division of Rheumatology, Department of Medicine, Perelman School of Medicine, University of Pennsylvania; A. Puravath, MD, Division of Rheumatology and Clinical Immunogenetics, Department of Internal Medicine, John P. and Kathrine G. McGovern Medical School, UTHealth; J.D. Reveille, MD, Division of Rheumatology and Clinical Immunogenetics, Department of Internal Medicine, John P. and Kathrine G. McGovern Medical School, UTHealth.

Address correspondence to Dr. M.C. Hwang, John P. and Kathrine G. McGovern Medical School, UTHealth, 6431 Fannin, MSB.5276, Houston, Texas 77030, USA. E-mail: Mark.C.Hwang@uth.tmc.edu

Accepted for publication July 19, 2019.

Personal non-commercial use only. The Journal of Rheumatology Copyright © 2020. All rights reserved.

Downloaded on April 23, 2024 from www.jrheum.org

in Rheumatology (ASAS/OMERACT) international groups have established 3 independent core sets of domains used to measure outcomes: (1) disease-controlling antirheumatic treatment, (2) symptom-modifying antirheumatic drugs and physical therapy, and (3) clinical record-keeping[6]. All 3 core sets include the domains of fatigue, function, pain, patient global assessment, and stiffness. These domains are important both from a research and clinical care level in AS.

Universal (or "generic") PRO measures represent an opportunity to compare disease burden and treatment effect across different chronic conditions using a common metric[7]. The US National Institutes of Health (NIH)–funded Patient Reported Outcomes Measurement Information System (PROMIS) incorporates both adult and pediatric PRO in physical, mental, and social health domains across a wide variety of chronic diseases and conditions as well as general population controls, potentially allowing for this type of comparison[8]. The physical health domains captured by PROMIS are particularly relevant in rheumatology[9]. The use of item response testing (IRT) methodology yielded computer adaptive tests (CAT), static profiles, and short forms (SF) PROMIS instruments that are publicly available for use. Investigators continue to explore how PROMIS measures can be incorporated into different aspects of medical research and care[10].

While the ASAS/OMERACT PRO measures are vital in the assessment of AS, the IRT methodology used in PROMIS potentially reduces redundancy, increases sensitivity by avoiding floor/ceiling effects, and decreases survey burden with its adaptive design. Additionally, the publicly available online data collection system (Assessment Center, www.assessmentcenter.net) may lower barriers to clinical research in AS through its accessibility and ease of use. The purpose of our study is to examine the reliability and validity of PROMIS SF in patients with AS.

MATERIALS AND METHODS

Patients. Subjects were recruited from a single center [University of Texas Health Science Center at Houston (UTHealth)] among patients currently enrolled in the Prospective Study of Ankylosing Spondylitis (PSOAS) observational cohort. All clinic patients at the UTHealth study site who met modified New York Classification Criteria for AS, were at least 18 years of age and fluent in English were eligible for participation. PSOAS is a multicenter observational study initiated in 2003 with continued enrollment encompassing 5 sites: Cedars-Sinai Medical Center, Los Angeles, California; the NIH Clinical Center, Bethesda, Maryland; the McGovern Medical School at UTHealth; the University of California, San Francisco; and the Princess Alexandra Hospital in Brisbane, Australia. The research followed the Declaration of Helsinki and was approved by the University of Texas Institutional Review Board (HSC-MS-07-0022). Each participating subject reviewed and signed an informed consent form.

Procedures. Patients were contacted by telephone, at their clinic visit, or at their study visit about participation and details regarding this ancillary study. After patients provided written informed consent, coordinators provided paper questionnaire packets in person and/or by e-mail for printing. A subset of patients was consecutively approached and asked to complete the same PROMIS SF > 48 h later to assess test-retest reliability from May to November 2018. We used data from a single patient visit per patient.

PRO assessments. Focus groups of patients with AS (3 groups of 5 patients) were asked whether the domains listed in the ASAS/OMERACT were important for their disease and if there were any additional domains they felt needed to be measured. Five academic rheumatologists at UTHealth involved in AS research and patient care were individually asked the same questions. After soliciting these opinions, we found that the core domains were well accepted among patients and rheumatologists. Mental health, specifically depression, was an important domain that was the most noted domain not covered in the ASAS/OMERACT core set.

PRO that are routinely collected in the PSOAS cohort include Patient Global assessment [0–100 numerical rating scale (NRS)], Patient Pain assessment (0–100 NRS), Bath Ankylosing Spondylitis Disease Activity Index (BASDAI), Bath Ankylosing Spondylitis Functional Index (BASFI), Ankylosing Spondylitis Disease Activity Score–C-reactive protein (ASDAS–CRP), and the Center for Epidemiologic Studies Depression Scale (CES-D)[11,12,13,14]. The BASDAI consists of 6 questions measured on a 0–10 scale covering 5 major symptoms of AS: fatigue, spinal pain, arthralgias/arthritis, enthesitis, and morning stiffness[11]. Additionally, Calin, *et al* developed the BASFI, a 10-question index measured on the mean of 0–10 scales focused on functional AS anatomical limitations[12]. The ASDAS is a newer disease activity measure designed specifically for AS and demonstrating high discriminant capacity; it incorporates acute-phase reactants [e.g., erythrocyte sedimentation rate (ESR) or CRP][14]. The CES-D is a 20-item questionnaire measuring depressive symptom severity[13]. These measures were termed "legacy" measures and served as comparators for the PROMIS measures addressing similar constructs.

As part of the NIH Roadmap initiative for the 21st century for medical research, the multicenter cooperative group referred to as PROMIS was formed. This group used modern advances in computer technology and item-response theory to create free-to-use measures for physical, mental, and social health domains[15]. Among the ways to administer PROMIS measures (on paper, by computer, or mobile application), we chose SF distributed in paper packets for ease of use in a clinical setting. Scoring manuals for PROMIS measures (www.assessmentcenter.net/Manuals.aspx) outline the SF development, report psychometric properties for each instrument in their study population, and describe how to identify PROMIS T scores based on short form raw summed item scores. We reported PROMIS T scores for all SF. PROMIS SF Versions 1.0/1.1 (assessmentcenter.net) were administered for Emotional Distress (ED)–Depression, Fatigue, Global Health, Pain Intensity, Pain Interference, and Physical Function and ranged from 3 to 12 questions per form (Supplementary Figure 1, available with the online version of this article). For the PROMIS Global v1.1 SF we reported the physical summary score. These domains were selected based on patient input, expert rheumatologist input, and the published ASAS/OMERACT core set for clinical record keeping[6]. Higher PROMIS scores represent more of the measured trait, so interpretation of directionality varied if the domain was a positive trait (higher scores better) versus symptom (higher scores indicate more severe symptoms). Time to complete was self-reported by patients upon completion of the PRO packet. Through open-ended critiques, we additionally solicited patient feedback on the PROMIS questionnaires regarding how well they addressed the measured domains and whether any important aspects of their disease were not being addressed.

Covariates. Sociodemographic information was drawn from the patients' data extracted from the PSOAS cohort and included age, sex, race/ethnicity, education, smoking habits, comorbidities, work status, and AS duration. Medication use, comorbidities, and serum inflammatory markers (e.g., CRP or ESR) were also recorded at each visit in addition to radiographs of the hips and the cervical and lumbar spine (the latter measured by the modified Stoke Ankylosing Spondylitis Scoring Scale), obtained at 2-year intervals over the course of followup.

Statistical analysis. Central tendency and distribution were calculated by mean (SD) or median [interquartile range (IQR)] for continuous normal versus non-normal data, respectively. Frequencies and percentages were

Personal non-commercial use only. The Journal of Rheumatology Copyright © 2020. All rights reserved.

Downloaded on April 23, 2024 from www.jrheum.org

descriptively reported for categorical variables. We examined histograms, and skewness and kurtosis statistics to assess normality[16]. For skewness and kurtosis, we looked at their Z scores by dividing their values over their standard error, with values > |1.96| considered significant. Spearman's correlation was used to examine PROMIS scores against legacy PRO for similar domains. The Kruskal-Wallis H test with Bonferroni correction for pairwise comparisons was used to compare PROMIS and legacy PRO domains stratified by ASDAS group levels. Intraclass correlation coefficient and Cronbach's alpha coefficient were used to assess test-retest reliability and internal consistency, addressing reliability; a correlation coefficient or alpha coefficient > |0.7| was considered acceptable. We hypothesized *a priori* that there would be moderate to strong correlation (ρ >|0.6|) between the PROMIS measures and the target legacy measures. All analyses were done with IBM SPSS version 24.

RESULTS

Patient characteristics. A total of 119 patients were enrolled and completed the surveys between September 2017 and January 2019. Twenty-four of the 88 patients (27.3%) from May 2018 through November 2018 completed the retest packet in addition. This sample included a diverse spectrum of AS characteristics (Table 1). Patients were mostly male (69%) and white (81%), with a mean (SD) age of 51 years (15). All patients met the modified New York Criteria for Ankylosing Spondylitis with a mean symptom duration of 25 years (± 13). In those who had available CRP laboratory values (90/119, 76%), over half (64%) had inactive or moderate disease by ASDAS.

Distributions of PRO. PROMIS and legacy scores are shown in Table 2. No significant kurtosis was noted in the legacy or PROMIS questionnaires. All legacy PRO were moderately to highly positively skewed (0.54 to 1.00), with the Patient Global, Patient Pain, BASFI, and CES-D significantly skewed ($p < 0.05$). Many of the instruments demonstrated floor effects with the proportion of patients at the lowest potential score in each of the legacy measures ranging from 5% to 17%: Global Health NRS (17%), CES-D (9%), BASDAI-Fatigue (5%), Pain (14%), and BASFI (11%). The PROMIS instruments showed a more normal distribution compared to legacy measures in physical domains (Figure 1, and Supplementary Figure 2, available with the online version of this article) with PROMIS Global, Fatigue, Pain Interference, Pain Intensity, Physical Function about symmetrical (–0.1 to 0.39 skew, $p > 0.05$). However, PROMIS ED-Depression had positive skew (1.01, $p < 0.05$) with significant floor effect (e.g., 54% with lowest possible value PROMIS-ED; Supplementary Figure 2). Floor effects for the rest of the PROMIS measures otherwise ranged from 1% to 31%: Global (1%), Fatigue (9%), Pain Interference (31%), Pain Intensity (11%), and Physical Function (1%). Among a subset that were sampled for time of completion, 35/41 (85%) stated that overall it was < 15 min to complete their PROMIS SF packets. Fourteen of the 119 patients (12%) raised concerns regarding the PROMIS SF addressing their disease.

Reliability. Test-retest reliability was tested among the 24 participants. The median (IQR) time between the 2 measures was 1 day (IQR 1–2). Correlations between the individual tests' 2 scores ranged from 0.80 (Pain Interference) to 0.98 (Physical Function). We also examined internal consistency using Cronbach's alpha. We found excellent consistency within the individual scales ranging from 0.91 (Global) to 0.98 (Pain Interference).

Construct validity: convergent validity and known groups validity with legacy measures. PROMIS Global Health, Physical Function, and Pain Intensity had very strong correlation (ρ value > 0.84, $p < 0.01$) with corresponding legacy measures (Global NRS, BASFI, and Pain NRS, respectively; Table 3). PROMIS Pain Interference and Fatigue showed strong correlation (ρ > 0.7, $p < 0.01$) with corresponding legacy measures (Pain NRS and BASDAI-Fatigue, respectively). The weakest correlation was seen with PROMIS Emotional Distress-Depression, which had moderate correlation (ρ = 0.68, $p < 0.01$) with CES-D.

In general, PROMIS scores measures worsened significantly ($p < 0.05$) with increased disease activity as defined by ASDAS categories ranging from inactive disease (ASDAS < 1.3) to high–very high disease activity (ASDAS ≥ 2.1) in the domains of Global, Fatigue, Pain Intensity, Pain Interference, Depression, and Physical Function (Table 4). In pairwise comparisons, PROMIS Global and Physical Function distinguished inactive, moderate, and

Table 1. Participant characteristics (n = 119).

Characteristic	N	Values
Age, yrs, mean ± SD	119	50.85 ± 14.77
Male	119	82 (69)
Race	119	
White		96 (81)
Other		23 (19)
Education	119	
High school		16 (13)
College		103 (87)
Employment status	119	
Full-time		76 (64)
Not working		31 (27)
Disabled		11 (9)
Self-reported depression	119	16 (13)
AS symptom duration, yrs, mean ± SD	119	25.47 ± 13.32
ASDAS*	90	
Inactive		26 (31)
Moderate		28 (33)
High–very high		30 (36)
Biologic DMARD usage†	110	66 (56)
Last mSASSS, median (IQR)‡	85	4 (0–34)

Values are n (%) unless otherwise specified. * Missing = 29 owing to lack of CRP laboratory data. † Missing = 9 because of incomplete medication list. ‡ Missing = 34 owing to incomplete radiographs. AS: ankylosing spondylitis; ASDAS: AS Disease Activity Score; DMARD: disease-modifying antirheumatic drug; mSASSS: modified Stoke AS Spine Score; IQR: interquartile range.

Personal non-commercial use only. The Journal of Rheumatology Copyright © 2020. All rights reserved.

Downloaded on April 23, 2024 from www.jrheum.org

Table 2. PROMIS and legacy measure scores in patients with AS.

	N	Mean	Median	SD	Range	Min	Max	Test-retest ICC (95% CI)*	Cronbach's Alpha (95% CI)
Patient Global									
PROMIS Global	119	45.62	44.90	8.90	44.20	23.50	67.70	0.898 (0.763–0.958)	0.910 (0.884–0.932)
Global VAS (NRS)	115	29.13	20.00	25.15	100.00	0.00	100.00		
Depression									
PROMIS Emotional Distress–Depression	118	45.28	38.20	8.48	31.50	38.20	69.70	0.859 (0.699–0.938)	0.936 (0.917–0.952)
CES-D	109	11.00	9.00	8.67	39.00	0.00	39.00		
Fatigue									
PROMIS Fatigue	118	51.07	49.80	10.46	44.70	33.10	77.80	0.901 (0.788–0.955)	0.972 (0.964–0.979)
BASDAI–Fatigue	115	4.12	3.00	2.78	10.00	0.00	10.00		
Pain									
PROMIS Pain Intensity	119	45.74	46.30	8.89	36.70	30.70	67.40	0.871 (0.729–0.941)	0.912 (0.881–0.937)
PROMIS Pain Interference	119	52.23	53.20	9.91	36.30	40.70	77.00	0.814 (0.616–0.915)	0.978 (0.972–0.984)
Pain VAS (NRS)	115	32.17	20.00	28.18	100.00	0.00	100.00		
Physical Function									
PROMIS Physical Function	112	46.56	45.20	9.80	41.00	25.10	66.10	0.957 (0.905–0.981)	0.922 (0.899–0.942)
BASFI	113	3.06	2.40	2.63	9.40	0.00	9.40		

* N = 25. AS: ankylosing spondylitis; VAS: visual analog scale; NRS: numerical rating scale; CES-D: Center for Epidemiologic Studies-Depression; PROMIS: Patient Reported Outcomes Measurement Information System; BASDAI: Bath AS Disease Activity Index; BASFI: Bath AS Functional Index.

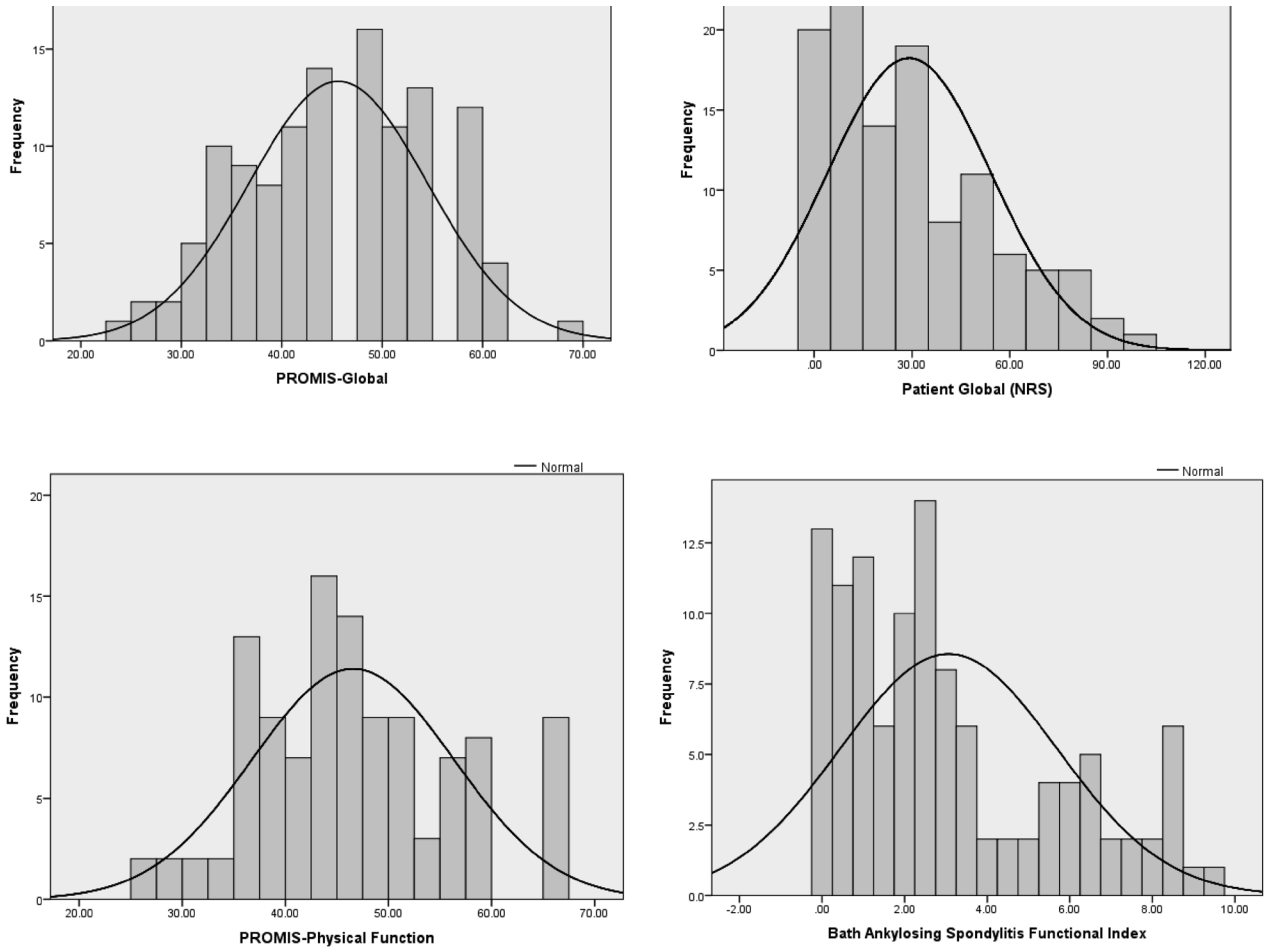

Figure 1. Distribution of select PROMIS (left) and legacy scores (right) in patients with ankylosing spondylitis (n = 119). PROMIS: Patient Reported Outcomes Measurement Information System; NRS: numerical rating scale.

Personal non-commercial use only. The Journal of Rheumatology Copyright © 2020. All rights reserved.

Downloaded on April 23, 2024 from www.jrheum.org

Table 3. Correlations between PROMIS and legacy measures in patients with AS.

	Global VAS (NRS)	CES-D	BASDAI: Fatigue	Pain VAS (NRS)	BASFI
PROMIS Global	–0.864*	—	—	—	—
PROMIS Emotional Distress–Depression	—	0.68*	—	—	—
PROMIS Fatigue	—	—	0.708*	—	—
PROMIS Pain Intensity	—	—	—	0.856*	—
PROMIS Pain Interference	—	—	—	0.746*	—
PROMIS Physical Function	—	—	—	—	–0.872*

* $p < 0.01$. PROMIS: Patient Reported Outcomes Measurement Information System; AS: ankylosing spondylitis; VAS: visual analog scale; CES-D: Center for Epidemiologic Studies Depression Scale; BASDAI: Bath AS Disease Activity Index; NRS: numerical rating scale; BASFI: Bath AS Functional Index.

Table 4. PROMIS and legacy measures by ASDAS disease activity levels.

	Inactive Disease (n = 25)		Moderate Disease Activity (n = 30)		High Disease Activity (n = 35)	
	Mean	SD	Mean	SD	Mean	SD
Patient Global						
PROMIS Global	53.46_a	5.73	47.31_b	6.19	38.39_c	6.17
Global VAS (NRS)	6.80_a	8.02	22.33_b	14.07	49.14_c	23.98
Depression						
PROMIS Emotional Distress–Depression*	41.71_a	6.26	46.02_a	9.72	46.95_a	8.17
CES-D	6.63_a	5.28	9.11_a	8.49	13.94_b	8.27
Fatigue						
PROMIS Fatigue	44.44_a	8.09	47.89_a	9.15	57.43_b	10.75
BASDAI-Fatigue	1.60_a	1.38	3.30_a	2.04	6.63_b	2.26
Pain						
PROMIS Pain Intensity	38.56_a	5.95	42.31_a	6.17	52.93_b	6.54
PROMIS Pain Interference	44.77_a	6.50	47.88_a	7.01	60.53_b	6.77
Pain VAS (NRS)	8.40_a	8.00	18.67_a	15.92	58.86_b	21.93
Physical Function						
PROMIS Physical Function	54.67_a	8.60	46.31_b	7.34	40.68_c	6.90
BASFI	0.97_a	1.02	2.55_b	1.73	4.99_c	2.65

Values in the same row not sharing the same subscript are significantly different at adjusted significance of $p < 0.05$. * $p > 0.05$ in PROMIS-Depression. PROMIS: Patient Reported Outcomes Measurement Information System; ASDAS: Ankylosing Spondylitis Disease Activity Score; VAS: visual analog scale; NRS: numerical rating scale; CES-D: Center for Epidemiologic Studies-Depression; BASDAI: Bath Ankylosing Spondylitis Disease Activity Index; BASFI: Bath Ankylosing Spondylitis Functional Index.

high–very high ASDAS-defined disease activity. PROMIS Fatigue, Pain Intensity, and Pain Interference were able to distinguish ASDAS inactive and moderate from high–very high disease activity. We observed these same patterns among the legacy measures (Patient Global, Pain, BASFI, BASDAI-Fatigue) that addressed physical domains. PROMIS ED-Depression measure was unable to distinguish across ASDAS-defined disease activity, unlike CES-D, which could distinguish high–very high disease compared to inactive disease.

DISCUSSION

To the best of our knowledge, this study is the first to examine the reliability and construct validity of PROMIS instruments in patients with AS in the context of ongoing clinical care. We selected PROMIS SF from a patient and clinician perspective based on patient and physician input as well as review of the ASAS/OMERACT clinical record-keeping domains. Among the 6 domains we studied (Depression, Fatigue, Global Health, Physical Function, Pain Intensity, and Pain Interference), 5 showed at least strong correlation ($\rho > |0.7|$) with the appropriate legacy AS measure. Additionally, in the physical domains of Global Health and Physical Function, PROMIS measures were able to discriminate inactive, moderate, and high–very high ASDAS activity groups. Similarly, in the other physical domains (i.e., Pain Intensity, Pain Interference, and Fatigue), the PROMIS measures could discriminate high–very high disease versus low-moderate disease activity groups. In depression, the only mental health domain, PROMIS Depression could not distinguish across disease activity levels, suggesting that depressive symptoms as defined in this SF may not be disease-related. A majority of patients also found these forms to take < 15 min to complete. These findings support the feasibility, reliability, and construct validity of PROMIS SF when assessing physical domains in AS outcomes.

PROMIS instruments have been evaluated across

Personal non-commercial use only. The Journal of Rheumatology Copyright © 2020. All rights reserved.

Downloaded on April 23, 2024 from www.jrheum.org

multiple physical, mental, and social health domains in other rheumatic diseases including juvenile idiopathic arthritis, osteoarthritis (OA), psoriatic arthritis, rheumatoid arthritis (RA), systemic lupus erythematosus (SLE), systemic sclerosis, and vasculitis[17-23]. While a majority of these studies have focused on PROMIS CAT, to date PROMIS SF have been studied in OA, RA, and SLE in similar fashion[24,25,26,27]. PROMIS instruments have also been studied in treatment response[28]. In addition, PROMIS measures have also correlated with objective measures. For example, Mahieu, *et al* demonstrated reliability and construct validity in PROMIS Fatigue with accelerometer-based measures of physical activity in patients with SLE[29]. This may suggest the potential cross-disease nature of these universal PRO.

Strengths of our study included use of a well-characterized cohort reflective of patients with AS in the United States, with AS legacy measures collected routinely at each study visit. All patients met modified New York criteria for ankylosing spondylitis, creating a homogeneous patient sample from a radiographic perspective. We evaluated the performance of PROMIS measures within the context of usual care.

Our study had limitations. The highly educated, largely white demographics of our patient sample may affect generalizability. Floor effects noted for the PROMIS Depression SF and CES-D may have been due to the low depression rate in our sample. Further, by including only patients who met modified New York criteria for AS, we excluded patients on the disease spectrum with nonradiographic axial spondyloarthritis (nr-axSpA). Thus, our study may not be generalizable to nr-axSpA patient populations. Because of potential rapid changes in underlying disease activity with medications, the 48-h type window was chosen to assess test-retest, a practice that may also artificially elevate the correlation observed. For both test-retest and time of completion, we acknowledge potential participation bias of those who volunteered this information. We also did not examine responsiveness of the PROMIS instruments in this study, and limited our study to English-speaking patients. Further, while the use of SF is feasible in all potential settings given its paper format, we did not study the CAT or profile PROMIS instruments.

Our study offers preliminary data in the study of PROMIS instruments in AS. Further study is required to see, separately, whether the PROMIS CAT can reduce redundancy, increase sensitivity by avoiding floor/ceiling effects, or decrease survey burden with its adaptive design in patients with AS. Future PROMIS SF studies in AS could include translations of PROMIS instruments, given the dynamic demographics seen in the US population. Further, longitudinal studies are required to study the responsiveness of PROMIS SF in patients with AS.

Our study offers evidence supporting the feasibility as well as the reliability and construct validity of 6 PROMIS instruments in AS clinical care. Additionally, our study demonstrates the effect and disease burden of AS across the domains studied relative to the general population, by comparing AS scores to population normative values. Because PROMIS measures are more widely used in clinical trials and US clinical care, the construct validity of these measures in axSpA will be increasingly relevant. Future work will examine the longitudinal construct validity and discrimination of these instruments in treatment initiation scenarios and continue to elucidate how PROMIS instruments can be used to understand the effect of AS in different cultural and societal contexts.

ACKNOWLEDGMENT

The authors are deeply grateful to the patients and their families for their participation and support in this study, and to the staff at the Division of Rheumatology and Clinical Immunogenetics, Center for Clinical Research and Evidence-based Medicine, and Center for Clinical and Translational Sciences at UTHealth. We also thank the PSOAS investigators for their continued collaboration.

ONLINE SUPPLEMENT

Supplementary material accompanies the online version of this article.

REFERENCES

1. Minnock P, Kirwan J, Bresnihan B. Fatigue is a reliable, sensitive and unique outcome measure in rheumatoid arthritis. Rheumatology 2009;48:1533-6.
2. Wells G, Li T, Maxwell L, Maclean R, Tugwell P. Responsiveness of patient reported outcomes including fatigue, sleep quality, activity limitation, and quality of life following treatment with abatacept for rheumatoid arthritis. Ann Rheum Dis 2008;67:260-5.
3. van Tuyl LH, Boers M. Patient-reported outcomes in core domain sets for rheumatic diseases. Nat Rev Rheumatol 2015;11:705-12.
4. Stolwijk C, van Onna M, Boonen A, van Tubergen A. Global prevalence of spondyloarthritis: a systematic review and meta-regression analysis. Arthritis Care Res 2016;68:1320-31.
5. van der Heijde D, Dougados M, Davis J, Weisman MH, Maksymowych W, Braun J, et al. Assessment in Ankylosing Spondylitis International Working Group/Spondylitis Association of America recommendations for conducting clinical trials in ankylosing spondylitis. Arthritis Rheum 2005;52:386-94.
6. Sieper J, Rudwaleit M, Baraliakos X, Brandt J, Braun J, Burgos-Vargas R, et al. The Assessment of Spondyloarthritis international Society (ASAS) handbook: a guide to assess spondyloarthritis. Ann Rheum Dis 2009;68 Suppl 2:ii1-44.
7. Universal health outcome measures for older persons with multiple chronic conditions. J Am Geriatr Soc 2012;60:2333-41.
8. Ader DN. Developing the Patient Reported Outcomes Measurement Information System (PROMIS). Med Care 2007;45:S1-2.
9. Khanna D, Krishnan E, Dewitt EM, Khanna PP, Spiegel B, Hays RD. The future of measuring patient-reported outcomes in rheumatology: Patient-reported Outcomes Measurement Information System (PROMIS). Arthritis Care Res 2011;63 Suppl 11:S486-90.
10. Baumhauer JF. Patient-reported outcomes — are they living up to their potential? N Engl J Med 2017;377:6-9.
11. Garrett S, Jenkinson T, Kennedy LG, Whitelock H, Gaisford P, Calin A. A new approach to defining disease status in ankylosing spondylitis: The Bath Ankylosing Spondylitis Disease Activity Index. J Rheumatol 1994;21:2286-91.
12. Calin A, Garrett S, Whitelock H, Kennedy LG, O'Hea J, Mallorie

Personal non-commercial use only. The Journal of Rheumatology Copyright © 2020. All rights reserved.

Downloaded on April 23, 2024 from www.jrheum.org

P, et al. A new approach to defining functional ability in ankylosing spondylitis: the development of the Bath Ankylosing Spondylitis Functional Index. J Rheumatol 1994;21:2281-5.
13. Radloff LS. The CES-D scale: a self-report depression scale for research in the general population. Appl Psychol Meas 1977; 1:385-401.
14. Machado P, Landewé R, Lie E, Kvien TK, Braun J, Baker D, et al. Ankylosing spondylitis disease activity score (ASDAS): Defining cut-off values for disease activity states and improvement scores. Ann Rheum Dis 2011;70:47-53.
15. Cella D, Yount S, Rothrock N, Gershon R, Cook K, Reeve B, 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:S3-11.
16. Bliss CI. Statistics in biology; statistical methods for research in the natural sciences. New York: McGraw-Hill; 1967.
17. Brandon TG, Becker BD, Bevans KB, Weiss PF. Patient-reported outcomes measurement information system tools for collecting patient-reported outcomes in children with juvenile arthritis. Arthritis Care Res 2017;69:393-402.
18. Broderick JE, Schneider S, Junghaenel DU, Schwartz JE, Stone AA. Validity and reliability of patient-reported outcomes measurement information system instruments in osteoarthritis. Arthritis Care Res 2013;65:1625-33.
19. Orbai AM, Mease PJ, de Wit M, Kalyoncu U, Campbell W, Tillett W, et al. Report of the GRAPPA-OMERACT Psoriatic Arthritis Working Group from the GRAPPA 2015 Annual Meeting. J Rheumatol 2016;43:965-9.
20. Bartlett SJ, Orbai AM, Duncan T, DeLeon E, Ruffing V, Clegg-Smith K, et al. Reliability and validity of selected PROMIS measures in people with rheumatoid arthritis. PLoS One 2015;10:e0138543.
21. Kasturi S, Szymonifka J, Burket JC, Berman JR, Kirou KA, Levine AB, et al. Validity and reliability of patient reported outcomes measurement information system computerized adaptive tests in systemic lupus erythematosus. J Rheumatol 2017;44:1024-31.
22. Khanna D, Maranian P, Rothrock N, Cella D, Gershon R, Khanna PP, et al. Feasibility and construct validity of PROMIS and "legacy" instruments in an academic scleroderma clinic. Value Health 2012;15:128-34.
23. Robson JC, Tomasson G, Milman N, Ashdown S, Boonen A, Casey GC, et al. OMERACT endorsement of patient-reported outcome instruments in antineutrophil cytoplasmic antibody-associated vasculitis. J Rheumatol 2017;44:1529-35.
24. Driban JB, Morgan N, Price LL, Cook KF, Wang C. Patient-reported outcomes measurement information system (PROMIS) instruments among individuals with symptomatic knee osteoarthritis: a cross-sectional study of floor/ceiling effects and construct validity. BMC Musculoskelet Disord 2015;16:253.
25. Bartlett SJ, Gutierrez AK, Butanis A, Bykerk VP, Curtis JR, Ginsberg S, et al. Combining online and in-person methods to evaluate the content validity of PROMIS fatigue short forms in rheumatoid arthritis. Qual Life Res 2018;27:2443-51.
26. Kasturi S, Szymonifka J, Burket JC, Berman JR, Kirou KA, Levine AB, et al. Feasibility, validity, and reliability of the 10-item Patient Reported Outcomes Measurement Information System Global Health Short Form in outpatients with systemic lupus erythematosus. J Rheumatol 2018;45:397-404.
27. Katz P, Yazdany J, Trupin L, Rush S, Helmick CG, Murphy LB, et al. Psychometric evaluation of the National Institutes of Health Patient-Reported Outcomes Measurement Information System in a multiracial, multiethnic systemic lupus erythematosus cohort. Arthritis Care Res 2019;71:1630-9.
28. Wohlfahrt A, Bingham CO 3rd, Marder W, Phillips K, Bolster MB, Moreland LW, et al. Responsiveness of Patient- Reported Outcomes Measurement Information System measures in rheumatoid arthritis patients starting or switching a disease-modifying antirheumatic drug. Arthritis Care Res 2019;71:521-9.
29. Mahieu MA, Ahn GE, Chmiel JS, Dunlop DD, Helenowski IB, Semanik P, et al. Fatigue, patient reported outcomes, and objective measurement of physical activity in systemic lupus erythematosus. Lupus 2016;25:1190-9.

Personal non-commercial use only. The Journal of Rheumatology Copyright © 2020. All rights reserved.

Downloaded on April 23, 2024 from www.jrheum.org