



Published in final edited form as:

J Pediatr Orthop. 2022 August 01; 42(7): e720–e726. doi:10.1097/BPO.0000000000002190.

Construct Validity of Pediatric PROMIS Computerized Adaptive Testing Measures in Children with Adolescent Idiopathic Scoliosis

Stuart L. Mitchell, MD^{1,2}, Kevin H. McLaughlin, DPT³, Keith R. Bachmann, MD⁴, Paul D. Sponseller, MD, MBA^{1,*}, Lisa M. Reider, PhD^{5,*}

¹Department of Orthopaedic Surgery, The Johns Hopkins University School of Medicine, Baltimore, MD

²Department of Orthopaedic Surgery, University of North Carolina, NC

³Department of Physical Medicine and Rehabilitation, The Johns Hopkins University School of Medicine, Baltimore, MD

⁴Department of Orthopaedic Surgery, University of Virginia, Charlottesville, VA

⁵Department of Health Policy and Management, The Johns Hopkins Bloomberg School of Public Health, Baltimore, MD

Abstract

Background: The use of patient-reported outcome measures, especially Patient-Reported Outcomes Measurement Information System (PROMIS) measures, has increased in recent years. Given this growth, it is imperative to ensure that the measures being used are validated for the intended population(s)/disease(s). Our objective was to assess the construct validity of 8 PROMIS computer adaptive testing (CAT) measures among children with adolescent idiopathic scoliosis (AIS).

Methods: We prospectively enrolled 200 children (aged 10–17 years) with AIS, who completed 8 PROMIS CATs (Anxiety, Depressive Symptoms, Mobility, Pain Behavior, Pain Interference, Peer Relationships, Physical Activity, Physical Stress Experiences) and the Scoliosis Research Society-22r questionnaire (SRS-22r) electronically. Treatment categories were observation, bracing, indicated for surgery, or postoperative from posterior spinal fusion. Construct validity was evaluated using known group analysis and convergent and discriminant validity analyses. ANOVA was used to identify differences in PROMIS T-scores by treatment category (known groups). Spearman's rank correlation coefficient (r_s) was calculated between corresponding PROMIS and SRS-22r domains (convergent) and between unrelated PROMIS domains (discriminant). Floor/ceiling effects were calculated.

Address all correspondence and reprint requests to: Stuart L. Mitchell, MD; 130 Mason Farm Rd, CB# 7055, UNC School of Medicine, Chapel Hill, NC 27599-7055; Stuart.L.Mitchell@gmail.com.

* Drs. Sponseller and Reider should be considered co-senior authors.

Conflict of Interest: The authors have no conflicts of interest related to this research.

Approval Statement: Institutional review board approval (IRB00165159) was received for this study.

Results: Among treatment categories, significant differences were found in PROMIS Mobility, Pain Behavior, Pain Interference, and Physical Stress Experiences and in all SRS-22r domains ($p < 0.05$) except Mental Health ($p = 0.15$). SRS-22r Pain was strongly correlated with PROMIS Pain Interference ($r_s = -0.72$) and Pain Behavior ($r_s = -0.71$) and moderately correlated with Physical Stress Experiences ($r_s = -0.57$). SRS-22r Mental Health was strongly correlated with PROMIS Depressive Symptoms ($r_s = -0.72$) and moderately correlated with Anxiety ($r_s = -0.62$). SRS-22r Function was moderately correlated with PROMIS Mobility ($r_s = 0.64$) and weakly correlated with Physical Activity ($r_s = 0.34$). SRS-22r Self-Image was weakly correlated with PROMIS Peer Relationships ($r_s = 0.33$). All unrelated PROMIS CATs were weakly correlated ($|r_s| < 0.40$). PROMIS Anxiety, Mobility, Pain Behavior, and Pain Interference and SRS-22r Function, Pain, and Satisfaction displayed ceiling effects.

Conclusions: Evidence supports the construct validity of 6 PROMIS CATs in evaluating AIS patients. Ceiling effects should be considered when using specific PROMIS CATs.

Level of Evidence: II, Prognostic

Keywords

adolescent idiopathic scoliosis; computer adaptive test; patient-reported outcomes; Patient-Reported Outcomes Measurement Information System; Scoliosis Research Society-22r health questionnaire

INTRODUCTION

Patient-reported outcome measures (PROMs) are important tools for evaluating orthopaedic patients.¹ In patients with adolescent idiopathic scoliosis (AIS), patient-reported outcomes (PROs) are frequently assessed using the widely accepted Scoliosis Research Society-22r (SRS-22r) health questionnaire.² Although it is psychometrically sound,³⁻⁶ the SRS-22r cannot be used to compare PROs in spinal deformities with other disorders because it is a disease-specific measure. Many anatomic or disease-specific PROMs (i.e., “legacy” measures) have deficiencies and cannot be used to compare PROs across conditions.¹ The Patient-Reported Outcomes Measurement Information System (PROMIS) is a set of person-centered measures developed to overcome some deficiencies of legacy PROMs.^{7,8} PROMIS was designed to be psychometrically sound, unidimensional, efficient, generalizable, and relevant across many conditions.^{7,8} PROMIS was developed using item response theory (IRT), enabling the use of computerized adaptive testing (CAT).^{1,9} IRT and CAT differentiate PROMIS from legacy measures because they enable more efficient data collection and reduce test-taker burden.^{1,9,10} These factors, in turn, improve the feasibility of using PROMs in routine clinical care. Incorporation of PROMs in routine clinical care improves patient engagement, patient-provider communication and shared decision-making, patient and provider satisfaction, and quality and performance of medical care.¹¹⁻¹⁴

Given the growing use of PROMIS and other PROMs in pediatric orthopaedic populations,^{2,15} it is imperative to ensure that these measures are valid in the population of interest. Most PROMs used in pediatric orthopaedics are not designed or validated for use in children.¹⁶⁻¹⁸ PROMIS is considered “partially validated” because the measures have

been validated in a general pediatric population^{12,17}. This allows comparison of scores in many conditions to a general population. However, PROMIS has not been fully validated in specific populations, such as AIS patients. Thus, we sought to contribute to the continuum of “validated” use of PROMIS in AIS. If proven valid for use in this population, PROMIS measures could be used to improve clinicians’ ability to monitor patients’ health-related quality of life (HRQoL) during routine clinical care. Our objective was to assess the construct validity of 8 pediatric PROMIS CATs measuring physical, mental, and social health in children with AIS.

MATERIALS AND METHODS

In this cross-sectional validation study, we prospectively enrolled English-speaking adolescents aged 10–17 years with AIS during routine outpatient visits at one US academic hospital from August 2018 to March 2020. Patients with scoliosis diagnosed before 10 years of age (e.g., juvenile idiopathic scoliosis) were excluded. Institutional review board approval was obtained. Informed consent was obtained from guardians, and assent was obtained from participants.

Participants completed 8 PROMIS CAT measures and the SRS-22r, in random order, on iPads, using the REDCap interface.^{15,19} The participants’ familiarity with using an iPad was not assessed, but a research assistant was available to assist with any technological issues. Additionally, participants’ understanding of the survey questions was not assessed because the SRS-22r has been validated in this population,^{3–6} and the PROMIS CATs assessed have been validated in children aged 8–17 years in the general US population. The PROMIS CAT measures were Physical Activity, Mobility, Anxiety, Depressive Symptoms, Peer Relationships, Physical Stress Experiences, Pain Behavior, and Pain Interference. These domains have been identified by surgeons, parents, and children as most relevant to individuals with AIS in a stakeholder survey. PROMIS measures are reported as T-scores (0–100 scale) with a mean (\pm standard deviation) of 50 ± 10 . All measures except Pain Behavior (from a clinical calibration cohort) are normalized to the general US population aged 8–17 years. PROMIS T-scores represent the amount of a concept being measured. Higher scores for negatively worded concepts, like Pain Behavior, indicate worse function. SRS-22r comprises 4 domains (Function, Pain, Self-Image, and Mental Health) and 2 management satisfaction questions. Each question has an integer score ranging from 1–5. The mean score for each domain and mean total score range from 1 (worst) to 5 (best).

Patient characteristics, medical history, and spinal curve location(s) and magnitude(s) were documented from medical records. Treatment categories were 1) observation (including physical therapy, chiropractic care, completed bracing); 2) bracing; 3) indicated for posterior spinal fusion (PSF) surgery; and 4) postoperative (after PSF). Patients who were undergoing observation or other forms of nonoperative care (other than current bracing) were included in the “observation” category to maximize data available for analysis because of the small number of patients in each group. Enrollment continued until the number of patients who had completed all surveys reached 200, which is the recommended sample size for initial validation studies.²⁰ Only participants who completed all surveys were included in our analysis (Figure 1).

Participant Characteristics

Mean (\pm standard deviation) age of the 200 participants (83% female) was 14 ± 1.6 years (Table I). The mean major curve magnitude, measured using the Cobb method, was $33\pm 15^\circ$, and most were main thoracic curves (62%). At enrollment, 48 participants (24%) were undergoing observation, 94 (47%) were undergoing bracing, 32 (16%) were indicated for surgery, and 26 (13%) had undergone PSF. The mean time since PSF in the postoperative group was 1.4 ± 1.6 years (range, 6 weeks to 5.3 years).

Statistical Analysis

Descriptive statistics were used to summarize demographic, radiographic, and clinical data. Construct validity was evaluated with known group analysis and convergent and discriminant validity analyses. Analysis of variance (ANOVA) was used to compare PROMIS T-scores by treatment category (known group analysis). Spearman's correlation coefficients (r_s) were calculated between corresponding PROMIS and SRS-22r domains (convergent) and between unrelated PROMIS domains (discriminant). Correlational analysis results from 0–0.39 are considered weak, 0.4–0.69 are moderate, 0.7–0.99 are strong, and 1 is perfect.^{21,22} We expected at least moderate ($|r_s| \geq 0.40$) correlation between SRS-22r Function and PROMIS Physical Activity and Mobility; between SRS-22r Mental Health and PROMIS Anxiety and Depressive Symptoms; and between SRS-22r Pain and PROMIS Physical Stress Experiences, Pain Behavior, and Pain Interference. PROMIS Peer Relationships was chosen as an experimental corresponding domain to SRS-22r Self-Image, with an expected moderate correlation. We expected weak correlation among unrelated PROMIS domains. Floor and ceiling effects were calculated as the proportion of individuals with the lowest (floor) and highest (ceiling) level of function for each PROMIS and SRS-22r domain.^{22–24} Floor or ceiling effects were considered present if $>15\%$ of participants scored the minimum or maximum observed scores, respectively.^{22,25,26} For negatively worded PROMIS constructs (e.g., Pain Interference), the minimum observed score was considered the ceiling and maximum was considered the floor. Analyses were performed using Stata, version 15, software (StataCorp LLC, College Station, TX). Significance was considered at $\alpha=0.05$.

RESULTS

Known Groups

PROMIS T-scores are summarized in Table II. Among treatment categories, we found significant differences in PROMIS Mobility, Physical Stress Experiences, Pain Behavior, and Pain Interference ($p<0.05$) and in all SRS-22r domains except Mental Health ($p=0.15$). In general, more dysfunction was observed in patients who were indicated for or had undergone surgery.

Convergent and Discriminant Validity

PROMIS Pain Interference ($r_s=-0.72$) and Pain Behavior ($r_s=-0.71$) were strongly correlated, and Physical Stress Experiences ($r_s=-0.57$) was moderately correlated with SRS-22r Pain (Table III). PROMIS Depressive Symptoms was strongly correlated ($r_s=-0.72$)

and Anxiety was moderately correlated ($r_s=-0.62$) with SRS-22r Mental Health. PROMIS Mobility was moderately correlated ($r_s=0.64$) and Physical Activity was weakly correlated ($r_s=0.34$) with SRS-22r Function. PROMIS Peer Relationships was weakly correlated ($r_s=0.33$) with SRS-22r Self-Image.

Discriminant validity was demonstrated by weak correlation among PROMIS CAT measures that should theoretically be unrelated. All unrelated PROMIS CAT measures were only weakly correlated ($|r_s|<0.40$; Table IV). For example, PROMIS Peer Relationships and PROMIS Physical Activity, a measure of self-reported amount of physical activity, had weak or no correlation with any other PROMIS CAT measure.

Floor/Ceiling Effects

Floor effects were absent for all PROMIS CAT measures and SRS-22r domains (<5%) (Table V). Ceiling effects were present for PROMIS Pain Interference (40%), Mobility (34%), and Anxiety (17%) and for SRS-22r Function (30%), Pain (26%), and Satisfaction (20%).

DISCUSSION

This is the largest, and only cross-sectional study to assess multiple aspects of construct validity of PROMIS CAT measures in a cohort of exclusively pediatric-aged AIS patients. Additionally, we evaluated more CAT measures than previous validation studies^{22,27,28} in this population. Our results provide initial construct validity for at least 6 PROMIS CAT measures to evaluate HRQoL in children and adolescents with AIS. Thus, these measures can comfortably be used in the evaluation of patients with AIS to assess their HRQoL. In the current versions, several PROMIS CAT measures exhibited ceiling effects in these typically high-functioning AIS patients and must be considered when used in this population.

PROMs have been a foundation of clinical research for years, but as health policy and reimbursement decisions are increasingly incorporating patient outcomes, it will become necessary to incorporate PROMs in routine clinical care of patients. The use of PROMs, such as PROMIS, in routine clinical care has many benefits to the patient and the care team. PROMs provide an opportunity to screen for certain symptoms or conditions or to serially monitor a patient's condition.²⁹ Serial monitoring can track a patient's progress in real time and can be used to make a "roadmap to recovery" after surgery or other treatment intervention that predicts functioning or symptoms in certain domains over time.³⁰ PROMs in routine clinical care can also improve patient engagement, patient-provider communication and shared decision-making, patient and provider satisfaction, and quality and performance of medical care.¹¹⁻¹⁴

Because PROMIS is a set of general PROMs (vs. disease-specific or anatomic), it may offer a more complete evaluation of a patient's HRQoL and overall well-being than the SRS-22r does.^{22,27} Unlike the SRS-22r, which is a disease-specific instrument, PROMIS measures can be used to compare PROs in patients with AIS to other disease states. PROMIS also provides additional information with constructs that are not captured by the SRS-22r and allows comparison with the general population. For example, we can accurately state

that, on average, the adolescents in our study felt that their peer relationships were no different than those of their healthy peers and were not significantly different based on treatment. Also, no patients screened positive for depression or an anxiety disorder because no patients scored ≥ 2 standard deviations above average/worse function (50 points) on either measure. This type of assessment cannot be made using the SRS-22r. Additionally, PROMIS can assess more nuanced ideas or constructs, as exemplified by PROMIS Physical Stress Experiences. This measure assesses consciously perceived sensations in response to stressors, which include physical arousal (e.g., sensory alertness, muscle potential), agitation (e.g., restlessness, fidgetiness), pain, sleep disturbance, and gastrointestinal distress.^{31,32} In our study, postoperative patients scored almost 1 standard deviation higher than the general population on Physical Stress Experiences, and this was also significantly higher than patients undergoing bracing.

Convergent and Discriminant Validity

We observed at least moderate correlation with corresponding SRS-22r domains among all but 2 PROMIS CAT measures (Peer Relationships, Physical Activity). Previous retrospective studies have assessed correlations between only PROMIS Mobility, Pain Interference, Depression, and Peer Relationships and their corresponding SRS-22r domains.^{22,27} Notably, the methods of these studies differed from ours. Fedorak et al.²⁷ assessed 113 pediatric patients with AIS who completed PROMIS short-forms (static measures), and Bernstein et al.²² included 64 pediatric patients and adults with AIS who completed PROMIS CAT measures. Although there are small differences among study findings, we all observed similar correlations among the selected PROMIS measures and their corresponding SRS-22r domains.

PROMIS Peer Relationships correlated weakly with SRS-22r Mental Health and Self-Image. Although PROMIS Peer Relationships was considered the corresponding measure to SRS-22r Self-Image, this comparison was made out of experimental curiosity because no PROMIS Self-Image measure exists and it was considered the closest match. Additionally, PROMIS Peer Relationships is likely an important aspect of HRQoL in patients with AIS given the temporal relationship with adolescent social development.

PROMIS Physical Activity also correlated only weakly with the corresponding SRS-22r Function domain. However, this finding likely reflects a difference in the constructs being measured by these domains. PROMIS Physical Activity aims to capture self-reported engagement in physical activity (e.g., amount, intensity) but not the types of activities being performed. This construct that is captured by the SRS-22r, the ability to engage in certain types of activities, is better captured by PROMIS Mobility. The construct captured by PROMIS Physical Activity is not assessed well, if at all, by the SRS-22r.

We also found appropriately weak correlation among all unrelated PROMIS CAT measures, as expected, which supports the PROMIS CAT's ability to discriminate against dissimilar constructs.

Known Group Analyses

Known group validity is a means of assessing an instrument's ability to differentiate among clinically distinct groups, often on the basis of severity.³³ The only published data regarding known groups were part of a recent study by Yau et al.,²⁸ who compared 5 PROMIS CAT measures with commonly used legacy PROMs. They compared PROMIS Mobility, Pain Interference, Physical Activity, Physical Stress Experiences, and Psychological Stress Experiences with SRS-22r and other legacy PROMs to ascertain differences according to spinal curve severity. They reported that certain legacy PROMs and PROMIS CAT measures were able to differentiate between patients with small curves (0°–40°) vs. patients with severe curves (>40°). We also found that certain PROMIS CAT measures and SRS-22r domains were able to differentiate between known groups on the basis of treatment category as a different assessment of severity. We compared treatment groups because our cohort included all forms of treatment, as well as postoperative patients. Assessment of curve magnitude becomes less relevant after surgical correction, so this comparison was considered less helpful than treatment categories.

Floor/Ceiling Effects

It is important for PROMs to have minimal floor/ceiling effects to ensure they can effectively differentiate among patients who are functioning at the lowest or highest ends of the spectrum, respectively.³⁴ Neither PROMIS nor the SRS-22r showed any floor effects. Four PROMIS and 3 SRS-22r domains displayed ceiling effects; thus, were unable to differentiate among the patients who were functioning well (at the highest end of the continuum). In the only other exclusively pediatric cohort, Fedorak et al.²⁷ also observed ceiling effects for PROMIS Mobility and Pain Interference. The ceiling effects we found may be explained in part by the method used to calculate floor/ceiling effects. We used a conservative analysis approach, thereby showing more floor/ceiling effects if they do, in fact, exist. An alternative method using the highest and lowest possible scores would have identified 0% ceiling and 0% floor effects for all PROMIS CAT measures analyzed because PROMIS T-scores have a theoretical range of 0–100. However, a score of 0 or 100 cannot be achieved in the current versions of any of the PROMIS CAT measures in our study. Additionally, the ceiling effects may also be related to the overall high functioning of these patients because most patients with AIS do not have major pain or functional limitations. Conversely, PROMIS Physical Activity, Depressive Symptoms, and Physical Stress Experiences did not display floor/ceiling effects, and therefore, should be able to differentiate among nearly all high- and low-functioning patients with AIS. Additionally, all PROMIS CAT measures and SRS-22r domains were able to differentiate among patients functioning at the lowest end of the spectrum.

Strengths and Weaknesses

We used a cross-sectional design, enrolled a large patient sample, and assessed 8 PROMIS CAT measures. Inclusion of patients undergoing all forms of treatment or monitoring for AIS improves the generalizability of our findings. Although it may be considered a weakness that postoperative patients were included at their first follow-up visit, 6 weeks after surgery, we think it is important to have included these patients to ensure that the

measures tested were validated in patients being treated along the entire spectrum of AIS. Furthermore, we assessed multiple aspects of construct validity. However, our study was performed at 1 urban tertiary referral center with patients from 1 surgeon. We included only English-speaking patients as most of our patient population speaks English as their primary language; therefore, our findings may not apply to non-English speaking patients and those from other cultures. Lastly, only 77% of patients who enrolled completed survey items, which may reflect cultural differences considering that 95% of patients (N=206) at a center in New York City completed all 8 of their measures.²⁸

One unaddressed concern is that PROMIS lacks a self-image domain. Self-image is an important aspect of AIS care and changes in response to treatment and could be an area for future PROMIS development.^{22,35} Others have suggested supplementing PROMIS with SRS-22r Self-Image questions.³⁵ Alternatively, one might use an appearance PROM that has been validated in AIS, such as the Spinal Appearance Questionnaire³⁶. Another area that remains unclear is which PROMIS CATs to collect. Certain CATs may be more useful than others for clinical decision-making based on patient population. This is an important topic for future investigation.

In pediatric patients with AIS, we found moderate to strong correlation for 6 of 8 PROMIS CAT measures with their corresponding SRS-22r domains and appropriately weak correlation among unrelated PROMIS CAT measures. Additionally, PROMIS Mobility, Physical Stress Experiences, Pain Behavior, and Pain Interference can differentiate among known groups of severity on the basis of treatment category. Importantly, 4 PROMIS CATs and similar corresponding SRS-22r domains displayed ceiling effects and should be considered when choosing these measures. PROMIS Self-Image CAT development is an opportunity for future research. Our study provides evidence of initial construct validity for selected pediatric PROMIS CAT measures to evaluate HRQoL in patients with AIS.

Acknowledgments:

The authors thank Anthony Carlini, MS, for his assistance with data collection and organization, Lucie Wiedefeld for her help with patient enrollment and data collection, and Kristen Venuti, CRNP, for her help with patient enrollment. Additionally, the authors thank Rachel Box, MS, for assistance with manuscript editing and submission.

Funding Statement:

This work was supported in part by T32 grant (no. AR067708) from the National Institutes of Health (Bethesda, MD); and support from the Coordinating Center of the Major Extremity Trauma and Rehabilitation Consortium (METRC).

REFERENCES

1. Brodke DJ, Saltzman CL, Brodke DS. PROMIS for orthopaedic outcomes measurement. *J Am Acad Orthop Surg.* 2016;24:744–749. [PubMed: 27661391]
2. Cutler HS, Guzman JZ, Al Maaieh M, et al. Patient reported outcomes in adult spinal deformity surgery: a bibliometric analysis. *Spine Deform.* 2015;3:312–317. [PubMed: 27927475]
3. Asher M, Min Lai S, Burton D, et al. Discrimination validity of the scoliosis research society-22 patient questionnaire: relationship to idiopathic scoliosis curve pattern and curve size. *Spine (Phila Pa 1976).* 2003;28:74–78. [PubMed: 12544960]

4. Asher M, Min Lai S, Burton D, et al. Scoliosis research society-22 patient questionnaire: responsiveness to change associated with surgical treatment. *Spine (Phila Pa 1976)*. 2003;28:70–73. [PubMed: 12544959]
5. Asher M, Min Lai S, Burton D, et al. The reliability and concurrent validity of the scoliosis research society-22 patient questionnaire for idiopathic scoliosis. *Spine (Phila Pa 1976)*. 2003;28:63–69. [PubMed: 12544958]
6. Berven S, Deviren V, Demir-Deviren S, et al. Studies in the modified Scoliosis Research Society Outcomes Instrument in adults: validation, reliability, and discriminatory capacity. *Spine (Phila Pa 1976)*. 2003;28:2164–2169; discussion 2169. [PubMed: 14501930]
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;63:1179–1194. [PubMed: 20685078]
8. 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:S3–S11.
9. Brodke DJ, Hung M, Bozic KJ. Item response theory and computerized adaptive testing for orthopaedic outcomes measures. *J Am Acad Orthop Surg*. 2016;24:750–754. [PubMed: 27623193]
10. Cheung EC, Moore LK, Flores SE, et al. Correlation of PROMIS with orthopaedic patient-reported outcome measures. *JBJS Rev*. 2019;7:e9.
11. Baumhauer JF, Bozic KJ. Value-based healthcare: patient-reported outcomes in clinical decision making. *Clin Orthop Relat Res*. 2016;474:1375–1378. [PubMed: 27052020]
12. Bernstein DN, Fear K, Mesfin A, et al. Patient-reported outcomes use during orthopaedic surgery clinic visits improves the patient experience. *Musculoskeletal Care*. 2019;17:120–125. [PubMed: 30615249]
13. Bozic KJ, Belkora J, Chan V, et al. Shared decision making in patients with osteoarthritis of the hip and knee: results of a randomized controlled trial. *J Bone Joint Surg Am*. 2013;95:1633–1639. [PubMed: 24048550]
14. Lowry KJ, Brox WT, Naas PL, et al. Musculoskeletal-based Patient-reported Outcome Performance Measures, Where Have We Been-Where Are We Going. *J Am Acad Orthop Surg*. 2019;27:e589–e595. [PubMed: 31232793]
15. Harris PA, Taylor R, Thielke R, et al. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform*. 2009;42:377–381. [PubMed: 18929686]
16. Phillips L, Carsen S, Vasireddi A, et al. Use of patient-reported outcome measures in pediatric orthopaedic literature. *J Pediatr Orthop*. 2018;38:393–397. [PubMed: 27603185]
17. Arguelles GR, Shin M, Lebrun DG, et al. The majority of patient-reported outcome measures in pediatric orthopaedic research are used without validation. *J Pediatr Orthop*. 2021;41:e74–e79. [PubMed: 32804862]
18. Truong WH, Price MJ, Agarwal KN, et al. Utilization of a wide array of nonvalidated outcome scales in pediatric orthopaedic publications: Can't we all measure the same thing? *J Pediatr Orthop*. 2019;39:e153–e158. [PubMed: 30300277]
19. Harris PA, Taylor R, Minor BL, et al. The REDCap consortium: building an international community of software platform partners. *J Biomed Inform*. 2019;95:103208. [PubMed: 31078660]
20. Frost MH, Reeve BB, Liepa AM, et al. What is sufficient evidence for the reliability and validity of patient-reported outcome measures? *Value Health*. 2007;10 Suppl 2:S94–S105. [PubMed: 17995479]
21. Akoglu H User's guide to correlation coefficients. *Turk J Emerg Med*. 2018;18:91–93. [PubMed: 30191186]
22. Bernstein DN, Papuga MO, Sanders JO, et al. Evaluating the correlation and performance of PROMIS to SRS questionnaires in adult and pediatric spinal deformity patients. *Spine Deform*. 2019;7:118–124. [PubMed: 30587304]
23. Beckmann JT, Hung M, Bounsanga J, et al. Psychometric evaluation of the PROMIS Physical Function Computerized Adaptive Test in comparison to the American Shoulder and Elbow

- Surgeons score and Simple Shoulder Test in patients with rotator cuff disease. *J Shoulder Elbow Surg.* 2015;24:1961–1967. [PubMed: 26321484]
24. Gullede CM, Lizzio VA, Smith DG, et al. What are the floor and ceiling effects of Patient-Reported Outcomes Measurement Information System computer adaptive test domains in orthopaedic patients? A systematic review. *Arthroscopy.* 2020;36:901–912.e7. [PubMed: 31919023]
 25. Anthony CA, Glass NA, Hancock K, et al. Performance of PROMIS instruments in patients with shoulder instability. *Am J Sports Med.* 2017;45:449–453. [PubMed: 28146398]
 26. Terwee CB, Bot SDM, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol.* 2007;60:34–42. [PubMed: 17161752]
 27. Fedorak GT, Larkin K, Heflin JA, et al. Pediatric Patient-Reported Outcomes Measurement Information System is equivalent to Scoliosis Research Society-22 in assessing health status in adolescent idiopathic scoliosis. *Spine (Phila Pa 1976).* 2019;44:E1206–E1210. [PubMed: 31574066]
 28. Yau A, Heath MR, Fabricant PD. Discrimination ability of Patient Reported Outcome Measurement Information System pediatric domains compared with Scoliosis Research Society-22r and legacy patient reported outcome measures in juvenile and adolescent idiopathic scoliosis. *Spine (Phila Pa 1976).* 2020;45:1713–1719. [PubMed: 32956254]
 29. Snyder CF, Aaronson NK. Use of patient-reported outcomes in clinical practice. *Lancet.* 2009;374:369–370. [PubMed: 19647598]
 30. Baumhauer JF. Patient-Reported Outcomes - Are They Living Up to Their Potential? *N Engl J Med.* 2017;377:6–9. [PubMed: 28679102]
 31. Bevans KB, Gardner W, Pajer KA, et al. Psychometric evaluation of the PROMIS® Pediatric Psychological and Physical Stress Experiences Measures. *J Pediatr Psychol.* 2018;43:678–692. [PubMed: 29490050]
 32. Laurent J, Catanzaro SJ, Joiner TE. Development and preliminary validation of the physiological hyperarousal scale for children. *Psychol Assess.* 2004;16:373–380. [PubMed: 15584796]
 33. Deshpande PR, Rajan S, Sudeepthi BL, et al. Patient-reported outcomes: a new era in clinical research. *Perspect Clin Res.* 2011;2:137–144. [PubMed: 22145124]
 34. Lim CR, Harris K, Dawson J, et al. Floor and ceiling effects in the OHS: an analysis of the NHS PROMs data set. *BMJ Open.* 2015;5:e007765.
 35. Raad M, Jain A, Huang M, et al. Validity and responsiveness of PROMIS in adult spinal deformity: the need for a self-image domain. *Spine J.* 2019;19:50–55. [PubMed: 30053523]
 36. Matamalas A, Bagó J, D'Agata E, et al. Body image in idiopathic scoliosis: a comparison study of psychometric properties between four patient-reported outcome instruments. *Health Qual Life Outcomes.* 2014;12:81. [PubMed: 24894714]

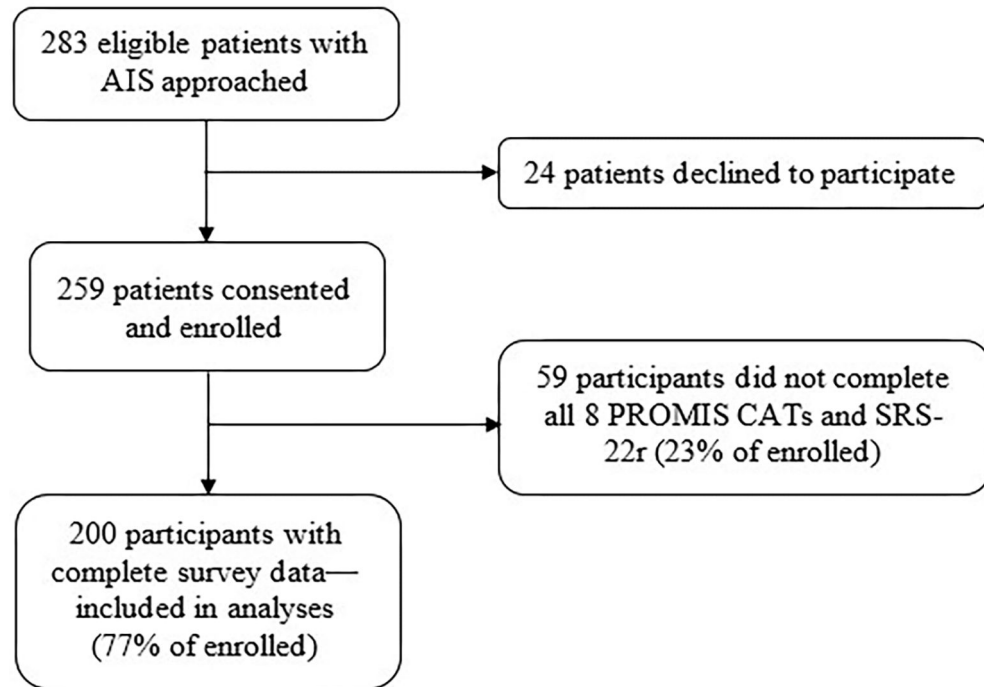


Figure 1. Flow diagram of all adolescent idiopathic scoliosis (AIS) patients approached for enrollment. (PROMIS CAT, Patient-Reported Outcomes Measurement Information System computerized adaptive testing; SRS-22r, Scoliosis Research Society-22r health questionnaire)

TABLE I.

Characteristics of 200 patients with adolescent idiopathic scoliosis who completed all 8 Patient-Reported Outcomes Measurement Information System computerized adaptive testing measures and the Scoliosis Research Society-22r health questionnaire.

Characteristic	N (%)	Mean \pm SD
Age, years		14 \pm 1.6
Female sex	166 (83)	
Race/ethnicity		
White	147 (74)	
Black	30 (15)	
Asian	11 (5.5)	
Other race/ethnicity [*]	12 (6.0)	
BMI		21 \pm 4.2
BMI percentile, %		
<5	6 (3.0)	
5–84.9	164 (82)	
85–94.9	14 (7.0)	
95	16 (8.0)	
1 Comorbidity [†]	44 (22)	
Treatment group		
Observation	48 (24)	
Bracing	94 (47)	
Indicated for surgery	32 (16)	
Postoperative	26 (13)	
Major curve magnitude, [°]		33 \pm 15
Major curve magnitude category, [°]		
<25	72 (36)	
25–44	88 (44)	
45	40 (20)	
Major curve location		
Proximal thoracic	43 (22)	
Main thoracic	125 (62)	
Thoracolumbar	17 (8.5)	
Lumbar	15 (7.5)	
Pain score [§]		1 (0–3) [‡]

BMI, body mass index; SD, standard deviation.

^{*} Other races/ethnicities were Hispanic, Native American/Native Alaskan, Pacific Island, and other.

[†] Self-reported comorbidities, which included any cardiac, respiratory, psychiatric, orthopaedic, gastrointestinal, neurologic, or “other problem.”

[‡]Data presented as median (interquartile range).

[§]On a Likert scale ranging from 0 (no pain) to 10 (worst possible pain).

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

TABLE II.

PROMIS CAT T-Scores of 200 patients with adolescent idiopathic scoliosis stratified by treatment category (i.e., known groups of severity)

Domains	Overall		Treatment Category, Mean (SD)					P*
	Mean (SD)	IQR	Observation (n=48)	Bracing (n=94)	Indicated for surgery (n=32)	Postoperative (n=26)		
PROMIS CAT								
Anxiety	48 (11)	32-68	46 (11)	47 (9.9)	49 (11)	51 (11)	0.39	
Depressive Symptoms	47 (11)	32-71	46 (11)	47 (10)	48 (11)	49 (9.8)	0.65	
Mobility	52 (8.7)	35-62	53 (7.6)	54 (8.4)	51 (8.5)	46 (9.2)	<0.001	
Pain Behavior	39 (12)	24-57	37 (12)	36 (11)	45 (8.9)	42 (13)	<0.001	
Pain Interference	42 (10)	32-66	41 (9.1)	40 (9.6)	46 (11)	48 (11)	<0.001	
Peer Relationships	52 (9.1)	34-66	53 (9.3)	53 (8.6)	48 (11)	53 (7.4)	0.07	
Physical Activity	48 (8.8)	31-67	49 (10)	48 (9.0)	47 (7.3)	44 (6.2)	0.09	
Physical Stress Experiences	54 (9.5)	36-73	55 (8.2)	53 (10)	55 (9.0)	59 (8.8)	0.04	
SRS-22r								
Function	4.5 (0.51)	4.2-5.0	4.6 (0.41)	4.5 (0.48)	4.4 (0.48)	4.2 (0.70)	<0.001	
Mental Health	4.0 (0.77)	3.4-4.6	4.1 (0.77)	4.0 (0.74)	3.8 (0.88)	3.7 (0.72)	0.15	
Pain	4.3 (0.68)	3.8-5.0	4.4 (0.57)	4.5 (0.64)	4.1 (0.62)	3.8 (0.80)	<0.001	
Satisfaction	3.9 (0.83)	3.2-4.5	3.8 (0.97)	4.0 (0.74)	3.5 (0.73)	4.5 (0.68)	<0.001	
Self-Image	3.9 (0.65)	3.4-4.4	4.2 (0.56)	3.8 (0.64)	3.5 (0.55)	4.2 (0.56)	<0.001	
Total	4.1 (0.49)	3.8-4.5	4.3 (0.47)	4.2 (0.50)	3.9 (0.47)	4.0 (0.45)		

IQR, interquartile range; PROMIS CAT, Patient-Reported Outcomes Measurement Information System computerized adaptive testing; SD, standard deviation; SRS-22r, Scoliosis Research Society-22r health questionnaire.

* Calculated with analysis of variance comparing treatment groups.

TABLE III.

Convergent validity analysis of PROMIS CAT measures and their corresponding SRS-22r domains among 200 patients with adolescent idiopathic scoliosis*

PROMIS CAT Measures	SRS-22r Domains			
	Function	Pain	Self-Image	Mental Health
Anxiety	-0.18 [†]	-0.39 [†]	-0.25 [†]	-0.62[†]
Depressive Symptoms	-0.35 [†]	-0.35 [†]	-0.42 [†]	-0.72[†]
Mobility	0.64[†]	0.63 [†]	0.42 [†]	0.43 [†]
Pain Behavior	-0.55 [†]	-0.71[†]	-0.34 [†]	-0.32 [†]
Pain Interference	-0.56 [†]	-0.72[†]	-0.33 [†]	-0.43 [†]
Peer Relationships	0.27 [†]	0.20 [†]	0.33[†]	0.38 [†]
Physical Activity	0.34[†]	0.06	0.09	0.03
Physical Stress Experiences	-0.42 [†]	-0.57[†]	-0.30 [†]	-0.50 [†]

PROMIS CAT, Patient-Reported Outcomes Measurement Information System computerized adaptive testing; SRS-22r, Scoliosis Research Society-22r health questionnaire.

* Values (r_s) generated using Spearman correlation tests. Negative values indicate an inverse relationship between domains. **Boldface** values indicate results of corresponding domain comparisons (e.g., PROMIS Physical Activity is corresponding to SRS-22r Function).

[†] Denotes significant difference ($p < 0.05$), indicating that domains are not independent.

TABLE IV.

Discriminant validity analysis of unrelated Patient-Reported Outcomes Measurement Information System computerized adaptive testing measures among 200 patients with adolescent idiopathic scoliosis *

Parameter	Physical Activity	Mobility	Peer Relationships	Anxiety	Depressive	Physical Stress Experiences	Pain Behavior	Pain Interference
Physical Activity	1							
Mobility	0.25 [†]	1						
Peer Relationships	0.11	0.23 [†]	1					
Anxiety	0.09	-0.27 [†]	-0.23 [†]	1				
Depressive Symptoms	0.01	-0.31 [†]	-0.26 [†]	0.66 [†]	1			
Physical Stress Experiences	0.05	-0.48 [†]	-0.10	0.55 [†]	0.52 [†]	1		
Pain Behavior	-0.07	-0.56 [†]	-0.14 [†]	0.34 [†]	0.31 [†]	0.57 [†]	1	
Pain Interference	-0.06	-0.70 [†]	-0.21 [†]	0.42 [†]	0.40 [†]	0.60 [†]	0.79 [†]	1

* Values (r_s) generated using Spearman correlation tests. Negative values indicate an inverse relationship between domains.

[†] Indicates significant value ($p < 0.05$) indicating that domains are not independent.

TABLE V.

Floor and ceiling effects* for pediatric PROMIS CAT and SRS-22r among 200 patients with adolescent idiopathic scoliosis

Domains	Ceiling		Floor	
	N [†] (%)	Value	N [‡] (%)	Value
PROMIS CAT				
Physical Activity	1 (0.50)	72	1 (0.50)	24
Mobility	67 (34)	62	1 (0.50)	33
Peer Relationships	27 (14)	66	1 (0.50)	19
Anxiety [§]	34 (17)	32	2 (1.0)	81
Depressive Symptoms [§]	20 (10)	32	1 (0.50)	75
Physical Stress Experiences [§]	13 (6.5)	36	1 (0.50)	81
Pain Behavior [§]	61 (30)	24	1 (0.50)	60
Pain Interference [§]	80 (40)	32	1 (0.50)	72
SRS-22r				
Function	59 (30)	5.0	2 (1.0)	3.0
Pain	51 (26)	5.0	2 (1.0)	2.4
Self-Image	18 (9.0)	5.0	1 (0.50)	2.0
Mental Health	16 (8.0)	5.0	1 (0.50)	1.0
Satisfaction	41 (20)	5.0	2 (1.0)	1.5

PROMIS CAT, Patient-Reported Outcomes Measurement Information System computerized adaptive testing; SRS-22r, Scoliosis Research Society-22r health questionnaire.

*Floor/ceiling effects are considered present if >15%.

[†]Number of patients with highest observed function.

[‡]Number of patients with the lowest possible score.

[§]Negatively worded PROMIS concepts (e.g., Pain Interference) will have lower scores (better function) that represent the ceiling value and higher scores (worse function) for the floor value.