5-27-2021

PROMIS Physical Function Instruments Compare Favorably to Legacy Patient Reported Outcome Measures in Upper and Lower Extremity Orthopedic Patients: A Systematic Review of the Literature.

Alexander C. Ziedas
Henry Ford Health, aziedas2@hfhs.org

Varag Abed
Henry Ford Health, vabed1@hfhs.org

Alexander J. Swantek
Henry Ford Health, aswante1@hfhs.org

Tahsin M. Rahman
Henry Ford Health, trahman1@hfhs.org

Austin G. Cross
Henry Ford Health, across1@hfhs.org

See next page for additional authors

Recommended Citation

Ziedas AC, Abed V, Swantek AJ, Rahman TM, Cross A, Thomashow K, and Makhni EC. PROMIS Physical Function Instruments Compare Favorably to Legacy Patient Reported Outcome Measures in Upper and Lower Extremity Orthopedic Patients: A Systematic Review of the Literature. Arthroscopy 2021.

Authors

Alexander C. Ziedas, Varag Abed, Alexander J. Swantek, Tahsin M. Rahman, Austin G. Cross, Katherine Thomashow, and Eric C. Makhni

Journal Pre-proof

PROMIS Physical Function Instruments Compare Favorably to Legacy Patient Reported Outcome Measures in Upper and Lower Extremity Orthopedic Patients: A Systematic Review of the Literature

Alexander C. Ziedas, BS, Varag Abed, BS, Alexander J. Swantek, BSE, Tahsin M. Rahman, MD, Austin Cross, BS, Katherine Thomashow, MS, Eric C. Makhni, MD MBA

PII: S0749-8063(21)00517-X

DOI: https://doi.org/10.1016/j.arthro.2021.05.031

Reference: YJARS 57654

To appear in: *Arthroscopy: The Journal of Arthroscopic and Related Surgery*

Received Date: 3 December 2020

Revised Date: 3 May 2021

Accepted Date: 11 May 2021

Please cite this article as: Ziedas AC, Abed V, Swantek AJ, Rahman TM, Cross A, Thomashow K, Makhni EC, PROMIS Physical Function Instruments Compare Favorably to Legacy Patient Reported Outcome Measures in Upper and Lower Extremity Orthopedic Patients: A Systematic Review of the Literature, *Arthroscopy: The Journal of Arthroscopic and Related Surgery* (2021), doi: https://doi.org/10.1016/j.arthro.2021.05.031.

PROMIS Physical Function Instruments Compare Favorably to Legacy Patient Reported Outcome Measures in Upper and Lower Extremity Orthopedic Patients: A Systematic Review of the Literature

Authors:

Alexander C. Ziedas, BS
Email: aziedas2@hfhs.org
Affiliation: Henry Ford Health System

Varag Abed, BS
E: vabed1@hfhs.org
Affiliation: Henry Ford Health System

Alexander J. Swantek, BSE
E: aswante1@hfhs.org
Affiliation: Henry Ford Health System

Tahsin M. Rahman, MD
E: trahman1@hfhs.org
Affiliation: Henry Ford Health System

Austin Cross, BS
E: across1@hfhs.org
Affiliation: Henry Ford Health System

Katherine Thomashow, MS
E: katherine.thomashow@med.wayne.edu
Affiliation: Henry Ford Health System

Corresponding: Eric C. Makhni, MD MBA
E: ericmakhnimd@gmail.com
P: 1-313-575-4293
Affiliation: Henry Ford Health System
Department of Orthopaedic Surgery
2799 W Grand Blvd, Detroit, MI, USA 48202

Institutional Review Board: This project did not require review by the institutional review board.

PROMIS Physical Function Instruments Compare Favorably to Legacy Patient Reported Outcome Measures in Upper and Lower Extremity Orthopedic Patients: A Systematic Review of the Literature

Abstract

Purpose: The purpose of this systematic review is to compare Patient-Reported Outcomes Measurement Information System (PROMIS) Physical Function (PF) with traditional ("legacy") patient reported outcome measures (PROMs) in regard to correlations, ease of use, and quality criteria for upper (UE) and lower extremity (LE) orthopedic conditions.

Methods: A systematic search of the PubMed/MEDLINE database was performed according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines to identify published articles that referenced the various PROMIS PF measures. Two authors independently reviewed selected studies. The search returned 857 studies, 85 of which were selected for independent review by two authors. Of these, 54 were selected for inclusion. Mixed linear models were performed to assess for differences between legacy PROMs and PROMIS measures.

Results: The combined sample size of all included studies yielded 6,074 UE and 9,366 LE patients. Overall, PROMIS PF measures demonstrated strong correlations with legacy PROMs among UE (weighted Pearson correlation, 0.624, standard error [SE] = 0.042; weighted Spearman correlation, 0.566, SE = 0.042) and LE patients (weighted Pearson correlation, 0.645, SE = 0.062; weighted Spearman correlation, 0.631, SE = 0.041). PROMIS PF questionnaires completed by UE patients had fewer questions than legacy PROMs (5.9 vs 17.7, P = 0.0093) and were completed in less time (90.5 vs 223.8 seconds, P = 0.084). PROMIS PF questionnaires completed by LE patients had fewer questions than legacy PROMs (4.81 vs 15.33, P < 0.001) and were completed in less time (63.6 vs 203.2 seconds, P = 0.0063). The differences for the reliability measures were not significant.

Conclusions: PROMIS PF scores correlate strongly with commonly used legacy PROMs in orthopedics, particularly in UE and LE patients. PROMIS PF forms can be administered efficiently and to a broad patient population while remaining highly reliable. Therefore, they can be justified for standardized use among orthopedic patients with UE and LE conditions, improving the ability to aggregate and compare outcomes in orthopedic research.

Level of Evidence: Level IV, systematic review of Level I-IV evidence.

Introduction

Patient-reported outcome measures (PROMs) are used in orthopedic surgery to assess clinical outcomes from the patient's perspective. There are many different validated PROMs reported in the orthopedic literature. Collectively, these PROMs are referred to as legacy measures. Within orthopedic research, legacy PROMs have proven to be useful in measuring different outcome variables, most notably the physical function of surgical patients.[1, 2] However, many legacy PROMs are only validated for particular orthopedic conditions or specific patient populations. Furthermore, there are several different legacy PROMs used to assess the same anatomic location, which limits the ability to compare and analyze studies that report PROMs.[3-5] These differing legacy PROMs pull key questions from different item banks, making data aggregation across orthopedic literature difficult and cumbersome. Although it can be beneficial to compare data from the same legacy instrument implemented in a similar manner, administering specific legacy PROMs to specific patients becomes time consuming and difficult to manage, both for the clinician and the patient in a busy orthopedic clinic setting. These questionnaires can contain several questions, leading to disruption of clinic workflow, survey fatigue in patients and, subsequently, low completion rates.[6] Finally, as many of these PROMs

are diagnosis-specific, they must be manually assigned to patients upon determination of a diagnosis. Therefore, they do not lend well to automation on a population health perspective. This is why it is important to implement a standardized method of administering and interpreting PROMs, both from a clinical and epidemiological standpoint.

Due to these challenges with legacy PROMs, there has been increased interest in National Institutes of Health (NIH) Patient-Reported Outcomes Measurement Information System (PROMIS) assessment instruments. These questionnaires are domain-specific instead of disease-specific. For example, PROMIS Physical Function (PROMIS PF) measures assess general physical function regardless of diagnosis, whereas legacy PROMs often focus on particular patient cohorts and diagnoses. Additionally, PROMIS Upper Extremity Physical Function-CAT (PROMIS PF-UE) was developed to evaluate physical function specifically in upper extremity patients. PROMIS PF-UE CAT 2.0 is the most updated iteration of PROMIS PF-UE used in clinical practice. With regards to physical function (PF), the PROMIS PF measures a patient's self-reported capability and reports a quantitative score, with 50 correlating to that of the reference population, and 10 points representing one standard deviation. PROMIS PF can be administered as a short-form (SF) or as a computer adaptive test (CAT). Computer adaptive testing uses item response theory to customize question delivery based on real-time patient answers, allowing for a higher level of precision while using fewer questions.[7]

This study will serve as an update to a previous systematic review comparing PROMIS with legacy PROMs in the field of orthopedic surgery.[8] Since this study by Fidai and colleagues was published in 2018, 19 additional UE studies and 20 additional LE studies comparing specific legacy scoring modalities with PROMIS PF forms have been published. With this substantial increase in relevant UE and LE studies, we aim to expand on answers to the following questions:

How well do PROMIS PF forms correlate with legacy PROMs in UE and LE orthopedic patients? Which PROM can be administered more efficiently? How do the floor and ceiling effects compare? Is PROMIS PF as reliable, if not more reliable, as legacy PROMs?

Therefore, the purpose of this systematic review is to compare PROMIS PF with traditional legacy PROMs in regard to correlations, ease of use, and quality criteria for UE and LE orthopedic conditions. We hypothesized that PROMIS PF would correlate strongly with legacy PROMs and remain highly reliable while having less question response burden.

Methods

One of the authors (A.Z.) performed a systematic electronic search under the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines using PubMed (MEDLINE). The initial search was conducted on July 25, 2020. This search was later reproduced by two independent authors (V.A., A.S.). The following search terms were used in the title, abstract, and keyword fields: ("Patient-Reported Outcomes Measurement Information System" OR "PROMIS") AND ("physical function" OR "pf"). The additional criterion of "physical function" was added because the PROMIS PF domain is the most thoroughly studied health domain in patients with musculoskeletal disorders.[7] Inclusion criteria were any orthopedic-related article Level I through Level IV published in or before July 2020. We excluded nonorthopedic articles, non-English articles, unpublished studies, studies with Level V evidence, letters to the editor, editorials, basic science articles, and conference proceeding abstracts. Non-English studies were excluded in order to avoid inconsistencies that may have been reported during the process of translating English PROMs used in non-English populations.

The search identified 857 potentially eligible studies. A study was deemed eligible if it was related to orthopedic surgery and the title or abstract included "Patient-Reported Outcomes Measurement Information System" or "PROMIS" along with "physical function" or "pf." One author (A.Z.) applied the inclusion and exclusion criteria to the studies. Subsequently, two independent authors (V.A., A.S.) duplicated and validated the initial screenings. After review of the title, abstract, and full text, 85 were marked for inclusion. A study was marked if it compared a PROMIS PF form with a legacy PROM completed among either UE, LE, or Spine patients. Of the 85 marked studies, three articles were excluded during data extraction based on the lack of appropriate metrics compared in this study. **Figure 1** shows the algorithm used. Two authors (A.Z., V.A.) sorted the studies based on the abstracts from the electronic search. The papers were categorized into one of three anatomic regions determined by the title and abstract: upper extremity, lower extremity, and spine. The full text of 26 UE and 28 LE articles were included in the final analysis. The authors were not blinded to the authors of the study, title, and journal of publication. Consensus regarding study inclusion was reached between the authors and the principal investigator. The data were extracted from the included papers by two authors (A.Z., V.A.) using a database created in Microsoft Excel (Microsoft, Seattle, WA, USA). The database was validated by two authors (V.A., A.S.) and again by the first author (A.Z.) before analysis.

The data extracted from the selected studies were author, year, title, journal, study design, PubMed ID, country, number of patients, and demographics of study participants including age, sex, level of education, study start date, end date, multicenter versus single center, studies compared, number of questions per study, floor and ceiling effects of each study, time to completion, and Spearman/Pearson correlations. In addition, item reliability, person reliability,

and Cronbach reliability data were extracted from 5 of the 26 UE studies and 6 of the 28 LE studies.

Quality Criteria for PROMs

The quality of PROMs is assessed on criterion of validity, reliability, and ability to detect change.[9] The validity (e.g., construct validity) of a questionnaire is determined through comparison with other established PROMs by correlation analysis.[10] Pearson and Spearman coefficients are the measures of correlation assessed in this study. Pearson correlations are more commonly used when both variables being assessed are normally distributed, while Spearman correlation is more commonly used when at least 1 variable is skewed or continuous. A generally accepted consensus for strength of correlation is stratified as r = 0.4 to 0.59 for moderate, r = 0.6 to 0.79 as strong, and r = 0.8 to 1.0 as very strong.[11]

Reliability is quantified by several different parameters. Cronbach's reliability measures internal consistency, where high Cronbach's alpha scores reflect higher levels of precision, but also redundancy within test items. Person reliability is a measure used in Rasch analysis, which determines whether a person's response to an item reflects a response the model would have predicted for that person.[12] Item reliability reflects how much the PROM question contributes to the total score variance, where higher values represent better reliability.

Floor and ceiling effects are another quality criterion for PROMs compared and assessed in this study. Minimal floor and ceiling effects (generally <15%) imply high levels of content validity, responsiveness, and applicability.[10] Therefore, a PROM with low floor and ceiling effects is useful in identifying differences in patients within the low and high score ranges.[13]

Statistical Analysis

Statistical analysis was conducted using Statistical Analysis System version 9.4 (SAS Institute Inc., Cary, NC, USA). Mixed linear models were used to assess the differences between the PROMIS and legacy test results for the outcomes of test completion time, average number of questions, Cronbach reliability, person reliability, item reliability and floor and ceiling effects. This method takes into account the possibility of a study having results from multiple PROMIS and legacy tests. In the models, test type (PROMIS vs legacy) was considered as the fixed effect and study was considered as the random effect. The mean and standard error for each outcome by test type were computed using these models. We report weighted averages for correlation coefficients generated between PROMIS PF measures and legacy PROMs using the study sample size as the weight. All correlation coefficients were converted to an absolute value, thus allowing us to only compare the magnitude (and not direction) of the correlation strength. Individual studies included in this review reported either Spearman or Pearson correlations to draw comparisons between PROMIS PF and legacy PROMs completed by UE and LE orthopedic patients. We generated separate weighted mean values depending on whether the correlation coefficient was generated using Spearman or Pearson procedures. A P value of $< .05$ was deemed statistically significant.

The combined sample size of all articles yielded 6,074 UE and 9,366 LE patients. In total, 7 different PROMIS PF forms (PF CAT, PF-UE CAT, PF short form 8a, PF-UE short form 8a, Global-10 PF, Mobility CAT, Peds PF-UE CAT) were compared with 39 unique legacy PROMs relevant to UE, while 7 different PROMIS PF forms (short forms 4a, 8a, and 10a, PF CAT, lower extremity CAT, mobility CAT, cancer CAT) were compared with 51 unique legacy PROMs relevant to LE.

Results

In total, we found 26 studies that evaluated a total of 6,074 patients with UE orthopedic diagnoses and 28 studies that evaluated a total of 9,366 patients with LE orthopedic diagnoses. Details regarding the UE and LE studies encompassed in this review are shown in **Table 1** and **Table 2**, respectively. PROMIS PF scores were correlated with legacy scores in all 26 UE studies, and 27 of the 28 LE studies. Of the 26 UE studies, 14 of them reported Pearson's correlation coefficients and 12 reported Spearman's correlation coefficients. Of the 27 LE studies, 11 Pearson and 16 Spearman correlations were reported. The correlations between PROMIS PF forms and legacy PROMs utilized in UE and LE studies are illustrated in **Figure 2**.. The Pearson and Spearman correlations between UE legacy scores and PROMIS PF scores are depicted in **Table 2** and **Table 3**, respectively. The Pearson and Spearman correlations between LE legacy scores and PROMIS PF scores are depicted in **Table 4** and **Table 5**, respectively. Overall, the weighted Pearson correlation was found to be 0.624 (standard error [SE] = 0.042) and the weighted Spearman correlation was 0.566 (SE = 0.042) among UE patients. The weighted Pearson correlation was found to be 0.645 (standard error [SE] = 0.062) and the weighted Spearman correlation was 0.631 (SE = 0.041) among LE patients. Weighted Pearson and Spearman correlations between PROMIS PF forms and highly represented UE and LE PF legacy PROMs are shown in **Table 6** and **Table 7**, respectively. Results comparing PROMIS forms to legacy PROMs completed by UE and LE orthopedic patients are shown in **Table 8** and **Table 9**, respectively.

This study also sought to compare the average time to completion and the average number of questions between lower extremity PROMIS PF measures and legacy PROMs. Eight

of the 26 UE studies analyzed time to completion among 1,301 UE patients who completed both a PROMIS PF form and a legacy PROM. Eleven of the 28 LE studies analyzed time to completion among 2,600 lower extremity patients who completed both a PROMIS PF form and a legacy PROM. **Figure 3** represents the comparison of time to completion between PROMIS PF and legacy PROMs completed by UE and LE patients. A mixed linear model was used since we did not have paired data for completion time. Not all studies reported an average completion time for both PROMIS and legacy PROMs. The difference in time to completion was found to be borderline significant (P = 0.084) for PROMIS PF forms (90.5 seconds, SE = 55.2) compared to legacy PROMs (223.8 seconds, SE = 37.0) completed by UE patients. Time to completion was found to be significantly less (P = 0.0063) for PROMIS PF forms (63.6 seconds, SE = 33.1) than legacy PROMs (203.2 seconds, SE = 23.4) completed by LE patients.

Sixteen of the 26 studies analyzed the average number of questions completed among 2,878 upper extremity patients who completed a PROMIS PF form. Eighteen of the 26 studies analyzed the average number of questions completed among 4,288 upper extremity patients who completed a legacy PF form. **Figure 4** represents the comparison of average number of questions between PROMIS PF and legacy PROMs completed by upper extremity patients. There were significantly fewer questions (P = 0.0093) in the PROMIS PF forms (5.9, SE = 3.3) than the legacy PROMs (17.7, SE = 2.1).

Twenty of the 28 studies analyzed the average number of questions completed among 4,256 LE patients who completed a PROMIS PF form. Twenty-four of the 28 studies analyzed the average number of questions completed among 8,099 lower extremity patients who completed a legacy PF form. **Figure** 4 represents the comparison of average number of questions between PROMIS PF and legacy PROMs completed by lower extremity patients. There were

significantly fewer questions ($P < 0.001$) in the PROMIS PF forms (4.81, SE = 2.27) than the legacy PROMs (15.33, SE = 1.28).

Mixed linear models were used in analyzing differences in Cronbach, person, and item reliability. This was done because not all UE and LE studies reported paired data for the reliability outcomes. No significant difference was found in the Cronbach, person, and item reliabilities between PROMIS PF forms and legacy PROMs completed by UE and LE patients. These findings are illustrated in **Figure** 5.

The floor and ceiling effects of legacy PROMs were compared with PROMIS PF forms. Fourteen UE studies, involving 3,116 UE patients reported floor and ceiling effects for PROMIS PF forms. Twelve studies reported floor and ceiling effects for legacy PROMs completed by 2,209 UE patients. **Figure 6 A** and **B** demonstrates the results comparing floor and ceiling effects between PROMIS PF forms and legacy PROMs completed by UE patients. Overall, PROMIS PF forms had less floor and ceiling effects (0.43%, SE = 0.79% and 2.29%, SE = 1.59%, respectively) than legacy PROMs (1.46%, SE = 0.55% and 4.39%, SE = 1.10%, respectively). The difference in floor and ceiling effects was not significantly different between PROMIS forms and legacy PROMs ($P = 0.305$ and $P = 0.299$, respectively) completed by UE patients.

Nineteen LE studies, involving 3,589 LE patients reported floor and ceiling effects for PROMIS PF forms. The same 19 studies reported floor and ceiling effects for legacy PROMs completed by 3,456 LE patients. **Figure 7 A** and **B** demonstrates the results comparing floor and ceiling effects between PROMIS PF forms and legacy PROMs completed by LE patients. Overall, PROMIS PF forms had less floor and ceiling effects (0.578%, SE = 1.302% and 0.821%, SE = 1.420%, respectively) than legacy PROMs (3.624%, SE = 0.827% and 5.710%,

SE = 0.901%, respectively) completed by LE patients. The difference in ceiling effects was found to be significantly different (P = 0.0094), while the difference in floor effects was nearly significant (P = 0.063).

Discussion

We demonstrate that PROMIS PF scores correlate strongly with legacy PROMs completed by both UE and LE orthopedic patients. Moreover, we found PROMIS PF to be quicker to administer and applicable to a broad patient population while remaining highly reliable. We base this conclusion on the results of our analysis comparing various quality and questionnaire criterion, which are used when interpreting any PROMs.

Typically, relative efficiency comparisons or responsiveness measured by an effect size comparison between a legacy PROM and PROMIS PF form would support the argument that these assessments equate on a clinical and epidemiological level. Several of the upper and lower extremity studies encompassed in this review measure the strength of the relationship between PROMIS and legacy questionnaires by incorporating effect size comparisons. Among the lower extremity studies included in this review, 7 reported effect size comparisons.[14-20] Likewise, 4 of the upper extremity studies reported effect size comparisons.[18, 21, 22]. Hung et al. demonstrated large effect sizes when evaluating PROMIS PF, HOOS JR, and KOOS JR instrument responsiveness in joint function, which supports the use of PROMIS PF in effectively assessing treatment change among arthroplasty patients.[23] The results of this review uncover the ease of use and versatile nature of PROMIS PF forms administered to upper and lower extremity patients, however, more research is needed to thoroughly verify and validate that PROMIS PF

forms equate to legacy PROMs on responsiveness effect size as well as relative efficiency comparisons among various, specific orthopedic conditions.

These results demonstrate PROMIS PF may have certain advantages over validated legacy PROMs with regards to decreased survey administration time and question burden. When compared to commonly used UE and LE legacy PROMs, early findings suggest that PROMIS PF can be used as a practical, standardized PROM applied to a variety of UE and LE patients, thereby minimizing variability in high-impact literature.[3, 24] However, additional research is needed to solidify PROMIS PF as an effective means of standardizing PROM measurements across literature. Additionally, despite administrative burden and variability seen among UE and LE legacy PROMs, they do provide useful anatomic, as well as condition-specific outcome measurements of PF.[1, 2] In some orthopedic cases, questions may arise that require a more focused intervention that PROMIS is unable to provide. Therefore, legacy PROMs aid in evaluating an UE or LE patient from a specific clinical standpoint, as opposed to a standardized, albeit validated measurement of PF produced using PROMIS PF forms.

There are a variety of legacy PROMs used to evaluate different UE and LE orthopedic conditions. However, heterogeneity exists among the different PROMIS PF forms utilized in the included studies. In total, 7 different PROMIS PF forms (PF CAT, PF-UE CAT, PF short form 8a, PF-UE short form 8a, Global-10 PF, Mobility CAT, Peds PF-UE CAT) were compared with 39 unique legacy PROMs relevant to UE, while 7 different PROMIS PF forms (short forms 4a, 8a, and 10a, PF CAT, lower extremity CAT, mobility CAT, cancer CAT) were compared with 51 unique legacy PROMs relevant to LE. Despite the heterogeneity among PROMIS PF forms, scores from different PROMIS PF measures are easily comparable and interpretable due to a common item bank utilized by each form.[25] The different versions of PROMIS PF forms are

heavily explained and validated in the literature.[25-28] Heterogeneity can introduce inconsistencies among different PROMs, but overall PROMIS PF forms performed consistently in measuring and reporting a patient's physical function.

Reducing administrative burden is a major parameter for successful implementation of any PROM. Decreasing the number of questions within a PROM helps to minimize user fatigue and simplify their scoring profile. Since many of these PROMs are distributed during clinic visits, reducing the time to completion of these questionnaires benefits both the patient and the clinician. Among studies that evaluated UE and LE orthopedic patients, we found that PROMIS PF forms have significantly fewer questions than legacy PROMs. Regarding time to completion, the PROMIS PF forms were completed in significantly less time than legacy PROMs. Our findings demonstrate PROMIS PF forms can reduce the administrative burden traditionally seen with legacy PROMs.

The reliability and floor and ceiling effects of the PROM questionnaire itself are important parameters when evaluating and comparing PROMs. A high reliability becomes important when administering PROMs because it indicates that the reproducibility of a subsequent test will not be altered by a patient's background characteristics.[19] No significant difference was found between PROMIS PF and legacy PROMs when comparing Cronbach, person, and item reliabilities in UE and LE orthopedic studies. This demonstrates that PROMIS PF is as reliable as the validated legacy PROMs that have been used traditionally in orthopedic practice.

Floor and ceiling effects are defined as the proportion of respondents scoring the highest (ceiling) or lowest (floor) possible score, therefore measuring the sensitivity and coverage of a questionnaire at each end of the scale.[13] For example, if a large proportion of patients receive the

lowest possible score on a questionnaire, then that indicates that all of those patients have the same level of health, which in turn indicates the inability of that instrument to differentiate among those at the low end of the spectrum. Furthermore, if a patient continues to improve over time and an instrument has a noticeable floor or ceiling effect, the instrument will not assess the change over time. Significant floor and ceiling effects have historically been set at 15%,[10] while other studies have stated that <10% or even 5% is an acceptable benchmark.[29, 30] Among UE and LE orthopedic patients, PROMIS PF forms had less floor and ceiling effects than legacy PROMs. Our findings show that PROMIS PF forms are able to adequately capture UE and LE patients at both ends of the scoring spectrum, while also being able to assess progress over time. This implies that PROMIS PF forms are better able than their legacy PROM counterparts to differentiate patients who are severely affected by the same orthopedic condition from those who are not as affected. This is likely due to the ability to administer PROMIS PF forms as computer adaptive tests. Subsequent questions within PROMIS PF CAT forms are customized based on the responses to previous questions, creating a dynamic and efficient scoring algorithm. Furthermore, this automation prevents patients from contradicting themselves while responding to the PROMIS PF form, thereby limiting response bias and collecting the most accurate responses. The ease of response of a CAT form minimizes the amount of time patients need to dedicate to completing the questions, which prevents survey fatigue and optimizes the opportunity to capture accurate responses, as opposed to recording rushed responses from a lengthy, redundant legacy questionnaire. With accurate question responses, PROMIS PF forms are able to accurately differentiate between patients whose physical function is either severely affected or not affected by a particular orthopedic condition, which is represented by floor and ceiling responses, respectively. This allows for more accurate measurements of physical function

and limits the opportunity for patients to score exceedingly low or high compared to other patients with the same condition.

These positive findings indicate that PROMIS PF forms should be used in practice by clinicians and researchers looking to assess the physical function of UE and LE orthopedic patients. Developing a standardized method of administering PROMs will aid in creating a valid and consistent interpretation of UE and LE physical function across orthopedic literature and within the orthopedic clinic setting. Furthermore, PROMIS PF can be administered as both short-forms and computer adaptive tests, which makes it an ideal platform for utilizing technology and virtual care to report outcomes among UE and LE orthopedic patients. PROMIS PF can also help to reduce the burdens of administering and collecting legacy PROMs in the busy clinic setting.

Limitations

Certain limitations are present within this systematic review. The search was limited to UE and LE studies indexed in PubMed/MEDLINE. Albeit a limitation, since the majority of systematic reviews include one or two additional databases, the isolated use of PubMed has been shown to be sufficient for publishing systematic reviews/meta-analyses within high-impact literature.[31] Therefore, this search method provides a comprehensive literature search, while remaining both efficient and reproducible. Only English papers were included in this review. Data extraction was not blinded as reviewers were able to view the authors, title, and journal of all articles reviewed in the course of this systematic review.

Conclusions

PROMIS PF scores correlate strongly with commonly used legacy PROMs in orthopedics, particularly in UE and LE patients. PROMIS PF forms can be administered efficiently and to a broad patient population while remaining highly reliable. Therefore, they can be justified for standardized use among orthopedic patients with UE and LE conditions, improving the ability to aggregate and compare outcomes in orthopedic research.

References

1. Hunt KJ, Hurwit D. Use of patient-reported outcome measures in foot and ankle research. *J Bone Joint Surg Am.* 2013;95:e118(111-119).
2. Cutler HS, Guzman JZ, Al Maaieh M, Connolly J, Skovrlj B, Cho SK. Patient Reported Outcomes in Adult Spinal Deformity Surgery: A Bibliometric Analysis. *Spine Deform.* 2015;3:312-317.
3. Makhni EC, Padaki AS, Petridis PD, et al. High Variability in Outcome Reporting Patterns in High-Impact ACL Literature. *J Bone Joint Surg Am.* 2015;97:1529-1542.
4. Makhni EC, Steinhaus ME, Morrow ZS, et al. Outcomes assessment in rotator cuff pathology: what are we measuring? *J Shoulder Elbow Surg.* 2015;24:2008-2015.
5. Makhni EC, Meyer MA, Saltzman BM, Cole BJ. Comprehensiveness of Outcome Reporting in Studies of Articular Cartilage Defects of the Knee. *Arthroscopy.* 2016;32:2133-2139.
6. Lizzio VA, Gulledge CM, Meta F, Franovic S, Makhni EC. Using a Web-Based Data Collection Platform to Implement an Effective Electronic Patient-Reported Outcome Registry. *Arthrosc Tech.* 2019;8:e535-e539.
7. Brodke DJ, Saltzman CL, Brodke DS. PROMIS for Orthopaedic Outcomes Measurement. *J Am Acad Orthop Surg.* 2016;24:744-749.
8. Fidai MS, Saltzman BM, Meta F, et al. Patient-Reported Outcomes Measurement Information System and Legacy Patient-Reported Outcome Measures in the Field of Orthopaedics: A Systematic Review. *Arthroscopy.* 2018;34:605-614.
9. Hawkins RJ. Recommendations for Evaluating and Selecting Appropriately Valued Outcome Measures. *Instr Course Lect.* 2016;65:587-591.
10. Terwee CB, Bot SD, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol.* 2007;60:34-42.
11. Chan YH. Biostatistics 104: correlational analysis. *Singapore Med J.* 2003;44:614-619.
12. Beckmann JT, Hung M, Bounsanga J, Wylie JD, Granger EK, Tashjian RZ. 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.
13. Gulledge CM, Smith DG, Ziedas A, Muh SJ, Moutzouros V, Makhni EC. Floor and Ceiling Effects, Time to Completion, and Question Burden of PROMIS CAT Domains Among Shoulder and Knee Patients Undergoing Nonoperative and Operative Treatment. *JB JS Open Access.* 2019;4.
14. Koltsov JCB, Greenfield ST, Soukup D, Do HT, Ellis SJ. Validation of Patient-Reported Outcomes Measurement Information System Computerized Adaptive Tests Against the Foot and Ankle Outcome Score for 6 Common Foot and Ankle Pathologies. *Foot Ankle Int.* 2017;38:870-878.
15. Nwachukwu BU, Beletsky A, Naveen N, et al. Patient-Reported Outcomes Measurement Information System (PROMIS) Instruments Correlate Better With Legacy Measures in Knee Cartilage Patients at Postoperative Than at Preoperative Assessment. *Arthroscopy.* 2020;36:1419-1428.

16. Janssen SJ, Paulino Pereira NR, Raskin KA, et al. A comparison of questionnaires for assessing physical function in patients with lower extremity bone metastases. *J Surg Oncol.* 2016;114:691-696.

17. Kenney RJ, Houck J, Giordano BD, Baumhauer JF, Herbert M, Maloney MD. Do Patient Reported Outcome Measurement Information System (PROMIS) Scales Demonstrate Responsiveness as Well as Disease-Specific Scales in Patients Undergoing Knee Arthroscopy? *Am J Sports Med.* 2019;47:1396-1403.

18. Ploetze KL, Dalton JF, Calfee RP, McDonald DJ, O'Keefe RJ, Cipriano CA. Patient-Reported Outcomes Measurement Information System physical function correlates with Toronto Extremity Salvage Score in an orthopaedic oncology population. *J Orthop Translat.* 2019;19:143-150.

19. Hung M, Franklin JD, Hon SD, Cheng C, Conrad J, Saltzman CL. Time for a paradigm shift with computerized adaptive testing of general physical function outcomes measurements. *Foot Ankle Int.* 2014;35:1-7.

20. Rothrock NE, Kaat AJ, Vrahas MS, et al. Validation of PROMIS Physical Function Instruments in Patients With an Orthopaedic Trauma to a Lower Extremity. *J Orthop Trauma.* 2019;33:377-383.

21. Beletsky A, Nwachukwu BU, Manderle BJ, et al. The Impact of Workers' Compensation on Patient-Reported Outcomes Measurement Information System Upper Extremity and Legacy Outcome Measures in Patients Undergoing Arthroscopic Rotator Cuff Repair. *Arthroscopy.* 2019;35:2817-2824.

22. Bernstein DN, Houck JR, Mahmood B, Hammert WC. Responsiveness of the PROMIS and its Concurrent Validity with Other Region- and Condition-specific PROMs in Patients Undergoing Carpal Tunnel Release. *Clin Orthop Relat Res.* 2019;477:2544-2551.

23. Hung M, Saltzman CL, Greene T, et al. Evaluating instrument responsiveness in joint function: The HOOS JR, the KOOS JR, and the PROMIS PF CAT. *J Orthop Res.* 2018;36:1178-1184.

24. Steinhaus ME, Makhni EC, Lieber AC, et al. Variable reporting of functional outcomes and return to play in superior labrum anterior and posterior tear. *J Shoulder Elbow Surg.* 2016;25:1896-1905.

25. Rose M, Bjorner JB, Gandek B, Bruce B, Fries JF, Ware JE, Jr. The PROMIS Physical Function item bank was calibrated to a standardized metric and shown to improve measurement efficiency. *J Clin Epidemiol.* 2014;67:516-526.

26. Schalet BD, Hays RD, Jensen SE, Beaumont JL, Fries JF, Cella D. Validity of PROMIS physical function measured in diverse clinical samples. *J Clin Epidemiol.* 2016;73:112-118.

27. Rose M, Bjorner JB, Becker J, Fries JF, Ware JE. Evaluation of a preliminary physical function item bank supported the expected advantages of the Patient-Reported Outcomes Measurement Information System (PROMIS). *J Clin Epidemiol.* 2008;61:17-33.

28. Kaat AJ, Buckenmaier CT, 3rd, Cook KF, et al. The expansion and validation of a new upper extremity item bank for the Patient-Reported Outcomes Measurement Information System® (PROMIS). *J Patient Rep Outcomes.* 2019;3:69.

29. Hung M, Voss MW, Bounsanga J, Gu Y, Granger EK, Tashjian RZ. Psychometrics of the Patient-Reported Outcomes Measurement Information System Physical Function

instrument administered by computerized adaptive testing and the Disabilities of Arm, Shoulder and Hand in the orthopedic elbow patient population. *J Shoulder Elbow Surg.* 2018;27:515-522.

30. Hung M, Hon SD, Franklin JD, et al. Psychometric properties of the PROMIS physical function item bank in patients with spinal disorders. *Spine (Phila Pa 1976).* 2014;39:158-163.

31. Saltzman BM, Leroux T, Meyer MA, et al. The Therapeutic Effect of Intra-articular Normal Saline Injections for Knee Osteoarthritis: A Meta-analysis of Evidence Level 1 Studies. *Am J Sports Med.* 2017;45:2647-2653.

32. Beckmann JT, Hung M, Voss MW, Crum AB, Bounsanga J, Tyser AR. Evaluation of the Patient-Reported Outcomes Measurement Information System Upper Extremity Computer Adaptive Test. *J Hand Surg Am.* 2016;41:739-744.e734.

33. Morgan JH, Kallen MA, Okike K, Lee OC, Vrahas MS. PROMIS Physical Function Computer Adaptive Test Compared With Other Upper Extremity Outcome Measures in the Evaluation of Proximal Humerus Fractures in Patients Older Than 60 Years. *J Orthop Trauma.* 2015;29:257-263.

34. Tyser AR, Hung M, Bounsanga J, Voss MW, Kazmers NH. Evaluation of Version 2.0 of the PROMIS Upper Extremity Computer Adaptive Test in Nonshoulder Upper Extremity Patients. *J Hand Surg Am.* 2019;44:267-273.

35. Jayakumar P, Overbeek CL, Ring DC. Relationship of age on enjoyment of physical activity in upper extremity illness. *Hand (N Y).* 2015;10:767-772.

36. Overbeek CL, Nota SP, Jayakumar P, Hageman MG, Ring D. The PROMIS physical function correlates with the QuickDASH in patients with upper extremity illness. *Clin Orthop Relat Res.* 2015;473:311-317.

37. Döring AC, Nota SP, Hageman MG, Ring DC. Measurement of upper extremity disability using the Patient-Reported Outcomes Measurement Information System. *J Hand Surg Am.* 2014;39:1160-1165.

38. Patterson BM, Orvets ND, Aleem AW, et al. Correlation of Patient-Reported Outcomes Measurement Information System (PROMIS) scores with legacy patient-reported outcome scores in patients undergoing rotator cuff repair. *J Shoulder Elbow Surg.* 2018;27:S17-s23.

39. Minoughan CE, Schumaier AP, Fritch JL, Grawe BM. Correlation of Patient-Reported Outcome Measurement Information System Physical Function Upper Extremity Computer Adaptive Testing, With the American Shoulder and Elbow Surgeons Shoulder Assessment Form and Simple Shoulder Test in Patients With Shoulder Pain. *Arthroscopy.* 2018;34:1430-1436.

40. Gausden EB, Levack AE, Sin DN, et al. Validating the Patient Reported Outcomes Measurement Information System (PROMIS) computerized adaptive tests for upper extremity fracture care. *J Shoulder Elbow Surg.* 2018;27:1191-1197.

41. Waljee JF, Carlozzi N, Franzblau LE, Zhong L, Chung KC. Applying the Patient-Reported Outcomes Measurement Information System to Assess Upper Extremity Function among Children with Congenital Hand Differences. *Plast Reconstr Surg.* 2015;136:200e-207e.

42. Anthony CA, Glass N, Hancock K, Bollier M, Hettrich CM, Wolf BR. Preoperative Performance of the Patient-Reported Outcomes Measurement Information System in Patients With Rotator Cuff Pathology. *Arthroscopy.* 2017;33:1770-1774.e1771.

43. Robins RJ, Anderson MB, Zhang Y, Presson AP, Burks RT, Greis PE. Convergent Validity of the Patient-Reported Outcomes Measurement Information System's Physical Function Computerized Adaptive Test for the Knee and Shoulder Injury Sports Medicine Patient Population. *Arthroscopy.* 2017;33:608-616.

44. Fu MC, Chang B, Wong AC, et al. PROMIS physical function underperforms psychometrically relative to American Shoulder and Elbow Surgeons score in patients undergoing anatomic total shoulder arthroplasty. *J Shoulder Elbow Surg.* 2019;28:1809-1815.

45. Beleckas CM, Gerull W, Wright M, Guattery J, Calfee RP. Variability of PROMIS Scores Across Hand Conditions. *J Hand Surg Am.* 2019;44:186-191.e181.

46. Saad MA, Kassam HF, Suriani RJ, Jr., Pan SD, Blaine TA, Kovacevic D. Performance of PROMIS Global-10 compared with legacy instruments in patients with shoulder arthritis. *J Shoulder Elbow Surg.* 2018;27:2249-2256.

47. Stoop N, Menendez ME, Mellema JJ, Ring D. The PROMIS Global Health Questionnaire Correlates With the QuickDASH in Patients With Upper Extremity Illness. *Hand (N Y).* 2018;13:118-121.

48. Kaat AJ, Rothrock NE, Vrahas MS, et al. Longitudinal Validation of the PROMIS Physical Function Item Bank in Upper Extremity Trauma. *J Orthop Trauma.* 2017;31:e321-e326.

49. Lu Y, Agarwalla A, Patel BH, et al. Relationship between the Patient-Reported Outcomes Measurement Information System (PROMIS) computer adaptive testing and legacy instruments in patients undergoing isolated biceps tenodesis. *J Shoulder Elbow Surg.* 2020;29:1214-1222.

50. Tyser AR, Beckmann J, Franklin JD, et al. Evaluation of the PROMIS physical function computer adaptive test in the upper extremity. *J Hand Surg Am.* 2014;39:2047-2051.e2044.

51. Nicholson AD, Kassam HF, Pan SD, Berman JE, Blaine TA, Kovacevic D. Performance of PROMIS Global-10 Compared With Legacy Instruments for Rotator Cuff Disease. *Am J Sports Med.* 2019;47:181-188.

52. Anthony CA, Glass NA, Hancock K, Bollier M, Wolf BR, Hettrich CM. Performance of PROMIS Instruments in Patients With Shoulder Instability. *Am J Sports Med.* 2017;45:449-453.

53. Dowdle SB, Glass N, Anthony CA, Hettrich CM. Use of PROMIS for Patients Undergoing Primary Total Shoulder Arthroplasty. *Orthop J Sports Med.* 2017;5:2325967117726044.

54. Shamrock AG, Wolf BR, Ortiz SF, et al. Preoperative Validation of the Patient-Reported Outcomes Measurement Information System in Patients With Articular Cartilage Defects of the Knee. *Arthroscopy.* 2020;36:516-520.

55. Papuga MO, Beck CA, Kates SL, Schwarz EM, Maloney MD. Validation of GAITRite and PROMIS as high-throughput physical function outcome measures following ACL reconstruction. *J Orthop Res.* 2014;32:793-801.

56. Padilla JA, Rudy HL, Gabor JA, et al. Relationship Between the Patient-Reported Outcome Measurement Information System and Traditional Patient-Reported Outcomes for Osteoarthritis. *J Arthroplasty.* 2019;34:265-272.

57. Nwachukwu BU, Beck EC, Chapman R, Chahla J, Okoroha K, Nho SJ. Preoperative Performance of the PROMIS in Patients Undergoing Hip Arthroscopic Surgery for Femoroacetabular Impingement Syndrome. *Orthop J Sports Med.* 2019;7:2325967119860079.

58. Moore MLG, Kortlever JTP, Keulen MHF, Brigati DP, Bozic KJ, Koenig KM. PROMIS PF correlates with HOOS, JR in patients with hip pain. *J Orthop.* 2020;21:58-61.

59. Miles M, Nadarajah V, Jauregui JJ, et al. Evaluation of the PROMIS Physical Function Computer Adaptive Test in Patients Undergoing Knee Surgery. *J Knee Surg.* 2020;33:810-817.

60. Kortlever JTP, Leyton-Mange A, Keulen MHF, et al. PROMIS Physical Function Correlates with KOOS, JR in Patients with Knee Pain. *J Knee Surg.* 2019.

61. Kollmorgen RC, Hutyra CA, Green C, Lewis B, Olson SA, Mather RC, 3rd. Relationship Between PROMIS Computer Adaptive Tests and Legacy Hip Measures Among Patients Presenting to a Tertiary Care Hip Preservation Center. *Am J Sports Med.* 2019;47:876-884.

62. Kohring JM, Pelt CE, Anderson MB, Peters CL, Gililland JM. Press Ganey Outpatient Medical Practice Survey Scores Do Not Correlate With Patient-Reported Outcomes After Primary Joint Arthroplasty. *J Arthroplasty.* 2018;33:2417-2422.

63. Hung M, Stuart AR, Higgins TF, Saltzman CL, Kubiak EN. Computerized Adaptive Testing Using the PROMIS Physical Function Item Bank Reduces Test Burden With Less Ceiling Effects Compared With the Short Musculoskeletal Function Assessment in Orthopaedic Trauma Patients. *J Orthop Trauma.* 2014;28:439-443.

64. Hung M, Nickisch F, Beals TC, Greene T, Clegg DO, Saltzman CL. New paradigm for patient-reported outcomes assessment in foot & ankle research: computerized adaptive testing. *Foot Ankle Int.* 2012;33:621-626.

65. Hung M, Baumhauer JF, Brodsky JW, et al. Psychometric Comparison of the PROMIS Physical Function CAT With the FAAM and FFI for Measuring Patient-Reported Outcomes. *Foot Ankle Int.* 2014;35:592-599.

66. Hoch JM, Lorete C, Legner J, Hoch MC. The Relationship Among 3 Generic Patient-Reported Outcome Instruments in Patients With Lower Extremity Health Conditions. *J Athl Train.* 2019;54:550-555.

67. Hancock KJ, Glass N, Anthony CA, et al. PROMIS: a valid and efficient outcomes instrument for patients with ACL tears. *Knee Surg Sports Traumatol Arthrosc.* 2019;27:100-104.

68. Hancock KJ, Glass N, Anthony CA, et al. Performance of PROMIS for Healthy Patients Undergoing Meniscal Surgery. *J Bone Joint Surg Am.* 2017;99:954-958.

69. Hajewski CJ, Baron JE, Glass NA, et al. Performance of the Patient-Reported Outcome Measurement Information System in Patients With Patellofemoral Instability. *Orthop J Sports Med.* 2020;8:2325967120915540.

70. Hafner BJ, Gaunaurd IA, Morgan SJ, Amtmann D, Salem R, Gailey RS. Construct Validity of the Prosthetic Limb Users Survey of Mobility (PLUS-M) in Adults With Lower Limb Amputation. *Arch Phys Med Rehabil.* 2017;98:277-285.

71. Gulbrandsen TR, Khazi ZM, Bollier M, et al. Preoperative Performance of Patient-Reported Outcomes Measurement Information System in Patients with Meniscal Root Tears. *J Knee Surg.* 2019.

72. Gausden EB, Levack A, Nwachukwu BU, Sin D, Wellman DS, Lorich DG. Computerized Adaptive Testing for Patient Reported Outcomes in Ankle Fracture Surgery. *Foot Ankle Int.* 2018;39:1192-1198.

73. 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.

Figure Legends

Figure 1. Preferred Reporting Items for Systematic Reviews and Meta-Analyses flow diagram

Figure 2. Results showing the weighted average strength of the correlation between Patient Reported Outcomes Measurement Information System and legacy for both Pearson and Spearman correlation procedures among upper and lower extremity patients. The means were weighted by the sample size used in the respective study. Each bar represents the mean ±2 standard errors. N = number of studies.

Figure 3. Results comparing patient time to completion between Patient-Reported Outcomes Measurement Information System and legacy outcomes for upper and lower extremity using a mixed linear model. Each bar represents the mean ±2 standard errors. N = number of studies. (*) represents $P < .05$

Figure 4. Average difference in number of questions between Patient-Reported Outcomes Measurement Information System and legacy outcomes for upper and lower extremity using a mixed linear model. Each bar represents the mean ±2 standard errors. N = number of studies. (*) represents $P < .05$. (**) represents $P < .001$

Figure 5. Average Cronbach reliability, person reliability, and item reliability between Patient-Reported Outcomes Measurement Information System (PROMIS) and legacy tests completed by (A) upper and (B) lower extremity patients. Mixed linear models were used to assess differences between PROMIS and legacy values. Each bar represents the mean ±2 standard errors. N = number of studies.

Figure 6. Results of (A) floor and (B) ceiling effects between Patient-Reported Outcomes Measurement Information System and legacy tests completed by upper extremity patients. Each bar represents the mean ±2 standard errors. N = number of studies.

Figure 7. Results of (A) floor and (B) ceiling effects between Patient-Reported Outcomes Measurement Information System and legacy tests completed by lower extremity patients. Each bar represents the mean ±2 standard errors. N = number of studies. (*) represents $P < .05$.

Table 1. Inclusion and Exclusion Criteria of Upper Extremity Studies Encompassed in This Review

Study	Level of Evidence	Sample Size (n)	Mean Age (yr)	Sex (n) M/F	Inclusion Criteria	Exclusion Criteria
Beckmann et al.[32] (2016)	III	379	47	171/208	>18 yr old; hand clinic patients	<18 yr old; shoulder injury
Beckmann et al.[12] (2015)	III	187	60	106/81	>18 yr old; clinical diagnosis of rotator cuff disease	<18 yr old; no rotator cuff disorder; unable to score surveys; data collected during visits other than initial patient encounter
Morgan et al.[33] (2015)	III	47	68	18/29	>60 yr old; displaced proximal humerus fracture	<60 yr old; nondisplaced fractures; isolated greater/lesser tuberosity fractures; pathologic fractures; neurologic injury; presented >4 wk after initial injury; failed previous treatment; systematic conditions affecting healing
Tyser et al.[34] (2014)	III	134	42	75/60	>18 yr old; clinic patient with upper extremity problem other than shoulder	<18 yr old; shoulder injury
Jayajumar et al.[35] (2015)	IV	98	50	47/51	>18 yr old; hand clinic patients	<18 yr old; pregnant women; unable to communicate in English
Overbeek et al.[36] (2014)	III	93	50	41/52	>18 yr old; hand clinic patients	<18 yr old; mental health conditions; unable to communicate in English
Doring et al.[37] (2014)	I	84	49	39/45	>18 yr old; orthopaedic clinic patients	<18 yr old; pregnant women; unable to give consent
Ploetze et al.[18] (2019)	III	97	53	Not explicitly stated	Patients with surgical treatment for primary and metastatic bone or soft tissue tumor while consecutively being evaluated in the outpatient setting	Patients whose preoperative evaluation took place only in the emergency department or inpatient setting, pregnant women, and non-English speaking individuals
Patterson et al.[38] (2018)	III	164	58	85/79	>18 yr old; patients with arthroscopic rotator cuff repair, both traumatic and atraumatic tears	<18 yr old; patients with revision rotator cuff repair or incomplete responses to 1 or more surveys such that the survey could not be scored
Minoughan et al.[39] (2018)	II	90	50.3	49/41	All patients presenting with a primary complaint of shoulder pain	Patients with a history of prior rotator cuff surgery, prior tendon transfer, current partial thickness tear, or history of shoulder surgery in the previous 6 months
Gausden et al.[40] (2018)	III	174	53	72/102	Patients with open reduction and internal fixation of a distal radius, elbow, humeral shaft, proximal humeral, or clavicular fracture	Not explicitly stated
Beletsky et al.[21] (2019)	III	122	53.6	71/51	Patients who completed all relevant legacy scores preoperatively and receipt of a primary rotator cuff repair for a full-thickness tear	Patients who did not complete any legacy PROM preoperatively, had revision rotator cuff repair, partial rotator cuff tears, or a receipt of significant concomitant procedures
Waljee et al.[41] (2014)	III	33	11.4	19/13	6-17 yr olds; patients with congenital hand difference, able to read and speak English, and had no previous history of cognitive impairment	Children were excluded if they were unable to complete data collection methods
Anthony et al.[42] (2017)	II	82	54	50/32	All patients with preoperative diagnosis of rotator cuff pathologies	Not explicitly stated
Robins et al.[43] (2017)	III	415	42.5	266/149	Patients with knee and shoulder related injuries who completed the PF-CAT score on an electronic tablet and either had completed all of the requested PRO scores or had started the process before opting out of completing the scores	Patients with repeat encounters or not having a completed PF-CAT score
Fu et al.[44] (2019)	III	179	65.6	103/76	Patients with osteoarthritis, having complete preoperative scores for ASES and PF-CAT, and undergoing primary anatomic total shoulder arthroplasty	Not explicitly stated
Bernstein et al.[22] (2019)	II	70	61	24/46	Patients with carpal tunnel syndrome who presented preoperatively and at 6 weeks or 3 months postoperatively	Not explicitly stated
Beleckas et al.[45] (2019)	III	1471	57.2	544/927	Hand clinic patients	Patients with missing valid PROMIS scores
Saad et al.[46] (2018)	III	161	64.5	85/76	>18 yr old; orthopedic clinic patients and able to speak English	<18 yr old
Stoop et al.[47] (2018)	III	112	50	54/58	>18 yr old	<18 yr old; patients who were pregnant
Kaat et al.[48] (2017)	III	424	47.3	180/228	>18 yr old; patients with isolated upper extremity fracture(s) treated with or without surgery in past 12 months, and able to speak English	<18 yr old; patients with pathological fractures, not community ambulatory before injury, and the presence of an additional lower extremity fracture
Lu et al.[49] (2020)	III	175	51.6	120/55	Patients with rotator cuff debridement for partial rotator cuff tears or impingement and have completed preoperative PROMs	Patients with full-thickness rotator cuff tears, receiving concurrent rotator cuff repair or shoulder arthroplasty, or a history of ipsilateral BT
Tyser et al.[50] (2014)	III	825	50.3	357/468	>18 yr old; orthopedic clinic patients	<18 yr old
Nicholson et al.[51] (2019)	II	323	57.7	174/149	>18 yr old; patients with shoulder impingement or a partial or full-thickness tear of the posterosuperior rotator cuff, with all clinical diagnoses confirmed by magnetic resonance imaging findings, and able to speak English	<18 yr old; patients with previous rotator cuff surgery
Anthony et al.[52] (2016)	II	70	27	52/18	Patients with shoulder instability	Patients with incomplete data
Dowdle et al.[53] (2017)	III	53	60.8	31/22	Patients with shoulder osteoarthritis	Patients with incomplete PROs

ASES, American Shoulder and Elbow Surgeons; BT, Biceps Tenodesis; CAT, Computer Adaptive Test; PF, Physical Function; PRO, Patient Reported Outcomes; PROM, Patient Reported Outcome Measures; PROMIS, Patient-Reported Outcomes Measurement Information System; Wk, week; Yr, Year

719 **Table 2.** Inclusion and Exclusion Criteria of Lower Extremity Studies Encompassed in This Review

Study	Level of Evidence	Sample Size (n)	Mean Age (yr)	Sex (n) M/F	Inclusion Criteria	Exclusion Criteria
Shamrock et al.[54] (2020)	III	275	34	150/125	>10 yr old; scheduled for osteochondral autograft or allograft transfer, microfracture, chondroplasty, autologous chondrocyte implantation, and allograft cartilage resurfacing	< 10 yr old; subjects undergoing concomitant ACL reconstruction, collateral ligament repair/reconstruction, meniscal repair, meniscal transplant, high tibial osteotomy, distal femoral osteotomy, or bilateral procedures under the same general anesthetic
Rothrock et al.[20] (2019)	III	402	45	225/177	>18 yr old; injury isolated to a lower extremity fracture, treatment with surgery, or nonsurgical closed treatment in the past 1–12 months	< 18 yr old; pathologic fractures, presence of an additional upper extremity fracture, and inability to ambulate independently in the community before the injury
Robins et al.[43] (2017)	III	450	36	267/183	Knee clinic patients	Repeat patient encounter
Ploetze et al.[18] (2019)	III	70	53	NA	Underwent surgical treatment for primary and metastatic bone or soft tissue tumor; consecutively evaluated in the outpatient setting by the orthopaedic oncology service	If preoperative evaluation took place only in the emergency department or inpatient setting; pregnant women
Papuga et al.[55] (2014)	II	106	30	49/57	>13 yr old; scheduled to undergo BTB autograft ACL reconstruction surgery	< 13 yr old; history of balance disorders, degenerative gait disorders, neuromuscular disorders, dementia, depression, or cognitive impairment
Padilla et al.[56] (2019)	III	3644	63	1403/2241	>18 yr old; hip/knee clinic patients	< 18 yr old; underwent revision TJA, bilateral TJA, or a partial joint arthroplasty
Nwachukwu et al.[15] (2020)	IV	250	38	143/107	Patients that underwent a primary cartilage procedure (i.e., debridement, allograft transplantation, microfracture, or autologous chondrocyte implantation [ACI]) for a focal cartilage defect	Concomitant ligamentous, meniscal, or bony procedures; biological augmentation; total knee arthroplasty within the follow-up period
Nwachukwu et al.[57] (2019)	II	197	33	46/151	Clinical and radiographic findings of symptomatic FAIS; surgical treatment with hip arthroscopic surgery for FAIS	History of bilateral hip surgery (including ipsilateral revision); hip arthroscopic surgery for an indication other than FAIS; signs of osteoarthritis (Tönnis grade >1); hip dysplasia (lateral center-edge angle <20 degrees); history of congenital hip disorders (slipped capital femoral epiphysis, developmental hip dysplasia, etc.); concomitant procedures during the time of surgery
Moore et al.[58] (2020)	III	213	56	73/140	>18 yr old; Hip clinic patients	< 18 yr old; cognitively impaired; traumatic hip injury etiology; prior hip replacement in the affected joint
Miles et al.[59] (2019)	III	412	39	232/180	>12 yr old; patients undergoing orthopaedic surgery	< 12 yr old
Kortlever et al.[60] (2019)	III	88	56	42/46	>18 yr old; knee clinic patients	< 18 yr old; concomitant pain in one or both hips
Koltsov et al.[14] (2017)	II	240	53	65/175	Scheduled to undergo 1 of 6 common foot and ankle surgeries: HV, AI, AA, HR, FF, osteochondral defects of the talus	Not explicitly stated
Kollmorgen et al.[61] (2019)	III	125	38	23/102	>18 yr old, < 80 yr old; Hip clinic patients	< 18 yr old, >80 yr old; repeat patient encounter
Kohring et al.[62] (2018)	II	540	64	238/302	Hip/knee arthroplasty patients	Multiple joint arthroplasty procedures
Kenney et al.[17] (2019)	II	76	49	40/36	>18 yr old; underwent primary knee arthroscopy with partial medial meniscectomy, partial lateral meniscectomy, chondroplasty, loose body removal, and/or synovectomy	< 18 yr old; revision surgery; concomitant ligamentous injury
Janssen et al.[16] (2016)	III	100	63	41/59	>18 yr old; metastatic bone lesion, myeloma, or lymphoma of lower extremity	< 18 yr old
Hung et al.[63] (2014)	III	153	49	83*/66*	>18 yr old; postoperative lower extremity trauma patients, or nonoperative lower extremity fracture care	< 18 yr old
Hung et al.[64] (2012)	III	287	47	115/172	Foot/ankle clinic patients	Not explicitly stated
Hung et al.[19] (2014)	I	126	44	49*/64*	>18 yr old; foot/ankle clinic patients	< 18 yr old
Hung et al.[65] (2014)	I	311	50	80*/178*	Scheduled to undergo 1 of 6 common foot and ankle surgeries: HV, AI, AA, HR, FF, HT	Not explicitly stated
Hoch et al.[66] (2019)	III	100	20	NA	> 18 yr old, < 35 yr old; intercollegiate or recreational athletes with lower extremity health condition	< 18 yr old, >35 yr old
Hancock et al.[67] (2019)	III	100	26	55/45	Patients indicated for operative management of an ACL tear	Simultaneous operations including microfracture, meniscus repair, osteotomy, or osteochondral allograft
Hancock et al.[68] (2017)	III	107	38	71/36	Scheduled to undergo primary meniscal repair, meniscectomy, or debridement	Simultaneous operation such as osteotomy, application of osteochondral allograft, or had grade-4 osteoarthritis
Hajewski et al.[69] (2020)	II	91	20	33/58	Patients indicated for operative management of patellofemoral instability	Not explicitly stated
Hafner et al.[70] (2017)	III	199	55	142/57	>18 yr old; amputation between the hip and ankle; amputation due to trauma, dysvascular complications, tumor, or infection; use of prosthesis to ambulate for at least 4 mo, English literacy	Excluded if they had another amputation (arm, contralateral leg, etc.)
Gulbrandsen et al.[71] (2019)	III	51	39	20/31	≥14 yr old; indicated for surgical repair of MRT	< 14 yr old
Gausden et al.[72] (2018)	III	132	47	79/53	Patients who underwent osteosynthesis for an unstable ankle fracture	Not explicitly stated
Driban et al.[73] (2015)	I	204	60	61/143	>40 yr old; >40 on at least 1 question in the WOMAC pain subscale; fulfill ACR OA criteria; radiographic evidence of OA; knee pain/disability on examination	Patients with experience in Tai Chi/yoga; serious medical conditions; intra-articular steroid injection or TKA in last 3 mo, MMSE score <24

720 AA, ankle arthritis; ACL, anterior cruciate ligament; ACR, American College of Rheumatology; AI, anterior instability; BTB, bone-
721 tendon-bone; FF, flat foot; HR, hallus rigidus; HT, hammertoe; HV, hallux valgus; MMSE, mini-mental status examination; OA, osteoarthritis;
722 TKA, total knee arthroplasty; WOMAC, Western Ontario and McMaster University Osteoarthritis Index.
723 *Sample size inconsistencies due to loss to follow-up and sample subsets.
724
725
726
727
728
729
730
731
732
733
734
735
736
737

738 **Table 3.** Pearson Correlations Between PROMIS and Upper Extremity Legacy Scores
739

Study	PROMIS	Legacy	Sample Size	Pearson	P Value

740

Beckmann et al.[32]	PROMIS PF CAT	DASH	379	-0.817	0.05
	PROMIS PF-UE CAT	DASH	379	-0.804	0.05
Beckmann et al.[12]	PROMIS PF CAT	ASES	187	0.581	0.001
	PROMIS PF CAT	SST	187	0.635	0.001
Tyser et al.[34]	PROMIS PF CAT	DASH	134	0.726	0.001
Jayajumar et al.[35]	PROMIS PF CAT	PACES	98	0.31	0.0018
Overbeek et al.[36]	PROMIS PF CAT	QuickDASH	93	-0.55	0.001
Doring et al.[37]	PROMIS PF CAT	QuickDASH	84	-0.57	0.001
	PROMIS PF-UE CAT	QuickDASH	84	-0.81	0.001
	PROMIS PF-Mobility CAT	QuickDASH	84	-0.43	0.001
Ploetze et al.[18]	PROMIS PF CAT	TESS	97	0.64	0.055
Patterson et al.[38]	PROMIS PF CAT	ASES	164	0.43	0.001
	PROMIS PF CAT	SST	164	0.51	0.001
	PROMIS PF-UE CAT v2.0	ASES	164	0.59	0.001
	PROMIS PF-UE CAT v2.0	SST	164	0.62	0.001
Minoughan et al.[39]	PROMIS PF-UE CAT v2.0	ASES	90	0.72	0.001
	PROMIS PF-UE CAT v2.0	SST	90	0.82	0.001
Anthony et al.[42]	PROMIS PF-UE CAT	ASES	82	0.77	0.01
	PROMIS PF-UE CAT	WORC	82	0.73	0.01
	PROMIS PF-UE CAT	Marx	82	0.23	0.04
	PROMIS PF-UE CAT	SF-36 PF	82	0.66	0.01
	PROMIS PF-UE CAT	SF-36 GH	82	0.3	0.01
	PROMIS PF-UE CAT	EQ-5D	82	0.73	0.01
	PROMIS PF CAT	ASES	82	0.55	0.01
	PROMIS PF CAT	WORC	82	0.61	0.01
	PROMIS PF CAT	Marx	82	0.34	0.01
	PROMIS PF CAT	SF-36 PF	82	0.77	0.01
	PROMIS PF CAT	SF-36 GH	82	0.5	0.01
	PROMIS PF CAT	EQ-5D	82	0.65	0.01
Fu et al.[44]	PROMIS PF CAT	ASES	179	0.487	0.001
Stoop et al.[47]	PROMIS Global-10 Physical Health scale	QuickDASH	112	-0.47	0.0001
Tyser et al.[50]	PROMIS PF-UE CAT v2.0	QuickDASH	825	-0.749	0.05
Dowdle et al.[53]	PROMIS PF-UE CAT	ASES	53	0.55	0.01
	PROMIS PF-UE CAT	SF-36 PF	53	0.53	0.01
	PROMIS PF-UE CAT	EQ-5D	53	0.48	0.01
	PROMIS PF-UE CAT	WOOS	53	0.34	0.01
	PROMIS PF-UE CAT	Marx	53	0.06	0.62
	PROMIS PF CAT	ASES	53	0.62	0.01
	PROMIS PF CAT	SF-36 PF	53	0.81	0.01
	PROMIS PF CAT	EQ-5D	53	0.64	0.01
	PROMIS PF CAT	WOOS	53	0.51	0.01
	PROMIS PF CAT	Marx	53	0.29	0.02

ASES, American Shoulder and Elbow Surgeons; CAT, Computer Adaptive Test; DASH, Disabilities of the Arm Shoulder and Hand; EQ-5D, EuroQol Five Dimensions Questionnaire; GH, General Health; Marx, Marx Activity Rating Scale; PACES, Physical Exercise During Adjuvant Chemotherapy Effectiveness Study; PF, Physical Function; PROMIS, Patient-Reported Outcomes Measurement Information System; QuickDASH, Quick Disabilities of Arm and Shoulder and Hand; SF-36, Short Form 36-Item Health Survey; SST, Simple Shoulder Test; TESS, Toronto Extremity Salvage Score; UE, Upper Extremity; v2.0, version 2.0; WOOS, Western Ontario Osteoarthritis of the Shoulder; WORC, Western Ontario Rotator Cuff

Table 4. Spearman Correlations Between PROMIS and Upper Extremity Legacy Scores

Study	PROMIS	Legacy	Sample Size	Spearman	P Value
Morgan et al.[33]	PROMIS PF CAT	Constant	47	0.52	0.001
	PROMIS PF CAT	DASH	47	-0.66	0.001
	PROMIS PF CAT	SMFA bothersome index	47	-0.68	0.001
	PROMIS PF CAT	SMFA functional index	47	-0.81	0.001
Gausden et al.[40]	PROMIS PF CAT	SF-36 PCS	174	0.71	0.001
	PROMIS PF CAT	SF-36 MCS	174	0.25	0.001
	PROMIS PF CAT	VAS Pain	174	-0.36	0.001
	PROMIS PF CAT	DASH	174	-0.76	0.001
	PROMIS PF CAT	Constant	174	0.5	0.002
	PROMIS PF CAT	UCLA	174	0.61	0.001
	PROMIS PF CAT	Mayo	174	0.36	0.02
Beletsky et al.[21]	PROMIS PF-UE CAT	QuickDASH	122	0.77	0.01
	PROMIS PF-UE CAT	ASES	122	0.61	0.01
	PROMIS PF-UE CAT	SANE	122	0.41	0.01
	PROMIS PF-UE CAT	SF-12 PCS	122	0.44	0.01
	PROMIS PF-UE CAT	VR12 PCS	122	0.52	0.01
	PROMIS PF-UE CAT	Constant	122	0.62	0.01
	PROMIS PF-UE CAT	VR6D	122	0.61	0.01
	PROMIS PF-UE CAT	BRS	122	0.25	0.01
	PROMIS PF-UE CAT	SF-12 MCS	122	0.42	0.01
	PROMIS PF-UE CAT	VR12 MCS	122	0.5	0.01
Waljee et al.[41]	PROMIS Ped PF-UE CAT	MHQ Overall	33	0.46	0.05
	PROMIS Ped PF-UE CAT	MHQ Satisfaction	33	0.41	0.05
	PROMIS Ped PF-UE CAT	MHQ Hand Appearance	33	0.17	
	PROMIS Ped PF-UE CAT	MHQ ADL	33	0.49	0.001
	PROMIS Ped PF-UE CAT	MHQ Pain	33	-0.46	0.05
	PROMIS Ped PF-UE CAT	DASH	33	-0.87	0.001
	PROMIS Ped PF-UE CAT	PODCI UE	33	0.89	0.001
	PROMIS Ped PF-UE CAT	PODCI Mobility	33	0.63	0.05
	PROMIS Ped PF-UE CAT	PODCI Sports	33	0.76	0.001
	PROMIS Ped PF-UE CAT	PODCI Pain	33	0.71	0.001
	PROMIS Ped PF-UE CAT	PODCI Happiness	33	0.4	
	PROMIS Ped PF-UE CAT	PODCI Global	33	0.8	0.001
	PROMIS PF-UE 8a	MHQ Overall	33	0.5	0.001
	PROMIS PF-UE 8a	MHQ Satisfaction	33	0.47	0.05
	PROMIS PF-UE 8a	MHQ Hand Appearance	33	0.32	0.05
	PROMIS PF-UE 8a	MHQ ADL	33	0.47	0.05
	PROMIS PF-UE 8a	MHQ Pain	33	-0.43	0.05
	PROMIS PF-UE 8a	DASH	33	-0.84	0.001
	PROMIS PF-UE 8a	PODCI UE	33	0.085	0.001
	PROMIS PF-UE 8a	PODCI Mobility	33	0.62	0.001
	PROMIS PF-UE 8a	PODCI Sports	33	0.77	0.001
	PROMIS PF-UE 8a	PODCI Pain	33	0.7	0.001
	PROMIS PF-UE 8a	PODCI Happiness	33	0.43	0.001
	PROMIS PF-UE 8a	PODCI Global	33	0.79	0.001
Robins et al.[43]	PROMIS PF CAT	SANE	415	0.5	0.0001
	PROMIS PF CAT	SST	415	0.64	0.0001
	PROMIS PF CAT	ASES	415	0.63	0.0001
Bernstein et al.[22]	PROMIS PF-UE CAT	MHQ Overall	70	0.57	0.001
	PROMIS PF-UE CAT	MHQ ADL	70	0.61	0.001
	PROMIS PF-UE CAT	MHQ ADL 2 Hands	70	0.69	0.001

	PROMIS PF-UE CAT	MHQ ADL Overall	70	0.69	0.001
	PROMIS PF-UE CAT	MHQ Work	70	0.64	0.001
	PROMIS PF-UE CAT	MHQ Pain	70	0.54	0.001
	PROMIS PF-UE CAT	MHQ Aesthetics	70	0.007	0.96
	PROMIS PF-UE CAT	MHQ Satisfaction	70	0.53	0.001
	PROMIS PF-UE CAT	MHQ Total	70	0.65	0.001
	PROMIS PF-UE CAT	BCTQ Symptom	70	0.74	0.001
	PROMIS PF-UE CAT	BCTQ Function	70	0.75	0.001
	PROMIS PF CAT	MHQ Overall	70	0.4	0.001
	PROMIS PF CAT	MHQ ADL	70	0.37	0.002
	PROMIS PF CAT	MHQ ADL 2 Hands	70	0.45	0.001
	PROMIS PF CAT	MHQ ADL Overall	70	0.44	0.001
	PROMIS PF CAT	MHQ Work	70	0.35	0.003
	PROMIS PF CAT	MHQ Pain	70	0.17	0.17
	PROMIS PF CAT	MHQ Aesthetics	70	0.05	0.67
	PROMIS PF CAT	MHQ Satisfaction	70	0.38	0.001
	PROMIS PF CAT	MHQ Total	70	0.36	0.002
	PROMIS PF CAT	BCTQ Symptom	70	0.4	0.001
	PROMIS PF CAT	BCTQ Function	70	0.37	0.002
Beleckas et al.[45]	PROMIS PF CAT	QuickDASH	1471	-0.66	
Saad et al.[46]	PROMIS Global-10 PF	ASES	161	0.57	0.0001
	PROMIS Global-10 PF	SANE	161	0.23	0.0045
	PROMIS Global-10 PF	WOOS	161	0.11	0.3743
	PROMIS Global-10 PF	EQ-5D	161	0.72	0.0001
Kaat et al.[48]	PROMIS PF-UE CAT	SF-36 PCS	424	0.59	0.001
	PROMIS PF-UE CAT	SMFA	424	-0.76	0.001
	PROMIS PF-UE CAT	QuickDASH	424	-0.82	0.001
	PROMIS PF 8a	SF-36 PCS	424	0.76	0.001
	PROMIS PF 8a	SMFA	424	-0.8	0.001
	PROMIS PF 8a	QuickDASH	424	-0.79	0.001
Lu et al.[49]	PROMIS PF-UE CAT	ASES	175	0.57	0.05
	PROMIS PF-UE CAT	SANE	175	0.42	0.05
	PROMIS PF-UE CAT	SF-12 PCS	175	0.35	0.05
	PROMIS PF-UE CAT	VR12 PCS	175	0.59	0.05
	PROMIS PF-UE CAT	VR6D	175	0.49	0.05
	PROMIS PF-UE CAT	VR12 MCS	175	0.3	0.05
	PROMIS PF-UE CAT	SF-12 MCS	175	0.27	0.05
	PROMIS PF-UE CAT	VAS Pain	175	0.35	0.05
	PROMIS PF-UE CAT	VAS Strength	175	0.5	0.05
	PROMIS PF-UE CAT	VAS Function	175	0.5	0.05
Nicholson et al.[51]	PROMIS Global-10 Physical Health	EQ-5D	323	0.7	0.0001
	PROMIS Global-10 Physical Health	ASES	323	0.62	0.0001
	PROMIS Global-10 Physical Health	WORC	323	0.47	0.0001
	PROMIS Global-10 Physical Health	SANE	323	0.41	0.0005
Anthony et al.[52]	PROMIS PF-UE CAT	ASES	70	0.71	0.01
	PROMIS PF-UE CAT	WOSI	70	0.63	0.01
	PROMIS PF-UE CAT	Marx	70	0.06	0.65
	PROMIS PF-UE CAT	SF-36 PF	70	0.78	0.01
	PROMIS PF-UE CAT	EQ-5D	70	0.66	0.01
	PROMIS PF CAT	ASES	70	0.67	0.01
	PROMIS PF CAT	WOSI	70	0.49	0.01
	PROMIS PF CAT	Marx	70	0.18	0.14
	PROMIS PF CAT	SF-36 PF	70	0.72	0.01
	PROMIS PF CAT	EQ-5D	70	0.59	0.01

ADL, Activities of Daily Living; ASES, American Shoulder and Elbow Surgeons; BCTQ, Boston Carpal Tunnel Questionnaire; BRS, Behavior Rating Scale; CAT, Computer Adaptive Test; Constant, Constant-Murley Shoulder Outcome Score; DASH, Disabilities of the Arm Shoulder and Hand; EQ-5D, EuroQol Five Dimensions Questionnaire; Marx, Marx Activity Rating Scale; Mayo, Mayo Elbow Performance Scale; MCS, Mental Component Score; MHQ, Michigan Hand Outcomes Questionnaire; PCS, Physical Component Score; PF, Physical Function; PODCI, Pediatrics Outcomes Data Collection Instrument; PROMIS, Patient-Reported Outcomes Measurement Information System; QuickDASH, Quick Disabilities of the Arm Shoulder and Hand; SANE, Single Assessment Numeric Evaluation; SF-12, Short Form 12-item Health Survey; SF-36, Short Form 36-Item Health Survey; SMFA, Short Musculoskeletal Function Assessment; SST, Simple Shoulder Test; UCLA, University California Los Angeles Activity Score; UE, Upper Extremity; VAS, Visual Analog Scale; VR12, Veterans Rand-12; VR6D, Veterans Rand 6 Dimension; WOOS, Western Ontario Osteoarthritis of the Shoulder; WORC, Western Ontario Rotator Cuff; WOSI, Western Ontario Shoulder Instability Index

Table 5. Pearson Correlations Between PROMIS and Lower Extremity Legacy Scores

Study	PROMIS	Legacy	Sample Size	Pearson	P Value
Rothrock et al.[20]	PROMIS PF 8a	PF-10	402	0.85	0.01
	PROMIS PF 8a	SMFA	402	-0.83	0.01

	PROMIS PF 8a	FAAM-ADL	402	0.87	0.01
	PROMIS PF 8a	FAAM Sport	402	0.82	0.01
	PROMIS PF 8a	UCLA	402	0.7	0.01
	PROMIS Mobility CAT	PF-10	402	0.8	0.01
	PROMIS Mobility CAT	SMFA	402	-0.77	0.01
	PROMIS Mobility CAT	FAAM-ADL	402	0.84	0.01
	PROMIS Mobility CAT	FAAM Sport	402	0.77	0.01
	PROMIS Mobility CAT	UCLA	402	0.61	0.01
Ploetze et al.[18]	PROMIS PF CAT	TESS	70	0.84	0.001
Papuga et al.[55]	PROMIS PF CAT	IKDC	106	0.8954	0.0001
Padilla et al.[56]	PROMIS PF CAT	KOOS-JR	3644	0.52	0.01
	PROMIS PF CAT	HOOS-JR	3644	0.59	0.01
Nwachukwu et al.[57]	PROMIS PF CAT	HOS ADL	197	0.801	0.001
	PROMIS PF CAT	HOS SS	197	0.675	0.001
	PROMIS PF CAT	mHHS	197	0.721	0.001
	PROMIS PF CAT	iHOT-12	197	0.722	0.001
	PROMIS PF CAT	VR-12 PCS	197	0.618	0.001
	PROMIS PF CAT	VR-12 MCS	197	0.721	0.001
	PROMIS PF CAT	VAS Pain	197	-0.365	0.001
Moore et al.[58]	PROMIS PF CAT	KOOS-JR	213	0.58	0.001
Kortlever et al.[60]	PROMIS PF CAT	KOOS-JR	88	0.74	0.001
Kenney et al.[17]	PROMIS PF CAT	IKDC	76	0.76	0.01
Hung et al.[64]	PROMIS PF-LE CAT	spFAAM	287	0.61	0.01
	PROMIS PF-LE CAT	FFI	287	0.6	0.01
Hung et al.[19]	PROMIS PF CAT	SF-36 PF	126	0.51	0.01
	PROMIS PF-LE CAT	SF-36 PF	126	0.7	0.01
Hung et al.[65]	PROMIS PF CAT	FAAM_ADL	311	0.792	
	PROMIS PF CAT	FFI-5pt	311	0.685	

CAT, computer adaptive test; FAAM-ADL, foot and ankle ability measure activities of daily living; FFI, foot function index; FFI-5pt, Foot Function Index-5pt Scale; HOS ADL, Hip Outcome Score Activities of Daily Living; HOS SS, Hip Outcome Score Sports Specific; HOOS-JR, Hip Injury and Osteoarthritis Outcome Score-Joint Replacement; iHOT-12, International Hip Outcome Tool-12; IKDC, International Knee Documentation Committee; KOOS-JR, Knee Injury and Osteoarthritis Outcome Score-Joint Replacement; LE, lower extremity; LEFS, Lower Extremity Functional Scale; mHHS, modified Harris Hip Score; MCS, Mental Component Summary; MTSS, Medial Tibial Stress Syndrome Score; PCS, Physical Component Summary; PF-10, physical function; PLUS-M, Prosthetic Limb Users Survey of Mobility; PS, Physical Function Short Form; QOL, quality of life; SF-36, Short Form 36-item Health Survey (RAND Corporation); SMFA, Short Musculoskeletal Function Assessment; spFAAM, sport module of Foot and Ankle Ability Measure; TESS, Toronto Extremity Salvage Score; UCLA, Single-item University of California, Los Angeles Activity Scale; VAS, visual analog scale; VR-12, Veterans RAND 12-item health survey; WOMAC, Western Ontario and McMaster University Osteoarthritis Index.

Table 6. Spearman Correlations Between PROMIS and Lower Extremity Legacy Scores

Study	PROMIS	Legacy	Sample Size	Spearman	P Value
Shamrock et al.[54]	PROMIS PF CAT	KOOS Symptoms	275	0.519	0.001

	PROMIS PF CAT	KOOS ADL	275	0.733	0.001
	PROMIS PF CAT	KOOS Sport	275	0.709	0.001
	PROMIS PF CAT	KOOS Pain	275	0.662	0.001
	PROMIS PF CAT	KOOS QOL	275	0.64	0.001
	PROMIS PF CAT	SF-36 PF	275	0.819	0.001
	PROMIS PF CAT	SF-36 PCS	275	0.766	0.001
	PROMIS PF CAT	EQ-5D	275	0.752	0.001
Robins et al.[43]	PROMIS PF CAT	SANE	450	0.6	0.0001
	PROMIS PF CAT	IKDC	317	0.75	0.0001
Nwachukwu et al.[15]	PROMIS PF CAT	IKDC	250	0.9	0.05
	PROMIS PF CAT	KOOS ADL	250	0.88	0.05
	PROMIS PF CAT	KOOS-JR	250	0.92	0.05
	PROMIS PF CAT	KOOS Pain	250	0.88	0.05
	PROMIS PF CAT	KOOS PS	250	-0.82	0.05
	PROMIS PF CAT	KOOS Sport	250	0.82	0.05
	PROMIS PF CAT	KOOS Symptoms	250	0.8	0.05
	PROMIS PF CAT	Marx	250	0.44	0.05
	PROMIS PF CAT	WOMAC Function	250	0.88	0.05
	PROMIS PF CAT	WOMAC Pain	250	0.93	0.05
	PROMIS PF CAT	WOMAC Stiffness	250	0.8	0.05
	PROMIS PF CAT	WOMAC Total	250	0.9	0.05
	PROMIS PF CAT	SF-12 PCS	250	0.86	0.05
	PROMIS PF CAT	VR-12 PCS	250	0.87	0.05
	PROMIS PF CAT	VR6D	250	0.83	0.05
	PROMIS PF CAT	VR-12 MCS	250	0.47	0.05
	PROMIS PF CAT	SF-12 MCS	250	0.27	0.05
	PROMIS PF CAT	BRS	250	0.13	0.05
Miles et al.[59]	PROMIS PF CAT	IKDC	412	0.71	0.001
Koltsov et al.[14]	PROMIS PF CAT	FAOS Activities	240	0.68	0.05
	PROMIS PF CAT	FAOS Sports	240	0.65	0.05
	PROMIS PF CAT	FAOS QOL	240	0.71	0.05
	PROMIS PF CAT	FAOS Pain	240	0.6	0.05
	PROMIS PF CAT	FAOS Symptoms	240	0.53	0.05
	PROMIS PF CAT	SF-12 PCS	240	0.29	0.05
	PROMIS PF CAT	SF-12 MCS	240	-0.02	0.05
Kollmorgen et al.[61]	PROMIS PF CAT	mHHS	125	0.71	
	PROMIS PF CAT	HOS-SS	125	0.81	
	PROMIS PF CAT	HOS-ADL	125	0.87	
	PROMIS PF CAT	iHOT-12	125	0.76	

	PROMIS PF CAT	VR6D	125	0.71	
Kohring et al.[62]	PROMIS PF CAT	PG	540	0.02	0.01
Janssen et al.[16]	PROMIS PF Cancer CAT	TESS	100	0.848	0.001
	PROMIS PF Cancer CAT	LEFS	100	0.867	0.001
	PROMIS PF Cancer CAT	MTSS	100	0.819	0.001
	PROMIS Neuro QOL-Mobility CAT	TESS	100	0.847	0.001
	PROMIS Neuro QOL-Mobility CAT	LEFS	100	0.843	0.001
	PROMIS Neuro QOL-Mobility CAT	MTSS	100	0.766	0.001
Hoch et al.[66]	PROMIS PF CAT	mDPA Total	100	-0.7	0.001
	PROMIS PF CAT	mDPA PSC	100	-0.7	0.001
	PROMIS PF CAT	mDPA MSC	100	-0.4	0.001
	PROMIS PF CAT	SF-12 PCS	100	0.65	0.001
	PROMIS PF CAT	SF-12 MCS	100	0.2	0.05
Hancock et al.[67]	PROMIS PF CAT	SF-36 PF	100	0.82	0.01
	PROMIS PF CAT	SF-36 GH	100	0.12	0.12
	PROMIS PF CAT	SF-36 Pain	100	0.51	0.01
	PROMIS PF CAT	KOOS Sport	100	0.7	0.01
	PROMIS PF CAT	KOOS ADL	100	0.74	0.01
	PROMIS PF CAT	KOOS Symptoms	100	0.54	0.01
	PROMIS PF CAT	KOOS Pain	100	0.58	0.01
	PROMIS PF CAT	KOOS QOL	100	0.49	0.01
	PROMIS PF CAT	Marx	100	0.08	0.46
	PROMIS PF CAT	EQ-5D	100	0.7	0.01
Hancock et al.[68]	PROMIS PF CAT	KOOS ADL	107	0.6	0.01
	PROMIS PF CAT	KOOS Sport	107	0.76	0.01
	PROMIS PF CAT	KOOS Symptoms	107	0.57	0.01
	PROMIS PF CAT	KOOS Pain	107	0.6	0.01
	PROMIS PF CAT	KOOS QOL	107	0.63	0.01
	PROMIS PF CAT	Marx	107	0.05	0.59
	PROMIS PF CAT	SF-36 PF	107	0.82	0.01
	PROMIS PF CAT	SF-36 GH	107	0.27	0.01
	PROMIS PF CAT	SF-36 Pain	107	0.6	0.01
	PROMIS PF CAT	EQ-5D	107	0.62	0.01
Hajewski et al.[69]	PROMIS PF CAT	SF-36 PF	91	0.78	0.01
	PROMIS PF CAT	SF-36 GH	91	0.12	0.26
	PROMIS PF CAT	KOOS ADL	91	0.68	0.01
	PROMIS PF CAT	KOOS Sport	91	0.58	0.01
	PROMIS PF CAT	KOOS Pain	91	0.62	0.01
	PROMIS PF CAT	KOOS QOL	91	0.53	0.01

	PROMIS PF CAT	KOOS Symptoms	91	0.47	0.01
	PROMIS PF CAT	AKPS	91	0.68	0.01
	PROMIS PF CAT	Marx	91	0.13	0.23
	PROMIS PF CAT	EQ-5D	91	0.6	0.01
Hafner et al.[70]	PROMIS PF 4a	PLUS-M	199	0.81	0.001
Gulbrandsen et al.[71]	PROMIS PF CAT	EQ-5D	51	0.72	0.0001
	PROMIS PF CAT	KOOS ADL	51	0.69	0.0001
	PROMIS PF CAT	KOOS Sport	51	0.54	0.0001
	PROMIS PF CAT	KOOS Pain	51	0.51	0.0001
	PROMIS PF CAT	KOOS Symptoms	51	0.48	0.0004
	PROMIS PF CAT	KOOS QOL	51	0.43	0.0015
	PROMIS PF CAT	WOMAC Function	51	0.69	0.0001
	PROMIS PF CAT	WOMAC Pain	51	0.53	0.0001
	PROMIS PF CAT	WOMAC Stiffness	51	0.5	0.0002
	PROMIS PF CAT	SF-36 PF	51	0.64	0.0001
	PROMIS PF CAT	Knee activity scale	51	0.18	0.208
Gausden et al.[72]	PROMIS PF CAT	FAOS Symptoms	132	0.46	0.001
	PROMIS PF CAT	FAOS Activities	132	0.63	0.001
	PROMIS PF CAT	FAOS QOL	132	0.61	0.001
	PROMIS PF CAT	FAOS Pain	132	0.56	0.001
	PROMIS PF CAT	FAOS Sports	132	0.62	0.001
	PROMIS PF CAT	OMAS	132	0.72	0.001
	PROMIS PF CAT	WAS	132	0.63	0.001
	PROMIS PF-LE CAT	FAOS Symptoms	132	0.5	0.001
	PROMIS PF-LE CAT	FAOS Activities	132	0.69	0.001
	PROMIS PF-LE CAT	FAOS QOL	132	0.63	0.001
	PROMIS PF-LE CAT	FAOS Pain	132	0.61	0.001
	PROMIS PF-LE CAT	FAOS Sports	132	0.65	0.001
	PROMIS PF-LE CAT	OMAS	132	0.73	0.001
	PROMIS PF-LE CAT	WAS	132	0.6	0.001
Driban et al.[73]	PROMIS PF 10a	SF-36 PF	204	0.79	0.05
	PROMIS PF 10a	WOMAC Function	204	-0.48	0.05

ADL, Activities of Daily Living; AKPS, Anterior Knee Pain Scale; BRS, Brief Resilience Scale; CAT, computer adaptive test; EQ-5D, EuroQol-5 Dimension questionnaire; FAOS, Foot and Ankle Outcome Score; FFI, foot function index; GH, General Health; HOS, Hip Outcome Score; HOOS, Hip Injury and Osteoarthritis Outcome Score; iHOT-12, International Hip Outcome Tool-12; IKDC, International Knee Documentation Committee; JR, Joint Replacement; KOOS, Knee Injury and Osteoarthritis Outcome Score; LE, lower extremity; LEFS, Lower Extremity Functional Scale; mHHS, modified Harris Hip Score; MCS, Mental Component Summary; mDPA, modified Disablement in the Physically Active Scale; MTSS, Medial Tibial Stress Syndrome Score; OMAS, Olerud and Molander Ankle Score; PCS, Physical Component Summary; PG, Press Ganey; PLUS-M, Prosthetic Limb Users Survey of Mobility; QOL, quality of life; SANE, Single Assessment Numerical Evaluation; SF-12, Short Form 12-item Health Survey (RAND Corporation); SF-36, Short Form 36-item Health Survey (RAND Corporation); SS, Sports Specific; TESS, Toronto Extremity Salvage Score; VR6D, Veterans RAND-6 Dimension; VR-12, Veterans RAND 12-item health survey; WAS, Weber Activity Scale; WOMAC, Western Ontario and McMaster University Osteoarthritis Index.

Table 7. Weighted Correlations of Highly Represented Upper Extremity Physical Function Legacy PROMs

PROMIS PF Form	Legacy PF Form	Number of Correlations Represented in the Literature	Weighted Pearson Correlation (SE)	Weighted Spearman Correlation (SE)	Correlation Strength
PROMIS PF-UE CAT	ASES	4	0.65 (0.05)		Strong
	Quick DASH	2	0.75 (0.03)		Strong
	SF-36 PF	2	0.60 (0.07)		Strong
	ASES	3		0.61 (0.04)	Strong
	Quick DASH	2		0.81 (0.03)	Very Strong
	Sane	2		0.42 (0.01)	Moderate
	SF-12 PCS	2		0.39 (0.05)	Weak
	VR-12 PCS	2		0.56 (0.04)	Moderate
PROMIS PF-UE CAT 2.0	ASES	2	0.64 (0.07)		Strong
	SST	2	0.69 (0.1)		Strong
PROMIS PF CAT	ASES	5	0.52 (0.03)		Moderate
	DASH	2	0.79 (0.05)		Strong
	EQ-5D	2	0.65 (0.01)		Strong
	Quick DASH	2	0.56 (0.01)		Moderate
	SF-36 PF	2	0.79 (0.02)		Strong
	ASES	2		0.64 (0.02)	Strong
	DASH	2		0.74 (0.05)	Strong
	Constant	2		0.50 (0.01)	Moderate

ASES, American Shoulder and Elbow Surgeons; CAT, Computer Adaptive Test; Constant, Constant-Murley Shoulder Outcome Score; DASH, Disabilities of the Arm Shoulder and Hand; EQ-5D, EuroQol Five Dimensions Questionnaire; PCS, Physical Component Score; PF, Physical Function; PROMIS, Patient-Reported Outcomes Measurement Information System; QuickDASH, Quick Disabilities of the Arm Shoulder and Hand; SANE, Single Assessment Numeric Evaluation; SF-12, Short Form 12-item Health Survey; SF-36, Short Form 36-Item Health Survey; SST, Simple Shoulder Test; UE, Upper Extremity; VR12, Veterans Rand-12

Table 8. Weighted Correlations of Highly Represented Lower Extremity Physical Function Legacy PROMs

PROMIS PF Form	Legacy PF Form	Number of Correlations Represented in the Literature	Weighted Pearson Correlation (SE)	Weighted Spearman Correlation (SE)	Correlation Strength

PROMIS PF CAT	KOOS-JR	3	0.53 (0.066)	Moderate
	IKDC	2	0.84 (0.068)	Very Strong
	KOOS ADL	6	0.75 (0.037)	Strong
	KOOS Sport	6	0.72 (0.043)	Strong
	SF-36 PF	5	0.80 (0.035)	Very Strong
	EQ-5D	5	0.70 (0.029)	Strong
	SF-12 PCS	3	0.59 (0.17)	Moderate
	IKDC	3	0.77 (0.058)	Strong
	FAOS Activities	2	0.66 (0.02)	Strong
	FAOS Sports	2	0.64 (0.015)	Strong
	WOMAC Function	2	0.85 (0.095)	Very Strong
	VR6D	2	0.79 (0.06)	Strong

ADL, Activities of Daily Living; EQ-5D, EuroQol-5 Dimension questionnaire; FAOS, Foot and Ankle Outcome Score; IKDC, International Knee Documentation Committee; JR, Joint Replacement; KOOS, Knee Injury and Osteoarthritis Outcome Score; PCS, Physical Component Summary; SF-12, Short Form 12-item Health Survey (RAND Corporation); SF-36, Short Form 36-item Health Survey (RAND Corporation); VR6D, Veterans RAND-6 Dimension; WOMAC, Western Ontario and McMaster University Osteoarthritis Index.

Table 9. Results comparing PROMIS forms to legacy PROMs completed by upper extremity orthopedic patients

	PROMIS				Legacy				
Outcomes	**Studies**	**Total participants**	**Mean**	**S.E.**	**Studies**	**Total participants**	**Mean**	**S.E.**	**p-value**
Completion time	8	1301	90.5	55.2	8	1301	223.8	37.0	0.084
Ave # of questions	16	2878	5.9	3.3	18	4288	17.7	2.1	0.0093
Ave Cronbach Reliability	3	1628	0.958	0.016	2	803	0.938	0.018	0.556
Ave Person Reliability	5	1599	0.914	0.080	4	774	0.673	0.080	0.122
Ave Item Reliability	5	1599	0.954	0.008	4	774	0.973	0.008	0.197
Ave Floor Effects (%)	14	3116	0.43	0.79	12	2209	1.46	0.55	0.305
Ave Ceiling Effects (%)	14	3116	2.29	1.59	12	2209	4.39	1.10	0.299

Table 10. Results comparing PROMIS forms to legacy PROMs completed by lower extremity orthopedic patients.

	PROMIS				Legacy				
Outcome	**Studies**	**Total participants**	**Mean**	**S.E.**	**Studies**	**Total participants**	**Mean**	**S.E.**	**p-value**
Completion time	11	2600	63.6	33.1	11	2600	203.2	23.4	0.0063

Ave # of questions	20	4256	4.81	2.27	24	8099	15.33	1.28	<0.001
Ave Cronbach Reliability	2	555	0.976	0.018	3	655	0.944	0.012	0.479
Ave Person Reliability	3	724	0.958	0.029	4	964	0.895	0.023	0.232
Ave Item Reliability	3	724	0.970	0.049	4	964	0.938	0.040	0.664
Ave Floor Effects (%)	19	3589	0.578	1.302	19	3456	3.624	0.827	0.063
Ave Ceiling Effects (%)	19	3589	0.821	1.420	19	3456	5.710	0.901	0.0094

941
942