Published in final edited form as:

Foot Ankle Int. 2019 January ; 40(1): 65–73. doi:10.1177/1071100718800304.

PROMIS and FAAM Minimal Clinically Important Differences in Foot and Ankle Orthopedics

Man Hung, PhD^{1,2,3}, Judith F. Baumhauer, MD, MPH⁴, Frank W. Licari, DDS⁵, Maren W. Voss, ScD¹, Jerry Bounsanga, BS¹, Charles L. Saltzman, MD¹

¹Department of Orthopaedic Surgery Operations, University of Utah, Salt Lake City, UT, USA

²Department of Family & Preventive Medicine, University of Utah, Salt Lake City, UT, USA

³Department of Population Health Sciences, University of Utah, Salt Lake City, UT, USA

⁴Department of Orthopaedic Surgery, University of Rochester, Rochester, NY, USA

⁵College of Dental Medicine, Roseman University of Health Sciences, South Jordan, UT, USA

Abstract

Background—Establishing score points that reflect meaningful change from the patient perspective is important for interpreting patient-reported outcomes. This study estimated the minimum clinically important difference (MCID) values of 2 Patient-Reported Outcomes Measurement Information System (PROMIS) instruments and the Foot and Ankle Ability Measure (FAAM) Sports subscale within a foot and ankle orthopedic population.

Methods—Patients seen for foot and ankle conditions at an orthopedic clinic were administered the PROMIS Physical Function (PF) v1.2, the PROMIS Pain Interference (PI) v1.1, and the FAAM Sports at baseline and all follow-up visits. MCID estimation was conducted using anchor-based and distribution-based methods.

Results—A total of 3069 patients, mean age of 51 years (range = 18–94), were included. The MCIDs for the PROMIS PF ranged from approximately 3 to 30 points (median = 11.3) depending on the methods being used. The MCIDs ranged from 3 to 25 points (median = 8.9) for the PROMIS PI, and from 9 to 77 points (median = 32.5) for the FAAM Sports.

Conclusions—This study established a range of MCIDs in the PROMIS PF, PROMIS PI, and FAAM Sports indicating meaningful change in patient condition. MCID values were consistent across follow-up periods, but were different across methods. Values below the 25th percentile of MCIDs may be useful for low-risk clinical decisions. Midrange values (eg, near the median) should be used for high stakes decisions in clinical practice (ie, surgery referrals). The MCID values within the interquartile range should be utilized for most decision making.

Level of Evidence—Level I, diagnostic study, testing of previously developed diagnostic measure on consecutive patients with reference standard applied.

Keywords

MCID; patient-reported outcomes; PROMIS; FAAM; physical function; pain; orthopedics

---

Introduction

Patient-reported outcomes (PROs) have emerged as key components in high quality clinical care and patient satisfaction. PROs have the capacity to improve the patient experience while offering reliable and valid test scores with minimal respondent burden.^{11,21} In orthopedics in particular, a patient's personal account of their improvement in pain or function is critical to understanding treatment effects. Recent developments in PROs include the use of computerized adaptive testing (CAT). CAT is an administration format that can minimize test burden by reducing the number of questions that need to be completed while still allowing maximized information gain on treatment outcomes.^{7,9}

The Patient-Reported Outcomes Measurement Information System (PROMIS) has developed CATs, generated after exhaustive review of pre-existing testing items, including categorization, revisions, and the application of item response theory to calibrate the individual items.¹² Across multiple orthopedic populations, PROMIS instruments have exhibited adequate coverage (the ability to detect high and low levels of pain and function) and precision (accurate categorization of pain or function) while lowering overall test burden.^{16,18–20} For these reasons, PROMIS instruments represent one of the best available options for measuring orthopedic patient outcomes.

PROMIS instruments can be considered an important contribution to clinical practice,⁵ but only if patient scores are meaningful. The minimal clinically important difference (MCID) is the smallest amount of change in score that can be considered meaningful from the patient perspective.^{8,24} An MCID is the smallest value of score beyond random or chance variation, indicating change that is real and not just measurement error. MCIDs have 2 useful applications. They can be an important benchmark of treatment effectiveness, which has implications for clinical decision making and helping patients set expectations of treatment outcomes. They are also useful in comparative effectiveness research and can help determine the sample size needed to study treatment effects. Meaningful interpretations of study results, enabled by an understanding of the MCID, are essential for research and practice on treatment outcomes.⁴⁷

Given the recent development of PROMIS instruments since 2004,⁶ the establishment of MCIDs for these tools has only begun among specific patient populations, including pediatrics⁴⁵ cancer⁴⁹ and multiple sclerosis.² Little MCID analysis has been done in orthopedics. In one previous orthopedic study, a PROMIS Pain Interference (PI) anchorbased mean change score range of 3.5 to 5.5 points was recommended for low back pain patients.¹ MCIDs can be determined 2 ways: with anchor-based or distribution-based methods. Anchor-based methods tie patient test scores to some other measures of the patients' response level or treatment effect.³⁴ In the distribution-based methods, statistical calculations based on the distribution of scores from the sample population determine the level of change that is beyond chance occurrence. Distribution methods focus on how much

scores vary between patients, rather than comparing patient scores to a patient-based anchor question.³⁴

The focus of the present study was to enable the use of PROs in a meaningful way in the foot and ankle patient population. The Foot and Ankle Ability Measure (FAAM) Sports is a well-recognized scale derived using standard test theory. Initial validation studies on the FAAM indicated a 5.7-point minimum detectable change (MDC),³³ though other research has shown as little as a 2 point difference as meaningful change.²⁶ MCIDs in the foot and ankle patient population for the PROMIS Physical Function (PF) and the PROMIS PI have not been investigated. This study examined MCIDs of the PROMIS PF, PI, and the FAAM Sports in a general foot and ankle orthopedic population using anchor-based and distribution-based approaches with cutoff values of varying precision. The goal was to provide clinicians and researchers with MCID values for each instrument that would be indicative of meaningful clinical change in foot and ankle patients.

Methods

Patient Sample

Consecutive patients seeking care at an orthopedic foot and ankle clinic between September 2013 and December 2017 were included if aged 18 or older. Clinic-wide approximately 1%–2% of patients refused to complete the PRO questionnaires. Follow-up data from return clinic visits were collected and analyzed at 4 follow-up periods, including (1) 3-month follow-up (ie, 80–100 days after initial assessment), (2) >3-month follow-up (ie, 90 days or more after initial assessment), (3) 6-month follow-up (ie, 170–190 days after initial assessment), and (4) >6-month follow-up (ie, 180 days or more after initial assessment) as common follow-up period in orthopedics.^{4,10,22,27–30,37,42,46} Depending on the timing of follow-up care, different patients were included in each analysis based on time elapsed since their first clinic visit. For the PROMIS PF, there were between 6491 and 8681 clinic patients who were not eligible for inclusion in each analysis. Those not included had baseline PROMIS PF scores ranging from 2 to 4 points higher, suggesting higher function. There were 3545 to 4795 clinic patients not in the PROMIS PI followup windows. The non-follow-up group had an average of 2 to 3 points lower PI scores, indicating less pain interference. There were 1795 to 2156 nonparticipants for the FAAM Sports with average baseline scores that were 6 to 14 points higher, indicating higher function. The first clinic visit established the baseline levels of function from which change scores were calculated at each time point.

A total of 3069 patients experiencing foot and ankle treatment between September 2013 and December 2017 at a single university orthopedic clinic were eligible for inclusion. Their average age was 51.3 years (SD=16.45), ranging from 18 to 94. Overall, 60.5% of the sample was female, 89.9% was Caucasian, and 4.8% reported Hispanic ethnicity (Table 1). Procedures for foot and ankle treatments included continuous infusion of anesthetic agent into sciatic nerve, removal of implants, fusion of ankle or foot joints, partial excision of foot or ankle, primary open repair of ruptured Achilles tendon, correction of hammertoe, correction of hallux rigidus, repair of disrupted collateral ligament of ankle, amputation of foot, open treatment of dislocated foot or ankle joints, and arthroscopy of ankle.

The goals of MCID determination were 2-fold: first, to provide a benchmark for meaningful interpretation of scores that could aid clinical treatment decisions and help in guiding patient expectations; and second, to allow for sample size estimation or respondent analysis for scientific research and discoveries. Besides speculation, there is no empirical evidence of disease severity, follow-up length, or specific patient groupings impacting MCID values to any large degree.^{25,36,44} Thus, it was deemed appropriate to conduct this study on a general sample of patients seeking care for a variety of foot and ankle conditions (both operative and nonoperative) to determine MCIDs that would not be treatment specific, as done in other studies.^{14,23,26}

Patient-Reported Outcome Measures

Three PRO instruments were administered via a web-based portal at the time of each clinic visit. The PROMIS PF v1.2 CAT and the PROMIS PI v1.1 CAT were administered to patients attending clinic visits between 2013 and 2017. The FAAM Sports was added to the administration battery in 2016 and administered through the end of the study period in 2017, resulting in fewer patient samples for the FAAM Sports MCID study.

The PROMIS PF v1.2 CAT draws from a 121-item bank, and the PROMIS PI v1.1 CAT has a 40-item bank. The CAT administration minimizes patient burden by selecting questions from the item bank that adapt based on prior patient answers.^{15,17} For example, patients would not be asked, “Are you able to run or jog for two miles (3 km)?” if they had previously answered “Unable to Do” to the question, “Are you able to stand for one hour?” The PROMIS instruments use standardized T scores with a mean of 50 and a standard deviation (SD) of 10 in the US general population.⁴¹ The range of possible scores on the PROMIS instruments is technically negative infinity to infinity, though for practical purposes scores primarily fall between 20 and 80 points. Higher scores on the PROMIS PF indicate higher patient function, whereas higher scores on the PROMIS PI indicate greater pain interference.

The FAAM consists of 2 subscales, the 21-item Activities of Daily Living subscale and the 8-item sports subscale. Only the Sports subscale was administered to this orthopedic foot and ankle patient population being evaluated in the clinics. The FAAM Sports items are scored on a 5-point scale from 0 (unable to do) to 4 (no difficulty at all), for a total score range from 0 to 32, with raw scores transformed into percentage scores.³² Higher percentage scores reflect higher levels of function.

Statistical Analysis

Patient characteristics and demographics were measured and described. The PROMIS and FAAM Sports scores from 4 follow-up time periods were compared individually to baseline scores. Change scores were calculated as the absolute value difference between the baseline score and the follow-up score.

For anchor-based methods, patient scores were anchored to their self-reports of meaningful change. At follow-up visits, patients answered 2 validated anchored questions: (1) Compared to your FIRST EVALUATION at the xxx: how would you describe your physical function now? (much worse, worse, slightly worse, no change, slightly improved, improved, much

improved), and (2) Compared to your FIRST EVALUATION at the xxx: how would you describe your episodes of PAIN now? (much worse, worse, slightly worse, no change, slightly improved, improved, much improved). The first anchor question addressed physical function, so it served as a reference standard for categorizing meaningful change for the PROMIS PF and the FAAM Sports, and the second anchor question related to pain served as a reference standard for the PROMIS PI. Patients were grouped as having meaningful change (much worse, worse, improved, much improved) or experiencing minimal or no-change (slightly worse, no change, slightly improved) as done in prior research.^{14,38}

The area under the receiver operating characteristic (ROC) curve was another anchor-based method used in this study to determine the optimal MCIDs. This method was used to maximize the point in which the instruments discriminate both specificity and sensitivity. The ROC cut-off point was based on the maximum of Youden's J index, defined as (sensitivity + specificity) – 1,⁵⁰ with equal weight given to sensitivity and specificity.

Distribution-based methods consisted of SD of the test scores and also the MDC based on standard errors of measurement (SEMs) of the follow-up scores. The $\frac{1}{2}$ SD^{40,43} and $\frac{1}{3}$ SD^{13,48} are 2 variation-based estimates of the MCID commonly used in the literature. The MDC was calculated at 3 confidence intervals, constituting the smallest score change that likely reflects a true change in condition with 90%, 95%, and 99% certainty. The formulas for calculating the MDC are as follows: $MDC_{@90\%} = 1.65 \times \text{square root of } (2) \times SEM$; $MDC_{@95\%} = 1.96 \times \text{square root of } (2) \times SEM$; $MDC_{@99\%} = 2.56 \times \text{square root of } (2) \times SEM$. The $SEM = SD \times \text{square root of } (1 - r)$, where r is a reliability coefficient represented by Cronbach alpha.

Statistical computing were performed using SPSS 25.0 (IBM SPSS Statistics for Windows; IBM Corp, Armonk, NY) and R 3.2.5 (R Development Core Team, Vienna, Austria).³⁹ Graphical visualization was created using Tableau 10.4 (Tableau Mission, Seattle, WA).

Results

Anchor-Based Methods

Mean Change—The mean change score for PROMIS PF in patients reporting meaningful change ranged from 7.8 to 9.8 across the 4 follow-up time points. For the PROMIS PI, the change scores ranged from 7.4 to 9.4. For the FAMM Sports, the range of change score was from 25.1 to 31.2 (Table 2).

ROC Curve—The area under the ROC curve was used to identify the optimal cut-off point between meaningful and nonmeaningful change groups. The true positive (sensitivity) rate is plotted as a function of the false positive (1 – specificity) rate for possible cut-off values, so the point at the top left corner of the graph represents the best estimate of identifying true change (see Figure 1). This optimal cutoff point serves as an estimate for MCID values. For the PROMIS PF, the cut-off point ranged from 7.9 to 13.2 for all follow-up time-points. For the PROMIS PI, the cut-off ranged from 5.5 to 12.4. For the FAAM Sports, the range was from 17.2 to 35.9 (see Table 2). The c-statistics at 3-month, >3-month, 6-month, and >6-month for the PF CAT are 0.531, 0.588, 0.520, and 0.530, respectively; for the PI CAT, the

values are 0.627, 0.584, 0.773, and 0.625, respectively; and for FAAM Sports, they are 0.577, 0.636, 0.507, and 0.757, respectively.

Distribution-Based Methods

Standard Deviation—The MCID values for the PROMIS PF ranged from 4.5 to 4.7 using the $\frac{1}{2}$ SD method, and ranged from 3.0 to 3.1 using the $\frac{1}{3}$ SD method across the 4 different follow-up time points. For the PROMIS PI, $\frac{1}{2}$ SD gave a range of 4.1 to 4.3, and $\frac{1}{3}$ SD gave a range of 2.8 to 3.0. For the FAAM Sports, its MCID values ranged from 12.0 to 13.3 using $\frac{1}{2}$ SD, and from 8.7 to 8.8 using $\frac{1}{3}$ SD (Table 3).

Minimum detectable change—The MDC methods yielded various MCID values depending on the precision levels (ie, $\text{MDC}_{@90\%}$, $\text{MDC}_{@90\%}$, and $\text{MDC}_{@90\%}$) desired. The MCID estimates ranged from 15.2 to 29.7 for the PROMIS PF, from 12.6 to 24.6 for the PROMIS PI, and from 40.1 to 76.6 for the FAAM Sports (Table 3).

Discussion

Information obtained from PRO data communicates the perspectives of patients and provides a mechanism for evaluating treatment effectiveness that is meaningful to patients. MCID scores calculated from specific patient populations are useful to help clinicians and patients gain perspective on relative levels of meaningful change. This study evaluated 3 PRO instruments at 4 follow-up periods applying 7 analytical methods, yielding a total of 28 MCID values for each instrument. We identified a range of MCIDs in the PROMIS PF, PROMIS PI, and FAAM Sports indicative of meaningful change in patient outcomes. Across different follow-up time periods, each specific method for MCID estimation provided relatively similar MCID values. However, across different methods, the MCID values were substantially different. MCID calculation methods, rather than the length of follow-up, were responsible for the large variation of MCID values.

The different methods for calculating MCIDs have different limitations and advantages. The use of an anchor question addressing perceptions of improvement or deterioration in the ROC method is an opportunity to minimize false-positives, when there is a change in a patient score that does not relate to patient perceptions of change in function. But anchor questions rely on recall and can introduce bias.^{8,35} The largest MCID values in this current study were obtained using the distributionbased method of statistically calculating detectable change with a high degree of precision. The $\text{MDC}_{@95\%}$ and $\text{MDC}_{@99\%}$ methods are extremely strict criteria that provide an extremely high precision but are not commonly found in other studies, so it was not surprising that the resulting high MCID values are not commonly found in the literature. This greater degree of precision may be appropriate when there is a need for definite assurance that important change has occurred, rather than just possible or probable assurance of important change.³¹ This precision and caution might be appropriate for studies evaluating high-cost, high-risk treatments. More stringent MCID cut-offs might also be useful to establish treatment effectiveness before a return to strenuous labor. However, a lower MCID cut-off might be appropriate for a low-risk referral to physical therapy.

There is no consensus on the best method for determining MCIDs.^{8,44} This study employed a variety of methods ranging from lenient to extremely strict precision criteria, producing low to extremely high MCID values, respectively. The MCIDs for the PROMIS PF in the foot and ankle orthopedic patient population could be as small as 3 score points, but as large as 30 points (median = 11.3) depending on the level of precision. For the PROMIS PI, the MCID values ranged from 3 points to 25 points (median = 8.9). For the FAAM Sports, the MCIDs ranged from 9 to 77 points (median = 32.5). This lack of agreement between MCID values calculated with differing methods is consistent with prior findings.³

As PROs are increasingly incorporated into clinical practice and administered and scored in real time, there is an opportunity for understanding MCIDs to inform clinical decision making. However, because there is no one-size-fits-all MCID calculation, it would be misleading to suggest a single fixed MCID value for clinical practice. Though simplifying MCIDs to a single number is tempting, a single value is often unstable and lacks usefulness.⁴⁷ For screening purpose, lower MCIDs (below the 25th percentile) may suffice, but for high-stakes decisions, at least midrange (eg, median) or higher should be considered. An example of a screening task is to find out whether a patient can return to normal work activities following ankle arthroscopy. Example of high-stakes decision is to decide whether a patient can return to high-level sport activities after ankle replacement. Figure 2 presents a visual summary of all of the MCID values and displays a shaded area including MCID values within the interquartile range (eg, between the 25th percentile and the 75th percentile). Those MCID values within the interquartile range would probably be used for most decision making. Ultimately, personal value judgments should be exercised in the selection of appropriate MCIDs under different circumstances.

Limitations

Our patient sample consisted of individuals seeking foot and ankle treatment for a wide range of conditions including amputations, fusions, reconstructions, repairs, arthroplasty, bunionectomy, and tendon transfers. Although some might consider the heterogeneity problematic, the broad patient sample is likely not a limitation because MCID determination in many cases is not dependent on disease severity or length of follow-up.^{25,36,44} Critically, the concept of MCID determination is not intended to be treatment specific.²⁵ The use of a general clinic population is a considered and wise approach for validation studies,⁴⁴ as the application of MCID in clinical practice will reflect the same wide range of patient conditions and contexts. It has also been recommended that MCID determination should be evaluated using multiple types of anchors (such as clinical end-points and alternative PRO measures in addition to global ratings of change) and with multiple patient samples.⁴⁰ Future research should evaluate the MCID of these instruments with alternative anchors and in other patient populations.

This study had limitations related to the sample characteristics, as they were not representative of the demography of the United States. Future research should be conducted in a more diverse sample to gain more insights. The mean change scores evaluated were based on patients' baseline scores on average. Patients with a lower baseline score may reach a higher change score more easily than those with a higher baseline score, as there is

less room for improvement when initial scores are high. Future work can focus on using Rasch-based methods to establish MCIDs linked to individual baseline scores.

Our patients were grouped into change and no-change groups based on a global rating of change that relies on retrospective recall of prior function. These types of questions have been criticized as more responsive to current health status than baseline levels of function.^{8,35} Furthermore, different anchor questions may produce different results. Most of the ROC curves had c-statistics slightly above 0.50, which was not informative in discriminating patients who had improvement versus no improvement. Only PI CAT and FAAM Sports had a c-statistics above 0.70 demonstrating higher discrimination in this clinical sample. Such low discrimination for most ROCs' follow-up periods could be related to the relatively short follow-up periods (6 months or less); this clinical sample probably requires 1 year or longer follow-up to show stronger effect. Low discrimination could also be due to inherent issues with the instruments or the specific anchor questions being used. It was because of these limitations in the anchor-based methods that this current study was designed to be comprehensive in nature from the outset, covering multiple anchor-based and distribution-based methods in order to triangulate results and provide alternative perspectives.

Conclusions

This study identified a range of MCIDs in the PROMIS PF, PROMIS PI, and FAAM Sports that were indicative of meaningful change in patients' foot and ankle conditions at 3-month follow-up and beyond. Notably, different methods to calculate MCIDs produced different results. Readers should select MCID values based on the clinical needs of their population.

Acknowledgment

This project was funded by the National Institute of Arthritis and Musculoskeletal and Skin Diseases of the National Institutes of Health under award number U01AR067138 (Principal Investigator Man Hung, PhD). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This project was funded by the National Institute of Arthritis and Musculoskeletal and Skin Diseases of the National Institutes of Health under award number U01AR067138.

Declaration of Conflicting Interests

The author(s) declared the following potential conflicts of interest with respect to the research, authorship, and/or publication of this article: Man Hung, PhD, Maren W. Voss, ScD, Jerry Bounsanga, BS, and Charles L. Saltzman, MD, report grants from NIH during the conduct of study. Judith F. Baumhauer, MD, MPH, is vice president of PROMIS Health Organization. ICMJE forms for all authors are available online.

References

1. Amtmann D, Kim J, Chung H, Askew RL, Park R, Cook KF. Minimally important differences for Patient Reported Outcomes Measurement Information System pain interference for individuals with back pain. *J Pain Res.* 2016;9:251–255. [PubMed: 27175093]
2. Askew RL, Kim J, Chung H, Cook KF, Johnson KL, Amtmann D. Development of a crosswalk for pain interference measured by the BPI and PROMIS pain interference short form. *Qual Life Res.* 2013;22(10):2769–2776. [PubMed: 23539469]

3. Beaton DE, van Eerd D, Smith P, et al. Minimal change is sensitive, less specific to recovery: a diagnostic testing approach to interpretability. *J Clin Epidemiol*. 2011;64(5):487–496. [PubMed: 21109396]
4. Carmont MR, Silbernagel KG, Nilsson-Helander K, Mei-Dan O, Karlsson J, Maffulli N. Cross cultural adaptation of the Achilles tendon Total Rupture Score with reliability, validity and responsiveness evaluation. *Knee Surg Sports Traumatol Arthrosc*. 2013;21(6):1356–1360. [PubMed: 22847247]
5. 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(11):1179–1194. [PubMed: 20685078]
6. Cella D, Yount S, Rothrock N, et al. The Patient-Reported Outcomes Measurement Information System (PROMIS): progress of an NIH Roadmap cooperative group during its first two years. *Med Care*. 2007;45(5)(suppl 1):S3.
7. Chakravarty EF, Bjorner JB, Fries JF. Improving patient reported outcomes using item response theory and computerized adaptive testing. *J Rheumatol*. 2007;34(6):1426–1431. [PubMed: 17552069]
8. Cook CE. Clinimetrics corner: the Minimal Clinically Important Change Score (MCID): a necessary pretense. *J Man Manip Ther*. 2008;16(4):E82–E83. [PubMed: 19771185]
9. Cook KF, O'Malley KJ, Roddey TS. Dynamic assessment of health outcomes: time to let the CAT out of the bag? *Health Serv Res*. 2005;40(5, pt 2):1694–1711. [PubMed: 16179003]
10. Cornell CN, Levine D, O'Doherty J, Lyden J. Unipolar versus bipolar hemiarthroplasty for the treatment of femoral neck fractures in the elderly. *Clin Orthop Relat Res*. 1998;348: 67–71.
11. Deutsch L, Smith L, Gage B, Kelleher C, Garfinkel D. Patient-Reported Outcomes in Performance Measurement: Commissioned Paper on PRO-Based Performance Measures for Healthcare Accountable Entities. Washington, DC: National Quality Forum; 2012.
12. DeWalt DA, Rothrock N, Yount S, Stone AA. Evaluation of item candidates: the PROMIS qualitative item review. *Med Care*. 2007;45(5)(suppl 1):S12–S21. [PubMed: 17443114]
13. du Bois RM, Weycker D, Albera C, et al. Six-minute-walk test in idiopathic pulmonary fibrosis: test validation and minimal clinically important difference. *Am J Respir Crit Care Med*. 2011;183(9):1231–1237. [PubMed: 21131468]
14. Franchignoni F, Vercelli S, Giordano A, Sartorio F, Bravini E, Ferriero G. Minimal clinically important difference of the disabilities of the arm, shoulder and hand outcome measure (DASH) and its shortened version (QuickDASH). *J Orthop Sports Phys Ther*. 2014;44(1):30–39. [PubMed: 24175606]
15. Hung M, Baumhauer JF, Latt LD, Saltzman CL, SooHoo NF, Hunt KJ. Validation of PROMIS (R) Physical Function computerized adaptive tests for orthopaedic foot and ankle outcome research. *Clin Orthop Relat Res*. 2013;471(11):3466–3474. [PubMed: 23749433]
16. Hung M, Clegg DO, Greene T, Saltzman CL. Evaluation of the PROMIS physical function item bank in orthopaedic patients. *J Orthop Res*. 2011;29(6):947–953. [PubMed: 21437962]
17. Hung M, Clegg DO, Greene T, Weir C, Saltzman CL. A lower extremity physical function computerized adaptive testing instrument for orthopaedic patients. *Foot Ankle Int*. 2012;33(4): 326–335. [PubMed: 22735205]
18. Hung M, Franklin JD, Hon SD, Cheng C, Conrad J, Saltzman CL. Time for a paradigm shift with computerized adaptive testing of general physical function outcomes measurements. *Foot Ankle Int*. 2014;35(1):1–7. [PubMed: 24101733]
19. 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(2):158–163. [PubMed: 24173018]
20. 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(8):439–443. [PubMed: 24378399]

21. Hung M, Zhang W, Chen W, et al. Patient-reported outcomes and total health care expenditure in prediction of patient satisfaction: results from a national study. *JMIR Public Health Surveil.* 2015;1(2):e13.
22. Ibrahim T, Beiri A, Azzabi M, Best AJ, Taylor GJ, Menon DK. Reliability and validity of the subjective component of the American Orthopaedic Foot & Ankle Society clinical rating scales. *J Foot Ankle Surg.* 2007;46(2):65–74. [PubMed: 17331864]
23. Irrgang JJ, Lubowitz JH. Measuring arthroscopic outcome. *Arthroscopy.* 2008;24(6):718–722. [PubMed: 18536080]
24. Jaeschke R, Singer J, Guyatt GH. Measurement of health status. Ascertaining the minimal clinically important difference. *Control Clin Trials.* 1989;10(4):407–415. [PubMed: 2691207]
25. Jones PW, Beeh KM, Chapman KR, Decramer M, Mahler DA, Wedzicha JA. Minimal clinically important differences in pharmacological trials. *Am J Respir Crit Care Med.* 2014;189(3):250–255. [PubMed: 24383418]
26. Kivlan BR, Martin RL, Wukich DK. Responsiveness of the foot and ankle ability measure (FAAM) in individuals with diabetes. *Foot (Edinburgh, Scotland).* 2011;21(2):84–87.
27. Kotsis SV, Chung KC. Responsiveness of the Michigan Hand Outcomes Questionnaire and the Disabilities of the Arm, Shoulder and Hand questionnaire in carpal tunnel surgery. *J Hand Surg Am.* 2005;30(1):81–86. [PubMed: 15680560]
28. Landauer F, Wimmer C, Behensky H. Estimating the final outcome of brace treatment for idiopathic thoracic scoliosis at 6-month follow-up. *Pediatr Rehabil.* 2003;6(3–4):201–207. [PubMed: 14713586]
29. Little DG, MacDonald D. The use of the percentage change in Oswestry Disability Index score as an outcome measure in lumbar spinal surgery. *Spine (Phila Pa 1976).* 1994;19(19):2139–2143. [PubMed: 7809744]
30. MacDermid JC, Richards RS, Donner A, Bellamy N, Roth JH. Responsiveness of the short form-36, disability of the arm, shoulder, and hand questionnaire, patient-rated wrist evaluation, and physical impairment measurements in evaluating recovery after a distal radius fracture. *J Hand Surg Am.* 2000;25(2):330–340. [PubMed: 10722826]
31. Man-Son-Hing M, Laupacis A, O'Rourke K, et al. Determination of the clinical importance of study results. *J Gen Intern Med.* 2002;17(6):469–476. [PubMed: 12133163]
32. Martin RL, Irrgang JJ. A survey of self-reported outcome instruments for the foot and ankle. *J Orthop Sports Phys Ther.* 2007;37(2):72–84. [PubMed: 17366962]
33. Martin RL, Irrgang JJ, Burdett RG, Conti SF, Van Swearingen JM. Evidence of validity for the Foot and Ankle Ability Measure (FAAM). *Foot Ankle Int.* 2005;26(11):9682–983.
34. McGlothlin AE, Lewis RJ. Minimal clinically important difference: defining what really matters to patients. *JAMA.* 2014;312(13):1342–1343. [PubMed: 25268441]
35. Norman GR, Stratford P, Regehr G. Methodological problems in the retrospective computation of responsiveness to change: the lesson of Cronbach. *J Clin Epidemiol.* 1997;50(8):869–879. [PubMed: 9291871]
36. Ostelo RW, Deyo RA, Stratford P, et al. Interpreting change scores for pain and functional status in low back pain: towards international consensus regarding minimal important change. *Spine (Phila Pa 1976).* 2008;33(1):90–94. [PubMed: 18165753]
37. Paatelma M, Kilpikoski S, Simonen R, Heinonen A, Alen M, Videman T. Orthopaedic manual therapy, McKenzie method or advice only for low back pain in working adults: a randomized controlled trial with one year follow-up. *J Rehabil Med.* 2008;40(10):858–863. [PubMed: 19242624]
38. Polson K, Reid D, McNair PJ, Larmer P. Responsiveness, minimal importance difference and minimal detectable change scores of the shortened disability arm shoulder hand (QuickDASH) questionnaire. *Man Ther.* 2010;15(4):404–407. [PubMed: 20434942]
39. R Development Core Team. *R: A Language and Environment for Statistical Computing.* Vienna, Austria: R Foundation for Statistical Computing; 2010.
40. Revicki D, Hays RD, Cella D, Sloan J. Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. *J Clin Epidemiol.* 2008;61(2):102–109. [PubMed: 18177782]

41. Rose M, Bjorner JB, Gandek B, Bruce B, Fries JF, Ware JE. The PROMIS Physical Function item bank was calibrated to a standardized metric and shown to improve measurement efficiency. *J Clin Epidemiol.* 2014;67(5):516–526. [PubMed: 24698295]
42. Segal NA, Glass NA, Teran-Yengle P, Singh B, Wallace RB, Yack HJ. Intensive gait training for older adults with symptomatic knee osteoarthritis. *Am J Phys Med Rehabil.* 2015;94(10)(suppl 1): 848–858. [PubMed: 25768068]
43. Sloan JA. Assessing the minimally clinically significant difference: scientific considerations, challenges and solutions. *COPD.* 2005;2(1):57–62. [PubMed: 17136963]
44. Terwee CB, Roorda LD, Dekker J, et al. Mind the MIC: large variation among populations and methods. *J Clin Epidemiol.* 2010;63(5):524–534. [PubMed: 19926446]
45. Thissen D, Liu Y, Magnus B, et al. Estimating minimally important difference (MID) in PROMIS pediatric measures using the scale-judgment method. *Qual Life Res.* 2016;25(1):13–23. [PubMed: 26118768]
46. Uchiyama S, Imaeda T, Toh S, et al. Comparison of responsiveness of the Japanese Society for Surgery of the Hand version of the carpal tunnel syndrome instrument to surgical treatment with DASH, SF-36, and physical findings. *J Orthop Sci.* 2007;12(3):249–253. [PubMed: 17530376]
47. Wright A, Hannon J, Hegedus EJ, Kavchak AE. Clinimetrics corner: a closer look at the minimal clinically important difference (MCID). *J Man Manip Ther.* 2012;20(3):160–166. [PubMed: 23904756]
48. Yost KJ, Eton DT. Combining distribution- and anchor-based approaches to determine minimally important differences: the FACIT experience. *Eval Health Prof.* 2005;28(2):172–191. [PubMed: 15851772]
49. Yost KJ, Eton DT, Garcia SF, Cella D. Minimally important differences were estimated for six patient-reported outcomes measurement information system-cancer scales in advancedstage cancer patients. *J Clin Epidemiol.* 2011;64(5):507–516. [PubMed: 21447427]
50. Youden WJ. Index for rating diagnostic tests. *Cancer.* 1950;3(1):32–35. [PubMed: 15405679]

Figure 1.

A visual representation of the receiver operating characteristic curves of the PROMIS PF, PI, and FAAM Sports at various follow-time periods. The c-statistics for the 3-month, >3-month, 6-month, and >6-month time points are 0.531, 0.588, 0.520, and 0.530 for the PF CAT, respectively; 0.627, 0.584, 0.773, and 0.625 for the PI CAT, respectively; and 0.577, 0.636, 0.507, and 0.757 for the FAAM Sports, respectively. (CAT, computerized adaptive test; FAAM Sports, Foot and Ankle Ability Measure Sports subscale; MDC, minimum

detectable change; PROMIS, Patient-Reported Outcomes Measurement Information System; PROMIS PF, PROMIS Physical Function; PROMIS PI, Pain Interference.)

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Figure 2.

MCID range based on method for PROMIS PF, PI, and FAAM Sports. The figure presents a visual summary of all of the MCID values and displays a shaded area including MCID values within the interquartile range (ie, between the 25th percentile and the 75th percentile). FAAM Sports, Foot and Ankle Ability Measure Sports subscale; MCID, minimal clinically important difference; PROMIS, Patient-Reported Outcomes Measurement Information System; PROMIS PF, PROMIS Physical Function; PROMIS PI, Pain Interference.

Table 1. Demographics of Patients ($n = 3069$).

Patient Characteristics	<i>n</i>	Percent	Mean (SD)	Range
Age, y			51.34 (16.45)	18–94
Gender				
Male	1211	39.5		
Female	1858	60.5		
Race				
White or Caucasian	2758	89.9		
Black or African American	42	1.4		
Asian	37	1.2		
American Indian and Alaska Native	24	0.8		
Native Hawaiian or Pacific Islander	19	0.6		
Other	151	4.9		
Missing	38	1.2		
Ethnicity				
Non-Hispanic	2189	93.9		
Hispanic	111	4.8		
Missing	31	1.3		

Table 2.

Anchor-Based Methods for Determination of MCID Values for the PROMIS Instruments and the FAAM Sports.^a

Instrument	n	No Change (SD)	n	Change (SD)	ROC
3-mo follow-up from baseline					
PROMIS PF	32	7.8 (5.4)	42	9.0 (7.3)	12.0
PROMIS PI	41	5.5 (4.4)	64	8.4 (6.6)	8.0
FAAM Sports	25	10.2 (16.2)	56	27.5 (22.9)	35.9
>3-mo follow-up from baseline					
PROMIS PF	219	6.1 (5.5)	337	8.4 (7.3)	7.9
PROMIS PI	187	5.5 (4.4)	321	7.4 (6.1)	12.4
FAAM Sports	96	17.9 (14.8)	190	29.5 (23.4)	25.5
6-mo follow-up from baseline					
PROMIS PF	8	8.0 (7.7)	19	9.8 (9.6)	13.2
PROMIS PI	9	4.4 (2.5)	22	9.4 (7.5)	5.5
FAAM Sports	4	19.3 (11.1)	19	25.1 (24.0)	33.9
>6-mo follow-up from baseline					
PROMIS PF	128	6.6 (6.5)	178	7.8 (7.7)	10.5
PROMIS PI	100	5.0 (3.9)	170	8.0 (6.7)	5.5
FAAM Sports	21	10.5 (7.0)	24	31.2 (25.3)	17.2

Abbreviations: FAAM Sports, Foot and Ankle Ability Measure Sports subscale; MCID, minimal clinically important difference; PROMIS, Patient-Reported Outcomes Measurement Information System; PROMIS PF, PROMIS Physical Function; PROMIS PI, Pain Interference.

^aMCID values are highlighted in bold.

Table 3.

Distribution-Based Methods for Determination of MCID Values for the PROMIS Instruments and the FAAM Sports.^a

Instrument	n	SD		MDC			
		1/2	1/3	90%	95%	99%	
3-mo follow-up from baseline							
PROMIS PF	808	4.5	3.0	16.0	19.0	24.8	
PROMIS PI	647	4.4	2.9	13.8	16.3	21.3	
FAAM Sports	256	13.0	8.7	49.4	58.7	76.6	
>3-mo follow-up from baseline							
PROMIS PF	2609	4.6	3.0	15.2	18.0	23.5	
PROMIS PI	2107	4.2	2.8	12.6	14.9	19.5	
FAAM Sports	875	13.3	8.8	48.1	57.2	74.7	
6-mo follow-up from baseline							
PROMIS PF	410	4.7	3.1	19.2	22.8	29.7	
PROMIS PI	327	4.1	2.8	15.9	18.9	24.6	
FAAM Sports	124	13.0	8.7	48.3	57.3	74.9	
>6-mo follow-up from baseline							
PROMIS PF	1825	4.5	3.0	14.7	17.5	22.8	
PROMIS PI	1667	4.3	2.9	13.2	15.7	20.5	
FAAM Sports	367	13.1	8.7	40.1	47.6	62.1	

Abbreviations: FAAM Sports, Foot and Ankle Ability Measure Sports subscale; MCID, minimal clinically important difference; MDC, minimum detectable change; PROMIS, Patient-Reported Outcomes Measurement Information System; PROMIS PF, PROMIS Physical Function; PROMIS PI, PROMIS Pain Interference.

^aMCID values are highlighted in bold.