Identifying Minimal and Meaningful Change in PROMIS® for Rheumatoid Arthritis: Use of Multiple Methods and Perspectives

Susan J. Bartlett, PhD[1,2,3], **Anna Kristina Gutierrez, MD**[3], **Kathleen M. Andersen, MS**[4,5], **Vivian P. Bykerk, MD**[6], **Jeffrey R. Curtis, MD MPH**[7], **Uzma J. Haque, MD**[3], **Ana-Maria Orbai, MD MHS**[3], **Michelle R. Jones, BA**[3], **Clifton O. Bingham III, MD**[3]

[1]Divisions of Clinical Epidemiology and Rheumatology, McGill University, CORE Centre, 5252 de Maisonneuve #3D.57, Montreal, QC Canada H4A 3S5

[2]RI-MUHC, Centre for Outcomes Research and Evaluation and MUHC Center for Outcomes Research, 5252 de Maisonneuve #3D.57, Montreal, QC Canada H4A 3S5

[3]Division of Rheumatology, Johns Hopkins University, 5200 Eastern Avenue, MFL Center Tower, Suite 4100, Baltimore, MD 21224

[4]Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, 615 North Wolfe Street #W6021, Baltimore, MD 21205

[5]Center for Drug Safety and Effectiveness, Johns Hopkins University, 615 North Wolfe Street #W6035, Baltimore, MD 21205

[6]Rheumatology, Hospital for Special Surgery, Weill Cornell Medical College, 525 East 71st St, 7th floor, New York, NY, USA 10021

[7]University of Alabama at Birmingham, Division of Clinical Immunology and Rheumatology, Birmingham AL US

Abstract

Background: Rheumatoid arthritis (RA) is chronic, painful, disabling condition resulting in significant impairments in physical, emotional, and social health. We used different methods and perspectives to evaluate the responsiveness of PROMIS® short forms (SFs) and identify minimal and meaningful score changes.

Methods: Adults with RA enrolled in a multi-site prospective observational cohort completed PROMIS Physical Function, Pain Interference, Fatigue, Participation in Social Roles/Activities SFs, PROMIS-29, and pain, patient global, and rated change in specific symptoms and RA (a little vs. lot better or worse) at the second visit. Physicians recorded joint counts, MD Global

Corresponding Author: Susan J. Bartlett, PhD, McGill/MUHC Centre for Outcomes Research and Evaluation, 5252 de Maisonneuve, #3d.57, Montreal, QC, H4A 3S5, Susan.bartlett@mcgill.ca.

Assessment, and change in RA at visit 2. We compared mean score differences for minimal and meaningful improvement/worsening using patient and MD change ratings and distribution-based methods, and visually inspected empirical cumulative distribution function curves by change categories.

Results: The 348 adults were mostly (81%) female with longstanding RA. Using patient ratings, generally 1–3 point differences were observed for minimal change and 3–7 points for meaningful change. Larger differences were observed with patient vs. physician ratings and for symptom-specific vs. RA change. Mean differences were similar among SF versions. Prespecified hypotheses about change in PROMIS Physical Function, Pain Interference, Fatigue and Participation and legacy scales were supported.

Conclusions: PROMIS SFs and the PROMIS-29 Profile are responsive to change and generally distinguish between minimal and meaningful improvement and worsening in key RA domains. These data add to a growing body of evidence demonstrating robust psychometric properties of PROMIS and supporting use in RA care, research, and decision-making.

Background

Rheumatoid arthritis (RA), the most common form of chronic inflammatory arthritis, affects up to 1% of adults.(1, 2) The joint pain, swelling, and damage associated with RA greatly affects physical, emotional, and social health, and significantly impairs health-related quality of life (HRQL).(3–5) People with RA commonly experience transient increases in disease activity and episodes of disease flares which may necessitate a change in treatment.

Patient reported outcomes (PROs) can help to quantify how people feel and function, and how well treatment attenuates symptoms and functional impacts. An important characteristic of PROs is their responsiveness or ability to detect small changes in symptoms associated with both improving and worsening RA.

Different terms are used to define the amount of change in a PRO that is detectable, perceptible, or meaningful to people with a health condition. The *smallest detectable change* (SDC) is change beyond measurement error as estimated by the standard error of the mean (SEM). The *minimally important difference* (MID) is the mean difference estimated in relation to external anchors (e.g., patient assessments of feeling *a little better* or *a little worse*) or with statistical distributions (i.e., 0.2 or 0.5 SD).(6, 7) Conversely, *minimally clinically important difference* (MCID) is the change that patients, clinicians, and other stakeholders view as meaningful (e.g., *a lot better* or *a lot worse*).(8, 9) MCIDs are often used to decide if treatment is sufficiently controlling RA inflammation and hence symptoms and functional impacts both for individual patients, in clinical trials and comparative effectiveness studies.

The Patient Reported Outcome Measurement Information System® (PROMIS®) is a family of measures developed using advanced psychometric methods to precisely and reliably assess physical, emotional, and social health across chronic conditions.(10) PROMIS measures have been calibrated to the general US clinical population and use a common T-score metric (mean=50; SD=10) across scales to facilitate interpretation. Higher scores reflect more of the symptom/ function measured. We have previously shown that PROMIS® can reliably and precisely capture symptoms and functional impacts that people with RA say affect everyday life and HRQL.(11–14) In this study, we used multiple methods to evaluate the responsiveness of selected PROMIS short forms (SFs) and PROMIS-29 profile and identify minimal and meaningful changes score changes from the perspective of patients and clinicians, using statistical distributions, and in relation to RA disease activity indicators.

Methods

Design:

We used data from the first two visits of a prospective observational cohort study of adults with RA receiving care at academic arthritis centers at Johns Hopkins (JH), the Hospital for Special Surgery (HSS), and the University of Alabama at Birmingham (UAB) from October 2015 - December 2017. Patients completed PROs on tablets in the waiting room then met with their treating rheumatologist. The study was conducted with central oversight from the Johns Hopkins IRB (IRB_00059930) and IRBs at each site. All participants provided written informed consent.

Participants.

Adults 18+ years of age who met 2010 ACR/EULAR criteria for RA (15) and were fluent in English were eligible to participate. Exclusion criteria were having other forms of inflammatory arthritis or medical or psychiatric problems (e.g., receiving treatment for cancer or severe depression) that would preclude participation.

Outcomes

Patient-reported outcomes (PROs).—PROMIS SFs included Physical Function (20a, v1.0), Pain Interference (8a, v1.0), Fatigue (both 7a and 8a, v1.0), Ability to Participate in Social Roles and Activities (8a, v2.0, "Participation"), and the Adult Profile-29 (v2.0) (www.healthmeasures.net). Patients also completed a pain numeric rating scale (NRS; 0–10) where higher values reflect more pain. The Patient Global Assessment, an 11-point NRS ("Considering all the ways arthritis affects you, how well are you doing today": 0=very well to 10=very poorly)(16) and an RA change rating ("Compared to your last visit would you say that your arthritis is: *a lot better, a little better, the same, a little worse, a lot worse*").[21] Symptom-specific change ratings were also obtained for Physical Function, Pain Interference, Fatigue, and Participation ("Compared to your last visit, would you say your [symptom] is: *a lot better, a little better, the same, a little worse, a lot worse*").

Physician-reported outcomes (CLIN-ROs).—Physicians provided a count of the number of swollen and tender joints (from a total of 28), an MD Global Assessment (0–10) with higher scores representing higher levels of RA disease activity, and an RA Transition

Question ("Compared to the last visit would you say that your patient's arthritis is: *a lot better*, *a little better*, *the same*, *a little worse*, *a lot worse*").

The Clinical Disease Activity Index (CDAI)(17), a composite index of disease activity was calculated for each visit. CDAI is widely used to classify patients by RA disease activity level (remission 2.8, low 2.8- 10, moderate 10- 22, and high >22) which guide treatment decisions.

Responsiveness.—We examined responsiveness based on *a priori* hypothesized changes in PROMIS and legacy scores.(18, 19) We used Cohen's descriptors where r of |0.5 to 1.0|, |0.3 to 0.5|, and |0.1 to 0.3| are suggestive of strong, moderate, and weak relationships, respectively.(20) We hypothesized the following relationships with PROMIS SFs would be observed: 1) mean scores on SFs and corresponding PROMIS-29 domains will be strongly and positively correlated at the first visit; 2a) mean differences in Pain Interference and Pain NRS will be moderately-to-highly correlated; 2b) mean differences in Physical Function and Pain NRS will be moderately-to-highly and inversely correlated; 3a) mean differences in Pain Interference, MD Global, and CDAI will be very weakly-to-weakly correlated; 3b) mean difference in Physical Function, MD Global, and CDAI will be very weakly-to-weakly and inversely correlated; 4a) mean differences in Pain Interference, Fatigue, and Patient Global will be moderately-to-strongly correlated; 4b) mean differences in Participation and Patient Global will be moderately-to-strongly and inversely correlated; 5a) mean differences in Fatigue, MD Global, and CDAI will be very weakly-to-weakly correlated; 5b) mean differences in Participation, MD Global, and CDAI will be very weakly-to-weakly and inversely correlated.

Statistical Methods

Descriptive statistics were calculated to summarize patient characteristics and study outcomes. We calculated Pearson correlations to assess relationships among PROMIS SF versions.

Minimal and meaningful change.—To identify MIDs, we calculated the mean group differences for each domain at the second visit using patient (RA and symptom) and physician (RA) ratings of *a little better* or *a little worse* at the second visit (6). To identify MCIDs, similar methods were used based on patient and physician ratings of *a lot better* or *a lot worse*. We also calculated mean group differences in traditional RA clinical indicators (joint counts, MD Global, Patient Global, Pain NRS, and CDAI). We also identified the mean group differences associated with 0.2 (2-point) and 0.5 (5-point) SD change, and calculated the SEM to identify scores changes exceeding measurement error.(6, 7, 19, 21) We generated empirical cumulative distribution function (eCDF) curves to visually examine discrimination among patient-reported categories for symptoms and RA change categories.

Analyses were completed using SPSS (IBM SPSS Statistics, V26.0). eCDF curves were constructed using Stata version 16.0 (StataCorp LLC).

Results

Participants

Participants were 348 middle-aged, mostly (81%) female adults who had been diagnosed with RA an average of 14 years earlier. Participants were diverse with respect to race, ethnicity, education, RA duration, and current disease activity (Table 1). Baseline demographic and sociodemographic characteristics of the analytic sample (i.e., participants with at least 1 follow up visit, n=284) were similar to the overall cohort (data not shown). At baseline, most were in CDAI LDA (56%) or MDA (30%); 6% were in remission and 8% in HDA. The mean (SD) time between visits was 4.6 (2.4) months and 90 (26%) participants reported a change in RA medications at visit 2.

PROMIS Scores

Mean PROMIS scores at both visits are shown in Supplementary Table 1. At visit 1, participants reported impairments in Physical Function and Pain Interference. Fatigue, Participation, Anxiety, Depression, and Sleep Disturbance scores were within 0.5 SD of the population normative range (e.g., 45 to 55).

Distribution-based Assessments

Using a cut point of 0.2 SD (2 points), 35–73% reported a MID in PROMIS scores (Supplementary Table 1); using a cut point of 0.5 SD (5 points), 14–35% reported a MCID. SEM were similar across domains and averaged 1–2 points (Supplementary Table 2); the highest SEMs (i.e., lowest precision) occurred in patients in either remission or high disease activity.

Patient Ratings: Minimal and Meaningful Change

Change in specific symptoms.—Mean group differences in PROMIS scores by symptom-specific change categories are shown in Table 2 and Supplementary Figure 1. Although PROMIS scores changed in the expected directions, the mean differences varied across domains, SFs, and by direction (worsening vs. improvement). When patients rated their Pain Interference and Fatigue *a lot better*, scores were about 6–7 points lower, while *a lot worse* was associated with mean differences of 3–6 points. A large change was also seen when Participation was rated as *a lot worse*, where scores dropped 6 points. Conversely, scores generally increased 2–4 points when patients rated Participation and Physical Function *a little better* or *a lot better*. eCDF curves by symptom change categories are shown in Figure 1.

Change in RA.—At the second visit, 42% of patients rated their RA as the same at the second visit, 28% reported improvement, and 30% worsening; individuals reporting their RA was worse since the last visit were more likely to also report their medication had been changed at visit 2 (p<.02). Using RA transition categories, changes in PROMIS scores were lower than those seen using symptom-specific anchors (Table 3 and Supplementary Figure 2). Generally, patients who reported their RA was *the same* had stable PROMIS scores (i.e., mean difference of |1| point). Participants who said they were *a little better* generally had a mean difference of 2–3 points. Patients who reported their RA was *a little worse* had a mean

difference of 1–3 points. When participants rated their RA as *a lot worse*, the largest changes were for Pain Interference (>4 points), Physical Function and Pain Interference (2–3 points). When patients rated their RA *a lot better*, Pain Interference improved   4 points. Notably, when patients rated their RA as *a lot better* or *a lot worse*, mean anxiety, depression, and sleep scores were stable and generally within 2 points of visit 1 scores. eCDF curves by patient-reported RA change categories are shown in Figure 2 and Supplementary Figures 3 and 4.

Physician Ratings: Minimal and Meaningful Change

Scores changes in relation to MD RA transition categories are shown in Table 4 and Supplementary Figure 5. At the second visit, patients and rheumatologists rated a similar proportion of patients as *worsening;* physicians were more likely to rate patients as *the same* or *a little better* (whereas patients were more likely to rate their RA as *a lot better*.

The largest differences occurred when Pain Interference and Fatigue were rated as *a lot better* and Fatigue and Participation were rated *a lot worse* (3 points). Mean differences were generally   2 points when physicians rated their patients *a little better* and <1 points for *a little worse,* except the 4-item Fatigue scale where mean scores increased nearly 3 points. As seen with patient ratings of change, Anxiety, Depression, and Sleep scores were mostly stable when physicians rated RA as improving or worsening, except when RA was *a lot worse* for depression (3 points) and *a lot better* for sleep impairment (−3.6 points).

Responsiveness

Overall, hypothesized changes in PROMIS and legacy scores for Physical Function, Pain Interference, Fatigue, and Participation were supported (Supplementary Table 3). Scores on different SF versions assessing the same domain were highly correlated, except for the 20- and 4- item versions of Physical Function, which were moderately associated. Changes in Pain Interference, Physical Function, Fatigue, and Participation were moderately-strongly correlated with changes in Pain NRS and Patient Global, but weakly correlated with MD Global and CDAI.

Discussion

This is the first study to compare minimal and meaningful change scores for PROMIS Physical Function, Pain Interference, Fatigue, and Participation SFs and the PROMIS-29 adult profile from multiple perspectives and using several methods in a prospective observational RA cohort. We also compared the responsiveness of several versions of PROMIS SFs within domains. When using patient anchors, which are widely viewed as most relevant (6), we found that 1–3 points was generally the change in PROMIS scores when patients reported they were *a little better* or *a little worse*. Meaningful change (MCIDs) was associated with at least 3–5 point change for worsening Fatigue, Pain Interference, and Participation, and Physical Function. These findings extend earlier work supporting the validity and reliability of PROMIS measures in RA. (11–14)

Assessing pain and function is recommended in the ACR RA core set of measures for clinical trials and as part of ongoing care.(22, 23) Fatigue assessment also has been

recommended (24) and is important when assessing disease flares.(3, 24) Participation provides information about active engagement in meaningful life situations. Minimal change scores are often used to evaluate new therapies and for comparative effectiveness studies. Meaningful change is often used to establish that a new treatment is working; when RA inflammation is adequately controlled, a change in treatment may be needed. Overall, triangulating meaningful change scores among patient, clinician, RA indicators and score distributions yielded similar results, mean score differences were larger for patient MIDs and MCIDs than physicians. Meaningful change estimates were largest for worsening Participation (6 points), and improved Pain Interference and Fatigue (6 points) when patients used symptom-specific anchors. Importantly, patient estimates of meaningful change were about twice as large when we asked how their symptom had changes vs. asking how their RA had changed. This suggests that using changes in disease to derive MIDs and MCIDs may not be generalizable to specific symptoms as they may yield conservative estimates, which in turn would overestimate the proportion of individuals experiencing meaningful change. Interestingly, the convention of using a 0.5 SD change (5 points on PROMIS scales) in the absence of empirical evidence of responsiveness most closely approximated meaningful change from the patient's perspective using symptom-specific anchors.

As expected, MIDs and MCIDs were generally similar among SF versions, and across perspectives and methods. Notably, MIDs and MCIDs were similar between the 4-item SFs incorporated in the PROMIS-29 Profile as compared with longer SFs. We have previously shown that the seven domains captured in the PROMIS-29 Profile address the major determinants of HRQL in people with RA.(3) Thus, our results support the use of PROMIS-29 as a screening tool to monitor HRQL in people with RA and other chronic diseases over time.(25) Although more items improve the precision of point estimates, particularly in those with low and high disease activity, increased measurement precision offered by the longer length profiles and short forms profiles must be weighed against concerns regarding feasibility and burden to respondents.

This is also the first study we are aware of that graphically examines PROMIS in relation to patient-reported changes in RA and specific symptoms using eCDF curves. Generally, the curves of patients who said they were *a lot better* or *a lot worse* were distinct from those reporting they were *a little better or worse* or *the same*. Using symptom-specific ratings also resulted in better discrimination as compared asking RA change between visits. Within PROMIS domains, longer forms generally had better separation of curves compared with briefer SFs with fewer items.

Fatigue is common in RA and is often one of the first indicators of an impending flare.(26) We compared the responsiveness, MIDs and MCIDs of 7- and 8-item PROMIS Fatigue measures. Although these two SFs are distinct (i.e., contain non-overlapping items), both have evidence of content and construct validity in RA.(12, 13) The 7- and 8-item SFs also had similar MIDs and MCIDs although mean differences were numerically larger for the 8-item version. eCDF curves also suggested both the versions could discriminate minimal from meaningful change, lending additional support to their utility to monitor fatigue over time in RA.

To examine responsiveness, we also evaluated changes in traditional RA clinical indicators and observed that changes in CDAI and joint counts were largest when using the physician change categories. This is not surprising since clinicians often base treatment decisions on the results of joint counts and their global impression of disease activity, 3 of 4 components of the CDAI. When physicians rated patients as "*a lot better*" CDAI dropped –9 points; similarly, when patients were rated as "*a lot worse*", CDAI increased 13 points. Notably, the mean difference in CDAI we observed was larger than MCIDs previously recommended (i.e., –6 points when starting with moderate disease activity, and 2 points for worsening when starting in remission/low disease activity).(27) The change in CDAI when using patient reports of feeling "*a lot better*" was –2.6 points, whereas "*a lot worse*" was 7 points. Others have also noted asymmetry of MCIDs for meaningful improvement vs. worsening. (28) Meaningful improvement or worsening of the Patient Global Assessment and Pain NRS using patient ratings of change averaged about 1 point. However, mean differences on these traditional patient indicators for *a little* vs. *a lot better* or *a little worse* vs. *a lot worse* were similar suggesting that these widely used single item rating scales cannot discriminate between minimal and meaningful change.

Our results are similar with those previously reported for PROMIS MIDs in RA and oncology. Hays et al. reported MIDS of about 2 points for the PROMIS 20-item physician function SF in RA patients.(21) In patients with high disease activity starting a new treatment, changes in PROMIS computer adaptive tests for pain interference and fatigue were 4–5 points, values slightly less than those we observed when patients and rheumatologists said they were *a lot better* for these domains. (29). In people with cancer, MIDs for PROMIS using anchor- and distribution-based methods were: 7-item fatigue (3–5 points); 10-item pain interference (4–6 points); 10-item physical function (4–6 points); and anxiety and depression (3–4.5 points). PROMIS score changes in people with cancer approached or exceeded those we observed when RA patients said they were *a lot better* or *a lot worse* in our study (30).

Strengths of this study include the diversity of participants across geographic, sociodemographic and RA characteristics. Many patients reported a change in their RA at the second visit; these reports were supported by changes in traditional disease activity indicators. We compared results for minimal and meaningful change using patient, clinician, and clinical anchors, and compared estimates using symptom-specific and the more generic (and widely used) RA change categories. We used eCDF curves to visualize scores associated with patient change categories, a technique recently recommended by the US Food and Drug Administration (31).

There are also limitations. We did not evaluate change in response to an intervention shown to be effective (internal responsiveness), although we did note that patients reporting a change in treatment at the second visit were more likely to report worsening PROs. It is notable, however, that our estimates for patients experiencing a lot of improvement were similar to those reported in an RA interventional study for several domains (29). In our study, patients were receiving guideline-based treatment at academic medical centers, and many had limited disease activity and low levels of symptoms. Data were collected at routinely scheduled clinic visits and the time between visits was variable. Participants

enrolled in an observational cohort and were familiar with completing PROs at visits. It is possible that responsiveness would differ in patients with early versus more established disease; however, we did not have large enough numbers of patients experiencing significant changes to examine this.

In summary, from different perspectives and using multiple methods in a real-world cohort of RA patients seen in clinical practice, we evaluated how PROMIS SFs and PROMIS-29 Profile scores changed in relation to changes in specific symptoms, RA status, and traditional disease activity indicators. Study results suggest meaningful change is associated with a 3–7 points change on PROMIS Physical Function, Pain Interference, Fatigue, and Participation SF scores and PROMIS Profile-29 domains. Although the PROMIS-29 includes only 4 items per domain, our results suggest it is sensitive to small changes and could help monitor the key determinants of HRQL in people with RA; longer SFs result in better discrimination. These findings contribute to the growing body of evidence supporting the use of PROMIS SFs in RA care, regulatory decisions, clinical trials, and comparative effectiveness studies. Our results provide a framework for clinicians to use in interpreting scores as PROMIS becomes increasingly incorporated across health care settings. Awareness of the changes in PROMIS T-scores that constitutes meaningful change in RA symptoms and functional impacts can help patients and clinicians decide whether a current treatment is adequately addressing their RA and helping improve HRQL and contribute to shared medical decision-making.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements:

Funding:

References

1. Myasoedova E, Crowson CS, Kremers HM, Therneau TM, Gabriel SE. Is the incidence of rheumatoid arthritis rising?: results from Olmsted County, Minnesota, 1955–2007. Arthritis Rheum 2010;62(6):1576–82. [PubMed: 20191579]

2. Myasoedova E, Davis JM 3rd, Crowson CS, Gabriel SE. Epidemiology of rheumatoid arthritis: rheumatoid arthritis and mortality. Current rheumatology reports. 2010;12(5):379–85. [PubMed: 20645137]

3. Bartlett SJ, Hewlett S, Bingham CO 3rd, Woodworth TG, Alten R, Pohl C, et al. Identifying core domains to assess flare in rheumatoid arthritis: an OMERACT international patient and provider combined Delphi consensus. Ann Rheum Dis 2012;71(11):1855–60. [PubMed: 22772326]

4. Sanderson T, Morris M, Calnan M, Richards P, Hewlett S. Patient perspective of measuring treatment efficacy: the rheumatoid arthritis patient priorities for pharmacologic interventions outcomes. Arthritis care & research. 2010;62(5):647–56.

5. Gossec L, Dougados M, Rincheval N, Balanescu A, Boumpas DT, Canadelo S, et al. Elaboration of the preliminary Rheumatoid Arthritis Impact of Disease (RAID) score: a EULAR initiative. Ann Rheum Dis 2009;68(11):1680–5. [PubMed: 19054825]

6. Revicki D, Hays RD, Cella D, Sloan J. Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. Journal of clinical epidemiology. 2008;61(2):102–9. [PubMed: 18177782]

7. Revicki DA, Cella D, Hays RD, Sloan JA, Lenderking WR, Aaronson NK. Responsiveness and minimal important differences for patient reported outcomes. Health and quality of life outcomes. 2006;4:70. [PubMed: 17005038]

8. Guyatt G, Walter S, Norman G. Measuring change over time: assessing the usefulness of evaluative instruments. Journal of chronic diseases. 1987;40(2):171–8. [PubMed: 3818871]

9. Guyatt GH, Osoba D, Wu AW, Wyrwich KW, Norman GR, Clinical Significance Consensus Meeting G. Methods to explain the clinical significance of health status measures. Mayo Clinic proceedings. 2002;77(4):371–83. [PubMed: 11936935]

10. Cella D, Riley W, Stone A, Rothrock N, Reeve B, Yount S, 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. Journal of clinical epidemiology. 2010;63(11):1179–94. [PubMed: 20685078]

11. Bartlett SJ, Orbai AM, Duncan T, DeLeon E, Ruffing V, Clegg-Smith K, et al. Reliability and Validity of Selected PROMIS Measures in People with Rheumatoid Arthritis. PLoS One. 2015;10(9):e0138543. [PubMed: 26379233]

12. Bingham Iii CO, Gutierrez AK, Butanis A, Bykerk VP, Curtis JR, Leong A, et al. PROMIS Fatigue short forms are reliable and valid in adults with rheumatoid arthritis. J Patient Rep Outcomes. 2019;3(1):14. [PubMed: 30790155]

13. Bartlett SJ, Gutierrez AK, Butanis A, Bykerk VP, Curtis JR, Ginsberg S, et al. Combining online and in-person methods to evaluate the content validity of PROMIS fatigue short forms in rheumatoid arthritis. Quality of life research : an international journal of quality of life aspects of treatment, care and rehabilitation. 2018;27(9):2443–51. [PubMed: 29797175]

14. Bingham CO 3rd, Bartlett SJ, Merkel PA, Mielenz TJ, Pilkonis PA, Edmundson L, et al. Using patient-reported outcomes and PROMIS in research and clinical applications: experiences from the PCORI pilot projects. Quality of life research : an international journal of quality of life aspects of treatment, care and rehabilitation. 2016;25(8):2109–16. [PubMed: 26914103]

15. Aletaha D, Neogi T, Silman AJ, Funovits J, Felson DT, Bingham CO 3rd, et al. 2010 Rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. Arthritis Rheum 2010;62(9):2569–81. [PubMed: 20872595]

16. Nikiphorou E, Radner H, Chatzidionysiou K, Desthieux C, Zabalan C, van Eijk-Hustings Y, et al. Patient global assessment in measuring disease activity in rheumatoid arthritis: a review of the literature. Arthritis research & therapy. 2016;18:251. [PubMed: 27793211]

17. Aletaha D, Smolen JS. The Simplified Disease Activity Index (SDAI) and Clinical Disease Activity Index (CDAI) to monitor patients in standard clinical care. Best practice & research Clinical rheumatology. 2007;21(4):663–75. [PubMed: 17678828]

18. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. Journal of clinical epidemiology. 2010;63(7):737–45. [PubMed: 20494804]

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

19. De Vet HC, Terwee CB, Mokkink LB, Knol DL. Measurement in Medicine. Cambridge, UK: Cambridge Univeristy Press; 2011.

20. Cohen J Statistical power analysis for the behavioural sciences. New York: Academic Press; 1988.

21. Hays RD, Spritzer KL, Fries JF, Krishnan E. Responsiveness and minimally important difference for the Patient-Reported Outcomes Measurement Information System (PROMIS) 20-item physical functioning short form in a prospective observational study of rheumatoid arthritis. Ann Rheum Dis 2015;74(1):104–7. [PubMed: 24095937]

22. Felson DT, Anderson JJ, Boers M, Bombardier C, Chernoff M, Fried B, et al. The American College of Rheumatology preliminary core set of disease activity measures for rheumatoid arthritis clinical trials. The Committee on Outcome Measures in Rheumatoid Arthritis Clinical Trials. Arthritis Rheum 1993;36(6):729–40.

23. Singh JA, Saag KG, Bridges SL Jr, Akl EA, Bannuru RR, Sullivan MC, et al. 2015 American College of Rheumatology Guideline for the Treatment of Rheumatoid Arthritis. Arthritis Rheumatol 2016;68(1):1–26.

24. Kirwan JR, Minnock P, Adebajo A, Bresnihan B, Choy E, de Wit M, et al. Patient perspective: fatigue as a recommended patient centered outcome measure in rheumatoid arthritis. J Rheumatol 2007;34(5):1174–7. [PubMed: 17477482]

25. Cella D, Choi SW, Condon DM, Schalet B, Hays RD, Rothrock NE, et al. PROMIS((R)) Adult Health Profiles: Efficient Short-Form Measures of Seven Health Domains. Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research. 2019;22(5):537–44. [PubMed: 31104731]

26. Hewlett S, Sanderson T, May J, Alten R, Bingham CO 3rd, Cross M, et al. 'I'm hurting, I want to kill myself': rheumatoid arthritis flare is more than a high joint count--an international patient perspective on flare where medical help is sought. Rheumatology (Oxford). 2011;2011 May 12. [Epub ahead of print].

27. Curtis JR, Yang S, Chen L, Pope JE, Keystone EC, Haraoui B, et al. Determining the Minimally Important Difference in the Clinical Disease Activity Index for Improvement and Worsening in Early Rheumatoid Arthritis Patients. Arthritis care & research. 2015;67(10):1345–53. [PubMed: 25988705]

28. Strand V, Singh JA. Newer biological agents in rheumatoid arthritis: impact on health-related quality of life and productivity. Drugs. 2010;70(2):121–45. [PubMed: 20108988]

29. Wohlfahrt A, Bingham CO 3rd, Marder W, Phillips K, Bolster MB, Moreland LW, et al. Responsiveness of Patient-Reported Outcomes Measurement Information System Measures in Rheumatoid Arthritis Patients Starting or Switching a Disease-Modifying Antirheumatic Drug. Arthritis care & research. 2019;71(4):521–9. [PubMed: 29885039]

30. Yost KJ, Eton DT, Garcia SF, Cella D. Minimally important differences were estimated for six Patient-Reported Outcomes Measurement Information System-Cancer scales in advanced-stage cancer patients. Journal of clinical epidemiology. 2011;64(5):507–16. [PubMed: 21447427]

31. U.S. Food and Drug Administration. Discussion Document for Patient-Focused Drug Development Public Workshop on Guidance 3: Select, Develop, or Modify Fit-for-purpose Clinical Outcome Assessments. Methods to Identify What is Important to Patients & Select, Develop or Modify Fit-for-Purpose Clinical Outcomes Assessments [Internet]. Workshop Date October 15–16, 2018. Available from: https://www.fda.gov/media/116277/download.

Significance and Innovation

- PROMIS measures can be used to follow relevant RA symptoms over time.

- 1–3 points generally reflect the change in PROMIS scores when RA patients report feeling 'a little better' or 'a little worse'.

- 3–5 points reflect meaningful worsening of PROMIS Pain Interference, Fatigue, Physical Function, and Participation; whereas, meaningful improvements for Physical Function and Participation were 2–4 points and for Pain Interference and Fatigue were up to 7 points.

- Patient estimates of meaningful change were about twice as large when using symptom-specific vs. generic anchors of change.
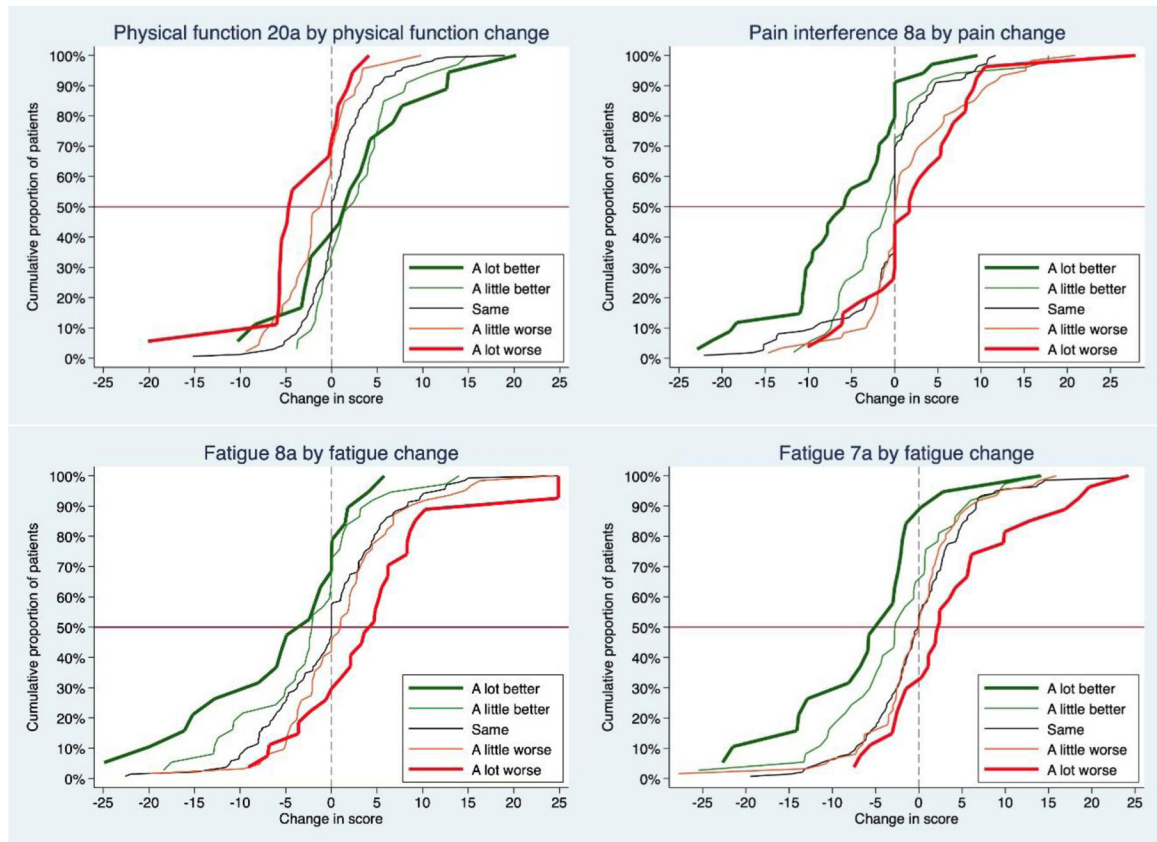
Figure 1.

Empirical cumulative distribution function curves illustrating change in PROMIS scores by patient symptom-specific categories.
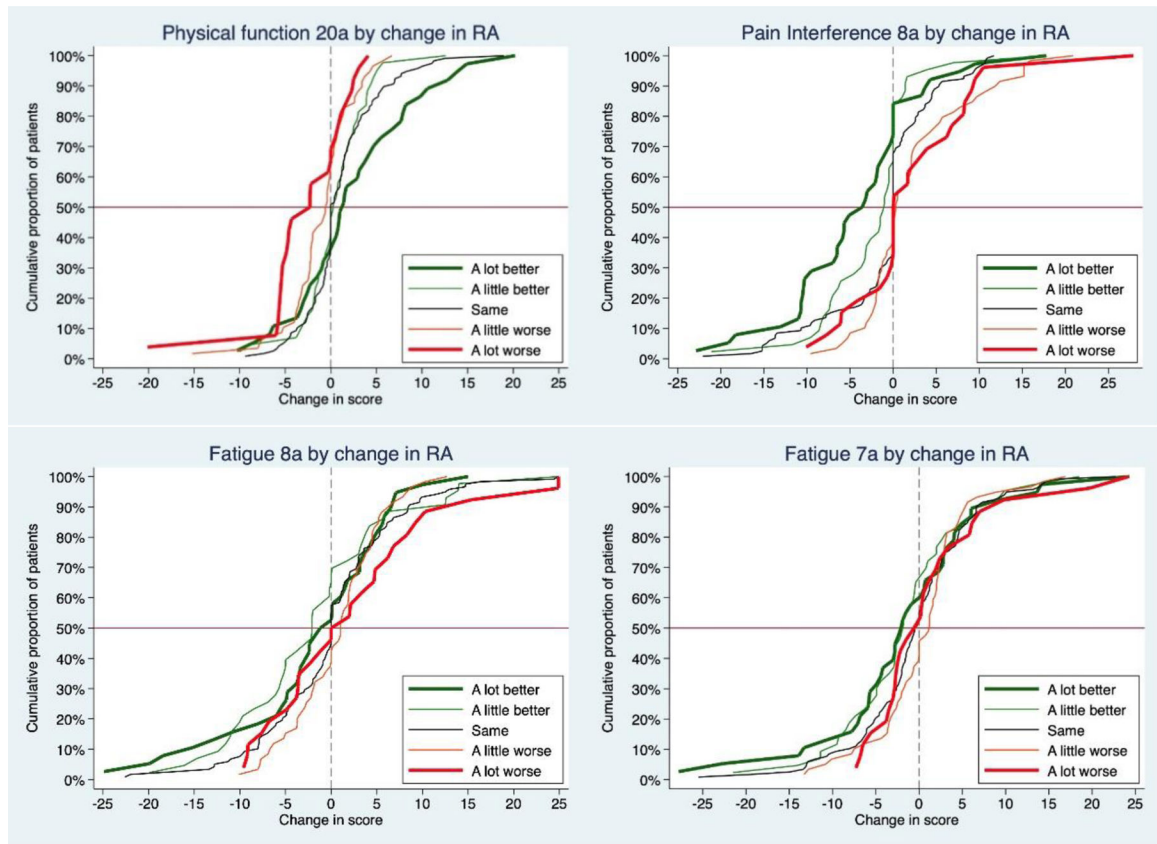
Figure 2.
Empirical cumulative distribution function curves illustrating change in PROMIS scores by RA change categories.

Table 1.

Baseline characteristics of participants (N=348).

Mean (SD) or %	All	Sites		
		HSS N = 82	JH N = 217	UAB N = 49
Age (years)	57 (14)	56 (15)	56 (14)	60 (10)
Female (%)	81%	85%	80%	78%
Race (%)				
American Indian or Native Alaskan	2%	<1%	3%	--
Asian	5%	9%	5%	--
Black	14%	11%	13%	20%
White	77%	71%	79%	80%
Other	2%	2%	2%	--
Declined	2%	6%	<1%	--
Education > High school (%)	79%	91%	76%	73%
Employment (%)				
Full time	36%	44%	35%	27%
Part time	8%	6%	9%	6%
Retired	23%	26%	25%	12%
Disabled because of RA	22%	14%	19%	49%
Other	11%	9%	12%	6%
Urban residence	80%	90%	76%	82%
RA duration (years)	14 (11)	13 (12)	13 (10)	17 (12)
Patient Global Disease Activity (0–100)	30 (27)	27 (23)	32 (28)	28 (27)
Patient Global Disease Activity (0–10)	3.2 (2.6)	2.7 (2.3)	3.2 (2.8)	3.8 (2.5)
Swollen Joints (0–28)	2.5 (3. 5)	2.0 (3.1)	2.8 (3.6)	2.1 (3.5)
Tender Joints (0–28)	2.2 (3.8)	1.4 (3.3)	2.2 (3.5)	3.9 (5.2)
MD Global Assessment (0–10)	1.9 (1.9)	3.0 (2.1)	1.5 (1.5)	2.0 (2.3)
Clinical Disease Activity Index (0–76)	10.3 (7.9)	10.2 (6.9)	10.1 (7.7)	11.6 (10.1)

[*] HSS, Hospital for Special Surgery; JH, Johns Hopkins; UAB, University of Alabama at Birmingham.

Table 2.

Mean group difference in PROMIS scores by symptom change categories at visit 2 as reported by patients.

	A Lot Better			A Little Better			The Same			A Little Worse			A Lot Worse		
	N	%	SD	N	%	SD	N	%	SD	N	%	SD	N	%	SD
Physical Function															
20a	18	6	2.5	7.6	33	12	2.8	4.6	166	59	0.6	4.0	46	17	−1.4	3.8	18	6	−3.4	5.3
4a	18	6	3.9	7.8	33	12	2.5	4.8	166	59	0.1	4.6	45	17	−1.8	4.4	18	6	−2.3	6.5
Pain Interference															
8a	34	12	−6.1	7.3	51	18	−1.1	6.1	113	39	−0.8	5.9	60	21	1.8	6.5	26	10	2.7	7.4
4a	34	12	−6.2	8.3	50	18	−1.1	6.4	112	39	−0.4	6.2	59	21	1.7	6.6	26	10	3.3	7.5
Fatigue															
7a	19	7	−5.9	8.5	37	13	−2.6	7.5	138	49	0.0	6.4	62	22	−0.3	76.5	27	10	4.3	8.3
8a	19	7	−6.1	10.2	37	13	−2.7	7.1	138	49	−0.3	7.1	62	22	1.3	6.9	27	10	5.6	11.0
4a	17	7	−6.8	9.9	37	13	−3.3	8.2	138	49	−0.4	7.4	60	22	1.2	7.2	27	10	5.2	11.5
Participation															
8a	20	7	2.2	8.0	36	13	2.2	4.9	163	58	0.3	5.6	44	16	−1.6	5.3	18	6	−6.1	9.1
4a	20	7	2.7	8.1	35	13	2.2	5.4	163	58	0.1	5.6	43	16	−1.2	5.2	17	6	−6.2	9.2

"Compared to your last visit would you say that your pain/fatigue/physical function/participation is…'a lot better' to 'a lot worse'. % reflects the proportion of total participants in each change category.

Table 3.

Mean group difference in PROMIS scores by RA change categories at visit 2 as reported by patients.

Domain	A Lot Better (13%)			A Little Better (15%)			The Same (42%)			A Little Worse (21%)			A Lot Worse (9%)		
	N		SD	N		SD	N		SD	N		SD	N		SD
PROMIS Short Forms															
Physical Function															
20a	37	2.6	6.7	43	0.5	3.7	118	1.1	4.2	60	−1.3	4.0	26	−2.7	4.8
4a	37	2.4	7.1	43	1.4	4.0	118	0.0	5.1	60	−0.9	4.6	25	−2.0	5.8
Pain Interference															
8a	38	−4.2	7.9	43	−2.3	5.7	118	−0.8	6.2	61	2.3	6.1	26	2.3	7.6
4a	37	−4.0	9.1	43	−2.6	5.7	118	−0.4	6.6	60	2.1	6.0	25	2.8	7.7
Fatigue															
7a	38	−1.6	9.4	43	−1.5	7.4	118	−0.3	7.0	61	1.0	6.2	26	1.3	7.5
8a	38	−2.1	9.4	43	−1.7	9.0	118	0.1	7.6	61	1.0	5.0	26	3.0	11.4
4a	37	−2.6	9.5	43	−2.3	9.0	118	0.1	8.1	60	0.6	5.6	25	2.6	12.0
Participation															
8a	38	2.1	6.3	43	1.5	5.6	118	−0.3	5.6	61	−0.5	6.4	26	−4.5	8.4
4a	37	2.6	6.8	43	1.9	5.5	118	−0.4	5.9	60	−0.6	5.9	25	−4.4	8.1
Depression 4a	37	−1.5	6.4	43	−1.5	6.1	118	−0.1	5.0	60	1.5	6.1	25	1.8	5.2
Anxiety 4a	37	−1.1	7.0	43	−2.4	6.6	118	0.0	7.4	60	1.1	6.4	25	1.3	6.6
Sleep 4a	37	−1.3	8.5	43	−1.6	5.4	118	0.5	5.8	60	−1.6	6.8	25	−0.3	7.4
Clinical Indicators															
CDAI [*]	37	−2.6	8.0	42	−1.9	5.6	113	0.2	5.0	62	2.3	9.4	24	7.0	12.5
Patient Global (0–10)	38	−0.8	1.8	43	−0.7	2.3	117	0.2	2.3	61	0.5	2.1	25	0.8	2.4
Pain (0–10)	37	−0.8	2.4	43	−0.8	2.0	119	−0.2	2.0	62	0.8	1.7	26	1.1	2.4
Physician Global (0–10)	38	−0.7	2.1	43	−0.5	1.2	118	−0.1	1.3	62	0.8	1.8	25	0.9	2.3
Swollen Joints (28)	38	−0.7	3.2	44	−0.8	2.9	119	0.0	3.4	63	0.3	4.2	26	2.3	6.3
Tender Joints (28)	38	−0.8	4.3	44	−0.5	2.5	118	0.0	2.1	63	1.3	5.5	26	2.5	6.7

Values are mean and SD.

[*] Clinical Disease Activity Index

Table 4.

Mean group difference in PROMIS scores by RA change categories at visit 2 as reported by rheumatologists.

	A Lot Better (5%)		A Little Better (21%)		Same (49%)		A Little Worse (18%)		A Lot Worse (7%)	
	N	SD	N	SD	N	SD	N	SD	N	SD
PROMIS Short Forms										
Physical Function										
20a	14 2.4	5.8	59 −0.1	4.3	138 0.8	4.6	49 0.0	5.4	19 −2.0	4.0
4a	14 0.8	4.7	59 0.4	5.0	138 0.1	5.5	49 0.7	5.8	18 −2.1	3.5
Pain Interference										
8a	15 −5.1	5.9	59 −1.9	6.0	138 −0.5	6.8	50 0.9	7.3	19 2.7	6.3
4a	14 −5.4	7.3	59 −1.6	6.5	138 −0.3	6.9	49 0.9	8.2	18 2.9	6.0
Fatigue										
7a	15 −3.1	4.1	59 −2.1	8.0	138 0.5	7.4	50 −0.6	6.5	19 2.6	7.0
8a	15 −3.5	4.6	59 −2.1	8.0	138 0.5	7.7	50 −0.2	7.7	19 4.9	10.9
4a	14 −5.2	4.4	59 −2.8	7.9	138 0.7	8.0	49 −0.6	8.5	18 4.8	11.5
Participation										
8a	15 1.3	4.9	59 1.0	6.8	138 −0.5	6.3	50 0.2	6.4	19 −3.8	5.4
4a	14 1.0	4.9	59 1.4	7.5	138 −0.6	6.0	49 0.4	6.3	18 −3.9	5.7
Anxiety 4a	14 −0.6	6.8	59 −1.2	6.9	138 0.0	7.4	49 0.5	6.4	18 1.0	6.6
Depression 4a	14 −1.5	6.8	59 −2.2	5.6	138 0.6	5.7	49 0.2	4.9	18 3.0	4.9
Sleep 4a	14 −3.6	8.5	59 −0.4	5.4	138 −0.2	6.3	49 −1.4	7.6	18 1.6	6.6
Clinical Indicators										
CDAI [*]	14 −8.9	9.3	59 −2.9	5.6	136 −0.4	4.9	52 5.6	8.9	19 12.9	7.0
Patient Global	15 −0.9	2.1	59 −0.8	2.1	137 0.1	2.2	50 0.2	1.8	18 1.8	2.9
Pain VAS	15 −0.9	1.9	58 −0.8	2.2	140 −0.1	1.9	50 0.5	2.1	19 1.5	2.3
Physician Global	14 −2.4	1.7	61 −0.7	1.5	141 −0.1	1.1	52 1.2	1.6	20 2.2	1.7
Swollen Joints (28)	15 −3.9	4.4	61 −0.9	2.8	144 −0.2	3.2	52 2.0	4.6	19 4.1	4.0
Tender Joints (28)	15 −3.8	4.9	61 −1.1	2.6	143 −0.1	3.0	52 2.3	4.4	19 6.8	4.9

Values are mean and SD.

[*] Clinical Disease Activity Index