UCLA
UCLA Previously Published Works

Title

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

Permalink

https://escholarship.org/uc/item/3362s5ck

Journal

Annals of the Rheumatic Diseases, 74(1)

ISSN

0003-4967

Authors

Hays, Ron D
Spritzer, Karen L
Fries, James F
et al.

Publication Date

2015

DOI

10.1136/annrheumdis-2013-204053

Peer reviewed

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

Ron D Hays, Karen L Spritzer, James F Fries, et al.

Updated information and services can be found at:

http://ard.bmj.com/content/early/2013/10/04/annrheumdis-2013-204053.full.html

These include:

References	This article cites 12 articles, 2 of which can be accessed free at: http://ard.bmj.com/content/early/2013/10/04/annrheumdis-2013-204053.full.html#ref-list-1
P<P	Published online October 4, 2013 in advance of the print journal.
Email alerting service	Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

Advance online articles have been peer reviewed, accepted for publication, edited and typeset, but have not not yet appeared in the paper journal. Advance online articles are citable and establish publication priority; they are indexed by PubMed from initial publication. Citations to Advance online articles must include the digital object identifier (DOIs) and date of initial publication.

To request permissions go to:
http://group.bmj.com/group/rights-licensing/permissions

To order reprints go to:
http://journals.bmj.com/cgi/reprintform

To subscribe to BMJ go to:
http://group.bmj.com/subscribe/

**Topic
Collections**

Articles on similar topics can be found in the following collections

Connective tissue disease (3245 articles)
Degenerative joint disease (3513 articles)
Immunology (including allergy) (3837 articles)
Musculoskeletal syndromes (3768 articles)
Rheumatoid arthritis (2456 articles)

Notes

Advance online articles have been peer reviewed, accepted for publication, edited and typeset, but have not not yet appeared in the paper journal. Advance online articles are citable and establish publication priority; they are indexed by PubMed from initial publication. Citations to Advance online articles must include the digital object identifier (DOIs) and date of initial publication.

To request permissions go to:
http://group.bmj.com/group/rights-licensing/permissions

To order reprints go to:
http://journals.bmj.com/cgi/reprintform

To subscribe to BMJ go to:
http://group.bmj.com/subscribe/

Clinical and epidemiological research

EXTENDED REPORT

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

Ron D Hays,[1,2] Karen L Spritzer,[1] James F Fries,[3] Eswar Krishnan[3]

Handling editor Tore K Kvien

[1]Department of Medicine, UCLA Division of General Internal Medicine & Health Services Research, Los Angeles, California, USA
[2]RAND, Santa Monica, California, USA
[3]Stanford ARAMIS Program, Stanford University School of Medicine, Palo Alto, California, USA

Correspondence to
Professor Ron D Hays, UCLA Department of Medicine, Division of General Internal Medicine & Health Services Research, 911 Broxton Avenue, Los Angeles, CA 90095-1736, USA; drhays@ucla.edu

ABSTRACT

Objective To estimate responsiveness (sensitivity to change) and minimally important difference (MID) for the Patient-Reported Outcomes Measurement Information System (PROMIS) 20-item physical functioning scale (PROMIS PF-20).

Methods The PROMIS PF-20, short form 36 (SF-36) physical functioning scale, and Health Assessment Questionnaire (HAQ) were administered at baseline, and 6 and 12 months later to a sample of 451 persons with rheumatoid arthritis. A retrospective change (anchor) item was administered at the 12-month follow-up. We estimated responsiveness between 12 months and baseline, and between 12 months and 6 months using one-way analysis of variance F-statistics. We estimated the MID for the PROMIS PF-20 using prospective change for people reporting getting 'a little better' or 'a little worse' on the anchor item.

Results F-statistics for prospective change on the PROMIS PF-20, SF-36 and HAQ by the anchor item over 12 and 6 months (in parentheses) were 16.64 (14.98), 12.20 (7.92) and 10.36 (12.90), respectively. The MID for the PROMIS PF-20 was 2 points (about 0.20 of an SD).

Conclusions The PROMIS PF-20 is more responsive than two widely used ('legacy') measures. The MID is a small effect size. The measure can be useful for assessing physical functioning in clinical trials and observational studies.

INTRODUCTION

Physical functioning is an especially important indicator of health for older individuals and one of the strongest predictors of healthcare utilisation and mortality. A physical functioning item bank was created for the Patient-Reported Outcomes Measurement Information System (PROMIS) project,[1] which consists of 124 items assessing mobility (lower extremity), dexterity (upper extremity), axial or central activity (neck and back function), and complex activities that overlap with more than one domain (daily living activities). The items were found to satisfy the item response theory unidimensionality assumption, and item parameters were estimated using a sample of over 21 000 subjects, which included about 1500 patients with rheumatoid arthritis and osteoarthritis.[2–4] The PROMIS physical functioning bank was

shown to have greater precision than existing measures. The PROMIS physical functioning items were recently translated and adapted for use in the Dutch culture.[5]

Item response theory makes it possible to estimate the underlying score using a subset of the items in the full bank. Subsets of the physical functioning items (short forms) can be chosen to minimise response burden. In a cross-sectional study, a 20-item short form was selected from the 'best' PROMIS items[3] which yielded more information (precise measurement) than the short form 36 (SF-36) physical functioning scale and Health Assessment Questionnaire (HAQ). But information about responsiveness (sensitivity to change), an important indicator of validity, for the PROMIS 20-item physical functioning measure has not yet been reported. Rheumatoid arthritis is a progressive disease and physical function tends to decline over time.

A responsive measure is sensitive to improvements, deteriorations and stability of health status over time.[6][7] This paper evaluates the responsiveness of the PROMIS 20-item physical functioning scale (PROMIS PF-20) in a prospective observational cohort of people with rheumatoid arthritis.

METHODS
Data sources and measures
Participants
A total of 451 patients participating in the Arthritis, Rheumatism and Aging Medical Information Systems (ARAMIS) cohorts during 2000–2002 accepted our invitation to participate in this study. There were no specific inclusion or exclusion criteria. ARAMIS is a multicentre longitudinal observational study in the USA that has been following patients who meet the American College of Rheumatology classification criteria.[8][9] These patients were followed over a year using semi-annual surveys. The study was approved by the Stanford University Institutional Review Board (IRB-17334).

An observational study of patients follows them as they receive whatever treatment their healthcare providers implement. Responsiveness can be estimated in this sort of study as long as there are enough subjects who get worse, stay the same, and get better over time.

Instruments

Physical functioning is a subdomain of physical health, which is in turn a subdomain of general health (http://www.nihpromis.org). The PROMIS definition of physical function is the ability to perform basic and instrumental activities of daily living. The PROMIS physical functioning items assess ability to perform, not whether or not an activity actually has been performed (box 1). The items assess capability and use the present tense and avoid attribution to disease or other limiting context. The PROMIS item bank assesses the latent trait of physical functioning ability.

Study participants were administered 19 of the 20 PROMIS PF-20 items. 'Are you able to wash your back' (item 4 in box 1) was not administered because of overlap with other similar items administered in the study. The correlation between scores estimated from the 19 items with the PROMIS PF-20 in the PROMIS wave 1 dataset was 0.998 (n=14 600).

The first 14 items shown in box 1 were administered with five response options: 'without any difficulty', 'with a little difficulty', 'with some difficulty', 'with much difficulty' and 'unable to do'. The last six items were administered using five other response options: 'not at all,' 'very little,' 'somewhat,' 'quite a bit,' and 'cannot do'.

In addition to the PROMIS PF-20, widely used self-report measures of physical functioning ('legacy' measures) were also administered to provide comparative information. These instruments were the 20-item HAQ[10] and the 10-item SF-36 physical functioning scale.[11]

Box 1   Content of PROMIS PF-20

1. Are you able to do chores such as vacuuming or yard work?
2. Are you able to push open a heavy door?
3. Are you able to dress yourself, including tying shoelaces and doing buttons?
4. Are you able to wash your back?
5. Are you able to dry your back with a towel?
6. Are you able to sit on the edge of a bed?
7. Are you able to wash and dry your body?
8. Are you able to get in and out of a car?
9. Are you able to squeeze a new tube of toothpaste?
10. Are you able to hold a plate full of food?
11. Are you able to run a short distance, such as to catch a bus?
12. Are you able to shampoo your hair?
13. Are you able to get on and off the toilet?
14. Are you able to transfer from a bed to a chair and back?
15. Does your health now limit you in doing vigorous activities, such as running, lifting heavy objects, participating in strenuous sports?
16. Does your health now limit you in bending, kneeling or stooping?
17. Does your health now limit you in lifting or carrying groceries?
18. Does your health now limit you in doing 2 hours of physical labor?
19. Does your health now limit you in walking more than a mile?
20. Does your health now limit you in climbing one flight of stairs?

An 'anchor' item was administered on both the two follow-up surveys: 'We would like to know about any changes in how you are feeling now compared with how you were feeling 6 months ago. How has your ability to carry out your everyday physical activities such as walking, climbing stairs, carrying groceries, or moving a chair got a lot better, got a little better, stayed the same, got a little worse, or got a lot worse?'

Administration

The PROMIS PF-20 and two legacy measures were self-administered at baseline and 6 and 12 months after baseline. As noted above, the anchor item was included in the 6-month and 12-month follow-up surveys. Surveys were administered by mail with three rounds of follow-up which included postcard and telephone reminders and multiple mailings of the survey. The attrition rate over the 1-year course of the study was 13%.

Statistical analysis

The anchor item at the 12-month follow-up assessment was used to categorise study participants into five retrospective ratings of change groups: 'lot better', 'little better', 'same', 'little worse' and 'lot worse'. Because the anchor referred to change over the last 6 months, we estimated change on the PROMIS PF-20 between the 6-month and 12-month post-baseline assessments according to change on the anchor. In addition, we examined change from baseline to the 12-month post-baseline assessment to see if there was consistency in responsiveness over a longer time period. This anchor was the independent variable in analyses of variance in which the PROMIS PF-20, SF-36 physical functioning scale and HAQ were dependent variables.

We computed correlations (product–moment and Spearman) between change on the PROMIS PF-20 and the anchor item. F-statistics from one-way analyses of variance were used as indicators of responsiveness.[12 13] In addition, we estimated the minimally important difference on the PROMIS PF-20 by looking at prospective change for the two subgroups that reported on the retrospective anchor item getting 'a little better' or 'a little worse'. Duncan multiple range tests were performed to identify when prospective change on the PROMIS PF-20 differed significantly by retrospectively reported change group. Finally, we conducted a sensitivity analysis for responsiveness by collapsing the 'a little' and 'a lot' categories so that the anchor items had three categories and computed F-statistics for the three physical functioning scales.

RESULTS

Forty-nine per cent of the sample reported an age of 64 years or younger, with 15% being 65–69, and 36% 70 or older; 81% were female; 87% were white; median educational level was 14 years (range 2–18); 6% were current smokers; median body mass index was 26.

Table 1 presents correlations among the PROMIS PF-20, SF-36 and HAQ physical functioning scales at baseline. Also provided are the means, SDs and range of scores. All three scales were strongly associated with one another; the HAQ was somewhat more strongly related to the PROMIS PF-20 than was the SF-36.

On the retrospective rating of change (anchor) item at the 12-month assessment, 21 people reported being 'a lot better', 35 'a little better', 252 'the same', 113 'a little worse' and 30 'a lot worse'. Product–moment (Spearman) correlations for prospective change with the anchor item were 0.35 (0.33) at 12 months and 0.34 (0.33) at 6 months for the PROMIS PF-20, 0.29 (0.32) at 12 months and 0.22 (0.26) for the SF-36 physical

Table 1 Correlations among physical functioning scales and descriptive statistics at baseline

	PF-20	SF-36	HAQ	Mean	SD	Minimum	Maximum
PF-20	1.00	0.84	−0.89	40.18	9.03	12.60	62.30
SF-36	0.84	1.00	−0.79	37.29	12.21	14.94	57.03
HAQ	−0.89	−0.79	1.00	0.88	0.71	0.00	3.00

HAQ, Health Assessment Questionnaire.

Table 3 Change on SF-36 physical functioning scale by self-reported retrospective rating of change

Interval (months)	Lot better (n=21)	Little better (n=35)	Same (n=252)	Little worse (n=113)	Lot worse (n=30)
12	4.99[a]	0.32[b]	0.46[b]	−3.86[c]	−4.74[c]
6	4.08[a]	−0.58[b,c]	0.89[b]	−2.34[c]	−3.47[c]

Cell entries in the same row that share a superscript letter do not differ significantly (p>0.05) from one another (Duncan's multiple range tests). SD of change was 7.74 for 12 months and 7.08 for 6 months.

functioning scale, and 0.29 (0.25) at 12 months and 0.29 (0.25) at 6 months for the HAQ.

Tables 2–4 show prospective change estimates for the PROMIS PF-20, SF-36 physical functioning scale and HAQ, respectively, by the retrospective anchor item for the 12-month and 6-month time intervals. F-statistics for prospective change in the PROMIS PF-20, SF-36 and HAQ physical functioning measures by the retrospective change item over 12 months were 16.64, 12.20 and 10.36, respectively (all p values< 0.0001). F-statistics for the 6-month change were 14.98, 7.92 and 12.90, respectively (all p values <0.0001).

For the three-category version of the anchor item (collapsing the 'a little' and 'a lot' response categories), F-statistics for prospective change in the PROMIS PF-20, SF-36 and HAQ physical functioning measures by the retrospective change item over 12 months were 30.71, 21.43 and 15.66, respectively (all p values <0.0001). F-statistics for the 6-month change were 23.54, 12.49, and 13.47, respectively (all p values <0.0001).

The estimates in table 2 show that the change on the PROMIS PF-20 at 12 months for those who were 'a lot better' on the anchor was significantly different from those reporting they were 'the same', 'a little worse', or 'a lot worse' on the anchor. In addition, those who reported they were a 'little worse' on the anchor differed significantly from those who reported they were 'the same' and those who were 'a lot worse'. Similar results were found for change at 6 months.

A change of about 2 points on a T-score metric (SD=10) is associated with a report of getting 'a little better' or 'a little worse', but change over 6 months for those reporting that they got 'a little worse' was about 1 point. Hence, the estimated minimally important difference for the PROMIS PF-20 appears to be about 0.20 (small effect size) of the baseline SD.

DISCUSSION
The American College of Rheumatology and other professional organisations have recommended that functional status in patients with rheumatoid arthritis be assessed at least annually to systematically identify patients not doing well and to benchmark physician performance. The PROMIS project was initiated to improve precision and the validity of health outcome measures. Previous analyses provided support for the greater precision of measurement of the PROMIS physical functioning measures compared with legacy measures.[3] This study provides support for the construct validity (responsiveness) of the PROMIS PF-20 compared with the SF-36 physical functioning scale and the HAQ. The PROMIS measures were also designed to minimise response burden. The PROMIS PF-20 is estimated to take about 5 min (using the Hays & Reeve[14] rule of thumb of 3–5 items per minute) to administer. We recommend that the PROMIS PF-20 be considered for this assessment and as an end point in studies of rheumatoid arthritis. Standard item parameters can be used to score the PROMIS PF-20 (see http://www.assessmentcenter.net) using 'response pattern scoring'. Raw score to T-score conversion tables are available at https://www.assessmentcenter.net/documents/PROMIS%20Physical%20Function%20Scoring%20Manual.pdf.

Bio-similar drugs for rheumatoid arthritis are expected to enter the market in the next few years. The regulatory pathway for approval of these drugs will involve performance of non-inferiority trials against the existing products. This study suggests that a change in the PROMIS PF-20 of 2 or more may be a minimally important difference for such trials. Although a change of two points does not necessarily warrant changes in therapy, the clinician can be confident that a change of this magnitude is non-trivial.

Future work is needed to evaluate the performance of the PROMIS PF-20 in additional samples and with other external anchors of change. In addition, research is needed to evaluate the extent to which the results reported here generalise to the full PROMIS item bank, computer-adaptive short form testing administration, and other static measures developed from it such as the 10-item PROMIS physical function

Table 2 Change on PROMIS PF-20 by self-reported retrospective rating of change

Interval (months)	Lot better (n=21)	Little better (n=35)	Same (n=252)	Little worse (n=113)	Lot worse (n=30)
12	1.94[a]	1.63[a,b]	0.27[b]	−1.68[c]	−3.20[d]
6	3.26[a]	1.96[a,b]	0.43[b,c]	−0.82[c]	−3.16[d]

Cell entries in the same row that share a superscript letter do not differ significantly (p>0.05) from one another (Duncan's multiple range tests). SD of change was 3.66 for 12 months and 3.76 for 6 months.

Table 4 Change on HAQ by self-reported retrospective rating of change

Interval (months)	Lot better (n=21)	Little better (n=35)	Same (n=252)	Little worse (n=113)	Lot worse (n=30)
12	−0.19[a]	−0.04[b]	0.02[b,c]	0.12[c]	0.28[d]
6	−0.29[a]	−0.08[b]	0.01[b]	0.03[b]	0.25[c]

Cell entries in the same row that share a superscript letter do not differ significantly (p>0.05) from one another (Duncan's multiple range tests). SD of change was 0.32 for 12 months and 0.28 for 6 months. HAQ, Health Assessment Questionnaire.

REFERENCES

1 Cella D, Riley W, Stone A, *et al*. The Patient-Reported Outcomes Measurement Information System (PROMIS) developed and tested its first wave of adult self-reported health outcome item banks: 2005–2008. *J Clin Epidemiol* 2010;63:1179–94.
2 Bruce B, Fries JF, Ambrosini D, *et al*. Better assessment of physical function: item improvement is neglected but essential. *Arthritis Res Ther* 2009;11:R191.
3 Fries JF, Cella D, Rose M, *et al*. Progress in assessing physical function in arthritis: PROMIS short forms and computerized adaptive testing. *J Rheumatol* 2009;36:2061–6.
4 Rose M, Bjorner JB, Becker J, *et al*. 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.
5 Voshaar MAH, ten Kooster PM, Taal E, *et al*. Dutch translation and cross-cultural adaptation of the PROMIS physical function item bank and cognitive pre-test in Dutch arthritis patients. *Arthritis Res Ther* 2012;14:R47.
6 Hays RD, Hadorn D. Responsiveness to change: an aspect of validity, not a separate dimension. *Qual Life Res* 1992;1:73–5.
7 Khanna D, Furst DE, Clements PJ, *et al*. Responsiveness of the SF-36 and the Health Assessment Questionnaire Disability Index (HAQ-DI) in a systemic sclerosis clinical trial. *J Rheumatol* 2005;32:832–40.
8 Arnett FC, Edworth SM, Blooch DA, *et al*. The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. *Arthritis Rheum* 1988;31:315–24.
9 Bruce B, Fries JF. The Arthritis, Rheumatism and Aging Medical Information System (ARAMIS): Still young at 30 years. *Clin Exp Rheumatol* 2005;5(Suppl 39):S163–7.
10 Fries JF, Spitz PW, Young DY. The dimensions of health outcomes: the Health Assessment Questionnaire, disability and pain scales. *J Rheumatol* 1982;9:789–93.
11 Ware JE, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med Care* 1992; 30:473–83.
12 Deyo RA, Diehr P, Patrick DL. Reproducibility and responsiveness of health status measures. *Controlled Clin Trials* 1991;12:142S–58S.
13 Hays RD, Revicki D. Reliability and validity (including responsiveness). In: Fayers P, Hays RD, eds. *Assessing quality of life in clinical trials: methods and practice*. 2nd edn. Oxford: Oxford University Press, 2005:25–39.
14 Hays RD, Reeve B. Measurement and modeling of health-related quality of life. In: Heggenhougen K, Quah S, eds. *International encyclopedia of public health*. Waltham, MA: Academic Press, 2008:241–51.