Minimally Important Difference of Health Assessment Questionnaire in Psoriatic Arthritis: Relating Thresholds of Improvement in Functional Ability to Patient-rated Importance and Satisfaction

PHILIP J. MEASE, J. MICHAEL WOOLLEY, BOJENA BITMAN, BRIAN C. WANG, DENISE R. GLOBE, and AMITABH SINGH

ABSTRACT. ***Objective***. To evaluate changes in function as measured by Health Assessment Questionnaire Disability Index (HAQ-DI) and the meaningfulness of the changes, in importance and satisfaction, in patients with psoriatic arthritis (PsA).

Methods. HAQ-DI was assessed at baseline and at Weeks 4, 12, and 24 in a randomized double-blind study of 205 patients with active PsA receiving etanercept 25 mg twice weekly or placebo. Concurrently, patients rated the importance of and satisfaction with their change in function on a 7-point scale (1 = not at all important/satisfied; 7 = extremely important/satisfied). Mean HAQ-DI improvement corresponding to ratings of minimally (2–3) or very (6–7) important or satisfied was determined using a posthoc linear mixed-model analysis. Patient importance ratings were used as an anchor to estimate minimally important difference (MID) for HAQ-DI; distribution-based estimates were also calculated.

Results. A total of 161 patients (69 placebo; 92 etanercept) had ≥ 1 HAQ-DI scores showing improvement from baseline and a corresponding importance or satisfaction rating. HAQ-DI improvements corresponding to importance scale ratings of 2 or 3 were 0.335 (95% CI 0.214, 0.455) and 0.360 (95% CI 0.263, 0.456), respectively, suggesting an MID of about 0.35. HAQ-DI improvements corresponding to satisfaction scale ratings of 2 and 3 were 0.293 (95% CI 0.230, 0.357) and 0.360 (95% CI 0.307, 0.413). For a given change in HAQ-DI, nearly two-thirds of patients indicated a lower rating for satisfaction than for importance. This trial was registered in the ClinicalTrials.gov registry (NCT00317499).

Conclusion. Our study suggests the MID for HAQ-DI in PsA is about 0.35. The results may also provide insight into patient satisfaction with changes in function and expectations for therapy. (First Release Sept 1 2011; J Rheumatol 2011;38:2461–5; doi:10.3899/jrheum.110546)

Key Indexing Terms:
PSORIATIC ARTHRITIS PATIENT SATISFACTION DISABILITY
HEALTH ASSESSMENT QUESTIONNAIRE OUTCOME ASSESSMENT

Psoriatic arthritis (PsA) is a complex and multifaceted chronic inflammatory disease. PsA may manifest with involvement of the peripheral and axial joints, skin, and nails, and enthesitis and dactylitis with variability in disease course and severity of symptoms. The goals of treatment are to alleviate disease signs and symptoms, inhibit structural damage of the joints, and maximize patient function and quality of life[1,2].

In contrast to rheumatoid arthritis (RA), for which disease response criteria are well defined, composite responder indices are still being developed in PsA. GRAPPA (Group for Research and Assessment of Psoriasis and Psoriatic Arthritis) and OMERACT (Outcome Measures in Rheumatology Clinical Trials) are making progress in this area and have recently identified several core domains for evaluating treatment response in patients with PsA[3,4]. One of these core domains is patient functional status. The Health Assessment Questionnaire Disability Index (HAQ-DI)[5] is validated in RA and is an accepted instrument for evaluating function in PsA[4,6]. Loss of function as assessed by HAQ-DI has been associated with a poor prognosis in patients with PsA[1].

From the Swedish Medical Center and University of Washington, Seattle, Washington; Amgen Inc., Thousand Oaks, California; Kforce Clinical Research, Tampa, Florida; and Pfizer Inc., Philadelphia, Pennsylvania, USA.

Supported by Immunex Corporation, a wholly owned subsidiary of Amgen Inc., and by Wyeth, which was acquired by Pfizer Inc. in October 2009.

P.J. Mease, MD, Swedish Medical Center and University of Washington; J.M. Woolley, PhD; B. Bitman, MS, Amgen Inc.; B.C. Wang, MS, Kforce Clinical Research; D.R. Globe, PhD, Amgen Inc.; A. Singh, PhD, Pfizer Inc.

Address correspondence to Dr. P.J. Mease, Swedish Medical Center and University of Washington, 1101 Madison St., Seattle, WA 98104, USA. E-mail: pmease@nwlink.com

Accepted for publication June 29, 2011.

For outcome measures to be of use in the clinic, physicians must have reliable benchmarks for evaluating them. To aid in interpretation, changes in outcome measures can be anchored by clinical markers or by patient perception. A common benchmark for evaluating changes is the minimally important difference (MID), which is generally considered to be the smallest change that has been defined in some way to be clinically important[7,8]. It is also relevant to consider the levels of change that are very important to patients or that minimally or greatly improve patient satisfaction. For measures of functional status such as HAQ-DI, the patients' perspective of the relevance of change may be a key component of the interpretation.

We have previously presented a preliminary estimate of the minimally important difference of HAQ-DI in PsA using data from a 24-week double-blind, randomized, placebo-controlled study of etanercept in patients with active PsA[9]. The goal of the current report is to expand on that analysis to help to interpret changes in the HAQ and the meaningfulness of these changes, in terms of importance and satisfaction, to patients with PsA.

MATERIALS AND METHODS

This trial, registered in ClinicalTrials.gov, identifier NCT00317499, was a double-blind, placebo-controlled, phase 3 study in patients with PsA. The study design and primary clinical and radiographic results have been described[10,11]. Briefly, patients between 18 and 70 years of age were eligible if they had active PsA inadequately responding to therapy. Patients had ≥ 3 swollen joints and ≥ 3 tender joints as well as a qualifying target lesion of stable plaque psoriasis. Patients receiving stable methotrexate (MTX) therapy for at least 2 months were eligible and could continue MTX at a stable dose of ≤ 25 mg/week. Patients were randomly assigned to receive etanercept 25 mg twice weekly as a self-administered subcutaneous injection or matching placebo for 24 weeks.

The study was conducted in accord with the International Conference on Harmonization Good Clinical Practice guidelines. Institutional review boards for each study center approved the study protocol and all patients provided written informed consent before any study-related procedures were performed.

Patient-reported outcomes including the HAQ-DI were measured at baseline and at 4, 12, and 24 weeks[12]. HAQ-DI is a commonly used measure of physical function that is based on patient-reported assessment of ability to perform tasks related to daily living. The HAQ-DI score is a continuous variable that ranges from 0 (no difficulty) to 3 (unable to perform)[5]. Concurrently, patients were asked to rate the importance of their improvement in functioning and their satisfaction with their improvement in functioning, each on a 7-point scale, with text anchors only at 1 (labeled not at all important/satisfied) and 7 (labeled extremely important/satisfied). Patients who had at least 1 HAQ-DI score that indicated improvement in functioning were included in the analysis. The analysis modifies and expands on a previously reported preliminary analysis of HAQ-DI in these patients[9].

Both anchor-based and distribution-based methods[13,14,15,16] were used to assist in interpreting changes in the HAQ-DI scores and to assess MID.

Anchor-based method. For the scales measuring importance/satisfaction with change, since changes scored at 1 point were labeled not at all important/satisfied, patient ratings of 2 to 3 were interpreted as being just above not important/satisfied and were assumed to indicate changes that are minimally important/satisfactory. Ratings of 6 to 7 were interpreted as indicating very important difference. The changes in HAQ-DI scores associated with these levels of importance and satisfaction (anchors) were determined using a repeated measures mixed-model analysis. The covariance structure was modeled as "unstructured," which does not assume any correlation pattern between the repeated measures.

Distribution-based method. Various distribution-based methods have been proposed for estimating MID, including a half standard deviation of the mean change[13], and 1.96 times the standard error of measurement (SEMT)[17]. The SEMT for the HAQ was calculated as described by Beaton, *et al*[8]:

$$\text{SEMT} = \partial_{\text{baseline HAQ}}\sqrt{1 - r_{\text{baseline HAQ}}}$$

where ∂ = standard deviation, r = Cronbach's alpha reliability coefficient
In this calculation, the estimated reliability coefficient was calculated using data from all 20 HAQ items without aids and devices. All statistical analyses were performed using SAS version 9.1 (SAS Institute, Cary, NC, USA).

RESULTS

Demographic data. A total of 205 patients (104 placebo, 101 etanercept) enrolled in the study. Of the enrolled patients, 161 (69 placebo, 92 etanercept) had at least 1 HAQ-DI score that showed improvement from baseline (i.e., a decreased score) and a corresponding response indicating their perception of the level of importance and/or satisfaction associated with the change in function, and thus were included in this analysis.

Baseline demographics and disease characteristics for patients included in the analysis were similar to those of the overall population (Table 1)[18]. Fewer placebo patients were included because of the criterion of HAQ-DI improvement from baseline, and a slightly smaller percentage of patients were men (48% in this analysis vs 51% in the overall population). The mean age of subjects was 47 years and mean disease duration was 9.1 years. The mean HAQ score at baseline was 1.2 (range 0.13–2.88) among these patients compared with a mean HAQ-DI score of 1.1 (range 0.00–2.88) for the overall population.

Changes in HAQ associated with importance and satisfaction. We analyzed the relationship between changes in the HAQ-DI score and patients' perception of the importance of those changes, as indicated on a 7-point scale. The mean change in HAQ-DI that corresponded to each rating on the importance scale as determined from a linear mixed model is shown in Table 2. Patient ratings of minimally important improvement (levels 2 to 3) corresponded to estimated improvements of 0.335 (95% CI 0.214, 0.455) for level 2 and 0.360 (95% CI 0.263, 0.456) for level 3. Combining these estimates leads to an estimate for the MID of 0.348 (the calculated estimate for a change in the importance item of 2.5), which we rounded to 0.35, but the number of responses in the 2 to 3 point range was small (n = 11). Patient ratings indicating very important improvement (levels 6 to 7; n = 291) corresponded to HAQ-DI improvements of 0.435 (95% CI 0.385, 0.485) to 0.460 (95% CI 0.404, 0.515) for 6 and 7-point changes, respectively.

Similarly, we analyzed the relationship between changes in the HAQ-DI score and patient satisfaction with those

Table 1. Baseline demographics and disease characteristics of a subset of patients in a randomized, double-blind, placebo-controlled trial of etanercept in patients with psoriatic arthritis. The table includes patients with at least 1 HAQ-DI score showing improvement from baseline.

Characteristic	Placebo, n = 69	Etanercept, n = 92	Total, n = 161
Sex, n (%)			
Men	28 (41)	50 (54)	78 (48)
Race, n (%)			
White	63 (91)	82 (89)	145 (90)
African American	2 (3)	3 (3)	5 (3)
Hispanic	3 (4)	6 (7)	9 (6)
Other	1 (1)	1 (1)	2 (2)
Age, mean (SD) yrs	46.7 (11.1)	46.9 (11.2)	46.8 (11.1)
Baseline HAQ-DI, mean (range)	1.17 (0.25–2.88)	1.15 (0.13–2.50)	1.16 (0.13–2.88)
Duration of PsA, yrs, mean (range)	9.5 (0.2–35.4)	8.8 (0.0–41.4)	9.1 (0.0–41.4)

HAQ-DI: Health Assessment Questionnaire Disability Index; PsA: psoriatic arthritis.

Table 2. HAQ-DI improvements associated with patient ratings of importance. Linearly predicted HAQ-DI improvement was calculated using a mixed model that accounts for within-patient variation over time by patient response as to level of importance of changes; post-baseline patient responses were collected at 4, 12, and 24 weeks. In the patient rating scale for importance of change in function, 1 = not at all important and 7 = extremely important.

Patient Rating for Importance	No. Ratings	Adjusted Mean Change in HAQ-DI (95% CI)
1	2	0.310 (0.165, 0.455)
2	2	0.335 (0.214, 0.455)
3	9	0.360 (0.263, 0.456)
4	23	0.385 (0.310, 0.460)
5	61	0.410 (0.352, 0.468)
6	81	0.435 (0.385, 0.485)
7	210	0.460 (0.404, 0.515)

HAQ-DI: Health Assessment Questionnaire Disability Index.

Table 3. HAQ-DI improvements associated with patient ratings of satisfaction. Linearly predicted HAQ-DI improvement was calculated using a mixed model that accounts for within-patient variation over time by patient response as to level of satisfaction with changes; post-baseline patient responses were collected at 4, 12, and 24 weeks. In the patient rating scale for importance of change in function, 1 = not at all important and 7 = extremely important.

Patient Rating for Satisfaction	No. Ratings	Adjusted Mean Change in HAQ-DI (95% CI)
1	26	0.227 (0.150, 0.304)
2	32	0.293 (0.230, 0.357)
3	51	0.360 (0.307, 0.413)
4	54	0.426 (0.379, 0.473)
5	68	0.493 (0.444, 0.541)
6	84	0.559 (0.503, 0.615)
7	73	0.625 (0.558, 0.693)

HAQ-DI: Health Assessment Questionnaire Disability Index.

changes (Table 3). Patient satisfaction ratings of 2 to 3 (n = 83 responses) corresponded to mean improvements in HAQ-DI of 0.293 (95% CI 0.230, 0.357) to 0.360 (95% CI 0.307, 0.413). Satisfaction ratings of 6 to 7 (n = 157) corresponded to mean improvements in HAQ-DI of 0.559 to 0.625.

Distribution-based estimates of MID for HAQ-DI were also evaluated and provided values similar to the estimates obtained using the anchor-based method (Table 4). Using a half SD estimate for importance gave a HAQ-DI change value of 0.293, while using 1.96-times the SEMT gave an estimate of 0.266.

Relationship between patient perception of importance and satisfaction. The relationship between the importance patients assigned to their improvement in function and their satisfaction with that improvement was also evaluated (Table 5). As noted above, only 11 of 388 responses to the importance of change in function were scored in the 2 to 3 point range. The vast majority of responses regarding importance of change were in the 6 to 7 point range. Satisfaction ratings associated with change in function were somewhat more evenly distributed across the scale. In 63% of responses, patients rated their level of satisfaction with the change lower than they rated the level of importance of that change.

DISCUSSION

Our study examined patient ratings of importance and satisfaction associated with changes in HAQ-DI over 24 weeks as part of a randomized clinical trial of etanercept in patients with active psoriatic arthritis[11]. The improvement in HAQ-DI score that was associated with minimal importance was 0.35 using anchor-based methods. The small number of subjects reporting change of small importance led to somewhat wide estimated CI around this estimate. These estimates were similar to estimates using distribution-based methods for determining MID. Minimal satisfaction was associated with a similar change in HAQ-DI of about 0.33.

Table 4. Minimal important differences in HAQ-DI as determined by distribution-based methods.

Standard Deviation of Baseline HAQ-DI	Half Standard Deviation	Standard Error of Measurement (SEMT)	1.96 × SEMT
0.586	0.293	0.136	0.266

HAQ-DI: Health Assessment Questionnaire Disability Index.

Table 5. Relationship of patient importance and patient satisfaction ratings. The top number in each table cell is the frequency of each pair of satisfaction/importance ratings; the bottom number is the percentage of all pairs of ratings represented by that satisfaction/importance pair.

	Patient Satisfaction Rating, Frequency Percentage							
	1	2	3	4	5	6	7	Total
Patient Importance Rating, Frequency Percentage								
1	0	1	0	0	0	1	0	2
	0	0.26	0	0	0	0.26	0	0.52
2	1	0	0	0	0	1	0	2
	0.26	0	0	0	0	0.26	0	0.52
3	0	0	6	1	2	0	0	9
	0	0	1.55	0.26	0.52	0	0	2.32
4	1	2	3	13	3	1	0	23
	0.26	0.52	0.77	3.35	0.77	0.26	0	5.93
5	0	6	19	11	13	8	4	61
	0	1.55	4.9	2.84	3.35	2.06	1.03	15.72
6	4	9	9	13	13	21	12	81
	1.03	2.32	2.32	3.35	3.35	5.41	3.09	20.88
7	20	14	14	16	37	52	57	210
	5.15	3.61	3.61	4.12	9.54	13.4	14.69	54.12
Total	26	32	51	54	68	84	73	388
	6.7	8.25	13.14	13.92	17.53	21.65	18.81	100

The MID for HAQ-DI in RA is accepted to be on the order of 0.22[19,20,21] (absolute value), and in ankylosing spondylitis the MID for improvement and worsening in HAQ-DI were recently estimated to be –0.136 and 0.220[22]. We previously published a preliminary report estimating an MID of 0.3 for the HAQ in PsA using patient data from the trial examined here[9]. The original estimate was calculated by using patient ratings of satisfaction with change, rather than importance of change, because of the much larger number of responses available in the 2- to 3-point range of the scale for patient satisfaction compared with importance. Because MID is more appropriately calculated with importance data than with satisfaction data, we believe the updated MID estimate of 0.35 reported here provides a more suitable estimation of MID in patients with PsA. Our value is somewhat higher than a recently reported MID estimate of 0.131 that was based on a group of almost 250 patients with PsA at a single clinic using methods based on an overall health status anchor[23]. Patients in that study had less functional disability at baseline (mean HAQ-DI = 0.732) than the patients in our study (mean HAQ-DI = 1.16), which may have contributed to the difference.

One of the goals of estimating the MID is to define a threshold for responder analyses. With 8 response categories for the HAQ-DI, changes would occur in increments of one-eighth (0.125) of a point. This implies that for responder analyses, changes in the HAQ-DI in the range of 0.250 to 0.375 are in essence the same, and the MID is likely in this range. Therefore responder analyses using a threshold of 0.375 may be appropriate for the HAQ-DI in PsA.

Recently in the literature, researchers have begun to explore the concept of "really important changes" as a complement to understanding minimally important changes. For example, a really important change in HAQ-DI in patients with RA was estimated to be in the range of 0.75 to 0.87 in 1 study[24]. In the present study in patients with PsA, change in HAQ-DI associated with very important improvement in function was about 0.45, a lower value than we expected. However, change in HAQ-DI associated with a very high level of satisfaction was higher, about 0.59. It appears that while patients may find the changes they experience to be important, they are less likely to be satisfied with them.

Our study had several limitations. The number of patient responses rating their functional change as of little importance was small, necessitating the use of a linear model to determine MID. However, given that multiple methods —

anchor-based and distribution-based — identified a similar range for MID, we believe our estimates are meaningful. Also, we examined only ratings associated with improvements in HAQ-DI. There is considerable debate about whether MID should be evaluated using both positive and negative changes in outcome measures. In some cases, MID has been shown to be different for improvements vs declines in function[22]. Thus, our estimates should only be used to evaluate improvements in HAQ-DI. Additionally, MID estimates may depend on baseline status of the patients studied. The patients in this trial had active PsA and a mean baseline HAQ-DI score of 1.1, considered moderate functional impairment, and so the MID may need to be interpreted in this context.

Our study examined thresholds of change in HAQ-DI that corresponded to patient perceptions of the importance of and their satisfaction with those changes. A HAQ-DI improvement of about 0.35 appears to be a minimally important change, while 0.45 could be considered very important. Our study may provide insight into patient perceptions of changes in function and their expectations with regard to therapy.

ACKNOWLEDGMENT

Holly Brenza Zoog, PhD, of Amgen Inc. provided medical writing support for this article.

REFERENCES

1. Ritchlin CT, Kavanaugh A, Gladman DD, Mease PJ, Helliwell P, Boehncke WH, et al. Treatment recommendations for psoriatic arthritis. Ann Rheum Dis 2009;68:1387-94.
2. Gottlieb A, Korman NJ, Gordon KB, Feldman SR, Lebwohl M, Koo JY, et al. Guidelines of care for the management of psoriasis and psoriatic arthritis: Section 2. Psoriatic arthritis: overview and guidelines of care for treatment with an emphasis on the biologics. J Am Acad Dermatol 2008;58:851-64.
3. Gladman DD, Landewe R, McHugh NJ, Fitzgerald O, Thaci D, Coates L, et al. Composite measures in psoriatic arthritis: GRAPPA 2008. J Rheumatol 2010;37:453-61.
4. Mease PJ. Assessment tools in psoriatic arthritis. J Rheumatol 2008;35:1426-30.
5. Stanford University Medical Center. ARAMIS: HAQ. [Internet. Accessed July 22, 2011.] Available from: http://aramis.stanford.edu/HAQ.html
6. Mease PJ, Antoni CE, Gladman DD, Taylor WJ. Psoriatic arthritis assessment tools in clinical trials. Ann Rheum Dis 2005;64 Suppl 2:ii49-54.
7. Beaton DE, Boers M, Wells GA. Many faces of the minimal clinically important difference (MCID): a literature review and directions for future research. Curr Opin Rheumatol 2002; 14:109-14.
8. Beaton DE, Bombardier C, Katz JN, Wright JG, Wells G, Boers M, et al. Looking for important change/differences in studies of responsiveness. OMERACT MCID Working Group. Outcome Measures in Rheumatology. Minimal clinically important difference. J Rheumatol 2001;28:400-5.
9. Mease PJ, Ganguly R, Wanke L, Yu E, Singh A. How much improvement in functional status is considered important by patients with active psoriatic arthritis: applying the Outcome Measures in Rheumatoid Arthritis Clinical Trials (OMERACT) group guidelines [abstract]. Ann Rheum Dis 2004;63 Suppl 1:SAT0015.
10. Mease PJ, Kivitz AJ, Burch FX, Siegel EL, Cohen SB, Ory P, et al. Continued inhibition of radiographic progression in patients with psoriatic arthritis following 2 years of treatment with etanercept. J Rheumatol 2006;33:712-21.
11. Mease PJ, Goffe BS, Metz J, VanderStoep A, Finck B, Burge DJ. Etanercept in the treatment of psoriatic arthritis and psoriasis: a randomised trial. Lancet 2000;356:385-90.
12. Mease PJ, Woolley JM, Singh A, Tsuji W, Dunn M, Chiou C-F. Patient-reported outcomes in a randomized trial of etanercept in psoriatic arthritis. J Rheumatol 2010;37:1221-7.
13. Norman GR, Sloan JA, Wyrwich KW. Interpretation of changes in health-related quality of life: the remarkable universality of half a standard deviation. Med Care 2003;41:582-92.
14. 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;61:102-9.
15. Wyrwich KW, Nienaber NA, Tierney WM, Wolinsky FD. Linking clinical relevance and statistical significance in evaluating intra-individual changes in health-related quality of life. Med Care 1999;37:469-78.
16. Hays RD, Woolley JM. The concept of clinically meaningful difference in health-related quality-of-life research. How meaningful is it? Pharmacoeconomics 2000;18:419-23.
17. Ware JE Jr, Bayliss MS, Rogers WH, Kosinski M, Tarlov AR. Differences in 4-year health outcomes for elderly and poor, chronically ill patients treated in HMO and fee-for-service systems. Results from the Medical Outcomes Study. JAMA 1996; 276:1039-47.
18. Mease PJ, Kivitz AJ, Burch FX, Siegel EL, Cohen SB, Ory P, et al. Etanercept treatment of psoriatic arthritis: safety, efficacy, and effect on disease progression. Arthritis Rheum 2004;50:2264-72.
19. Redelmeier DA, Lorig K. Assessing the clinical importance of symptomatic improvements. An illustration in rheumatology. Arch Intern Med 1993;153:1337-42.
20. Kosinski M, Zhao SZ, Dedhiya S, Osterhaus JT, Ware JE Jr. Determining minimally important changes in generic and disease-specific health-related quality of life questionnaires in clinical trials of rheumatoid arthritis. Arthritis Rheum 2000;43:1478-87.
21. Wells GA, Tugwell P, Kraag GR, Baker PR, Groh J, Redelmeier DA. Minimum important difference between patients with rheumatoid arthritis: the patient's perspective. J Rheumatol 1993;20:557-60.
22. Wheaton L, Pope J. The minimally important difference for patient-reported outcomes in spondyloarthropathies including pain, fatigue, sleep, and Health Assessment Questionnaire. J Rheumatol 2010;37:816-22.
23. Kwok T, Pope JE. Minimally important difference for patient-reported outcomes in psoriatic arthritis: Health Assessment Questionnaire and pain, fatigue, and global visual analog scales. J Rheumatol 2010;37:1024-8.
24. Wolfe F, Michaud K, Strand V. Expanding the definition of clinical differences: From minimally clinically important differences to really important differences. Analyses in 8931 patients with rheumatoid arthritis. J Rheumatol 2005;32:583-9.