Minimally Important Difference for Patient-reported Outcomes in Psoriatic Arthritis: Health Assessment Questionnaire and Pain, Fatigue, and Global Visual Analog Scales

TIFFANY KWOK and JANET E. POPE

ABSTRACT. Objective. To determine the minimally important difference (MID) for the Health Assessment Questionnaire-Disability Index (HAQ-DI), pain, fatigue, sleep, and global visual analog scale (VAS; 0–100 mm) in patients with psoriatic arthritis (PsA) using a patient-reported overall health status anchor. Patient-reported outcomes are often used to gauge the effect of PsA in clinical trials. There is currently no knowledge about the MID for patient-reported outcomes in PsA.

Methods. Patients with a diagnosis of PsA who had answered questions about outcomes at 2 consecutive visits and an overall health status question (“How would you describe your overall status since your last visit: much better, better, the same, worse, much worse?”) were included. MID was calculated as the mean change between visits for those who rated their disease as “better” or “worse.”

Results. Two hundred patients met inclusion criteria, of whom 17.5% rated their status as “better” and 25.0% rated their status as “worse” than the previous visit. MID estimates for improvement/worsening (SD) respectively were −0.131 (0.411)/0.131 (0.309) for HAQ-DI, −9.37 (24.37)/13.96 (22.05) for pain VAS, −8.15 (23.52)/3.63 (27.62) for fatigue VAS, −10.97 (29.74)/13.81 (27.32) for sleep VAS, and −8.41 (21.17)/11.53 (21.03) for global VAS. Spearman’s ρ correlation coefficients between the anchor and mean change were 0.374 (HAQ-DI), 0.448 (pain VAS), 0.239 (fatigue VAS), 0.326 (sleep VAS), 0.490 (global VAS); p < 0.01.

Conclusion. This is the first study investigating MID of patient-reported outcomes in PsA. MID for HAQ-DI, pain, and global VAS were shown to be the best predictors for a patient’s perception of overall changes in disease status. (First Release March 15 2010; J Rheumatol 2010;37:1024–8; doi:10.3899/jrheum.090832)

Key Indexing Terms:
PSORIATIC ARTHRITIS
PATIENT-REPORTED OUTCOMES
MINIMALLY IMPORTANT DIFFERENCE
HEALTH ASSESSMENT QUESTIONNAIRE-DISABILITY INDEX
HEALTH-RELATED QUALITY OF LIFE
Rheumatology Clinical Trials (OMERACT) working group of the Group for Research and Assessment of Psoriasis and Psoriatic Arthritis that should be included in clinical trials and longitudinal observational studies16,17. In addition, fatigue was also a domain that was highly recommended for inclusion in clinical trials16,17.

Because these patient-centered outcomes are used in assessing individuals with PsA, it is important to determine how much of a change in score correlates to patients’ changed perception of their disease in clinical practice. The minimally important difference (MID) is defined as the smallest change in a patient-reported outcome score that patients perceive as a relevant change18. Determination of MID may aid in physicians’ consideration of making changes to a patient’s existing management plan, as well as determination of clinically relevant outcomes and sample sizes in future research19. MID estimates can be derived using an anchor-based approach, employing clinically relevant indicators that are tied to the changes in patient-reported outcome scores20. To date, there are no reports of the MID of any patient-reported outcomes in PsA.

We studied patients with PsA to determine the MID for the HAQ-DI, pain, fatigue, sleep, and global VAS in patients with PsA using a patient-reported overall health status anchor. Our secondary objective was an exploratory analysis of MID estimates of these patient-reported measures stratified by gender, use of biologics, and use of DMARD. We hypothesized that the MID for each patient-reported outcome of interest could be bidirectionally different for improvement and worsening. We expected MID results to be similar to those of arthropathies with similar effects on HRQOL such as rheumatoid arthritis (RA)5-9.

MATERIALS AND METHODS

Multiple data are collected routinely on patients seen at our rheumatology clinic at St. Joseph’s Hospital, affiliated with the University of Western Ontario and serving a referral region of about 1 million. The clinic is heavily weighted toward seeing inflammatory arthritis. Patients are primarily English-speaking from a wide catchment area and represent the full spectrum of disease severity, and are generally followed 1–4 times a year depending on disease severity and flares. All patients with PsA are followed. The University of Western Ontario Ethics Committee approved the study.

Demographic, treatment-related, and patient-reported outcome variables were collected from patient charts of those with a diagnosis of PsA as per billing code from the practice of 1 rheumatologist at our clinic. Treatment-related variables were diagnosis of PsA, disease duration, time between visits, use of prednisone, use of DMARD, and use of biologics. Patient-reported outcome variables collected were HAQ-DI and VAS for pain, fatigue, sleep, and global status. Global status was defined as overall health anchor at the second visit, which asked regarding arthritis, “How would you describe your overall status since your last visit: much better, better, the same, worse, much worse”. Global ratings of change such as the anchor used in our study are well accepted subjective anchors in HRQOL research22.

Patients were included in the study if they had a diagnosis of PsA meeting the Moll and Wright criteria23, were seen in the clinic at least once between January 2007 and October 2008, and had completion of the overall health status anchor, HAQ-DI, and pain, fatigue, sleep, and global VAS at 2 consecutive clinic visits. Exclusion criteria were an absence of PsA diagnosis and lack of appropriate anchor, HAQ, and VAS data.

For each measurement, change scores were calculated as the difference in scores between 2 consecutive visits (visit 2 — visit 1), with a negative change signifying improvement. Calculation of the MID was the mean change of patient-reported outcome measures between visits in those who rated their disease as “better” or “worse” as per the anchor question. These categories represent a minimal change in disease status, as there are also options for patients to rate their disease as “much better” or “much worse.” Thus, a negative value for MID indicates improvement and positive value indicates decline in symptoms. This is consistent with MID determination described in the literature24,25. A Spearman correlation coefficient between the anchor and mean change score of > 0.37 was desired in order to correlate with a large effect size according to Cohen’s rule of thumb26. Autocorrelation measures of the change in scores between visits were performed in order to assess the reliability of the MID estimate. Therefore, we performed correlations between each change compared to an anchor (the anchor was the patient global from last visit, which was much worse, worse, same, better, or much better). The changes of interest were better or worse and the correlation was between the changed patient global and the change of each outcome of interest such as change in HAQ, fatigue, pain, sleep, etc. between visits.

An exploratory analysis was also undertaken to investigate MID for the different patient-reported outcomes as stratified by gender, DMARD use, and biologic use. This may potentially show that individuals of different genders or who undergo different treatment regimens require different amounts of improvement or worsening in order to consider a change to represent a MID. The decision to undergo exploratory analyses for these variables specifically was done only after data were collected.

Data were analyzed using SPSS v. 11.0. A p value of < 0.05 was considered significant. Results were presented as mean (SD) unless specified otherwise.

RESULTS

Two hundred forty-nine patients were identified with a billing code corresponding to PsA between January 2007 and October 2008. Twenty-six did not meet Moll and Wright criteria for PsA, and 23 did not have required data for inclusion in our study. A total of 200 patients were thus included. These patients had a mean age of 51.05 (14.08) and disease duration of 11.21 (7.88) years, and 58.5% were women. Mean time between visits was 8.28 (5.94) months. DMARD use was seen in 64.0%, with 18.5% using biologics and 4.0% using prednisone. In terms of the overall health status anchor, 17.5% reported feeling better and 25.0% reported feeling worse at the second consecutive visit over the time they were studied. In addition, 11.0% said they felt “much better,” 44.5% “same,” and 2.0% felt “much worse.” Descriptive statistics are displayed in Table 1.

Mean baseline, followup, and change scores of HAQ-DI, pain VAS, fatigue VAS, sleep VAS, and global VAS are found in Table 2. Mean change scores for each patient-reported outcome by anchor response are found in
It is important to determine MID for different patient-reported outcomes in order to ascertain how much of a change in score is meaningful to the patient. The MID may vary based on patient demographics, management strategy, severity of disease, and choice of anchors used to evaluate MID. This is the first study to determine MID estimates for patient-reported outcomes in PsA. In our study, MID was estimated for patients from a single rheumatology practice. Patient perception of disease status as compared to the previous visit (patient-reported anchor) was normally distributed, with almost half of our patients noting no change between visits, 17.5% and 25.0% of patients noting improvement and worsening, respectively, between visits, and the remaining few noting significant improvement or worsening. Patients reported low mean HAQ-DI, pain, fatigue, sleep, and global VAS scores overall. Means for these scores were consistent (within 1 SD) with means obtained in previous clinical studies of patients with PsA.

For HAQ-DI, the MID for improvement/worsening had bidirectionally equal magnitudes with no overlap in confidence intervals, at –0.131 (0.411; –0.272, 0.0105) and 0.131 (0.309; 0.0436, 0.219), respectively. This means that the same amount of change in HAQ-DI score is required for patients to subjectively report feeling better or worse. An adequate effect size, with a Spearman correlation coefficient of 0.374, was found. Autocorrelation of the change in scores between visits of –0.046 indicates that the reliability of the MID measure calculated in our study is high. Based on this MID, it is likely that many patients with PsA perceived even a single change on a HAQ-DI scale (change of 0.125) as disease improvement or worsening. Compared to other rheumatologic diseases that have been studied in our research practice, the MID of HAQ-DI in PsA is similar to that found in patients with (improvement/worsening) RA (–0.090/0.150), ankylosing spondylitis (–0.136/0.220), and systemic lupus erythematosus (–0.08/0.14). The reported changes are an average change, so not all patients would have equal perceptions.

All other MID for the various patient-reported health status outcomes showed bidirectional differences with improvement/worsening, and 95% CI were nonoverlapping, except for fatigue VAS. A similar magnitude of improvement and worsening within all scores was needed for patients to perceive a better or worse status of their disease. An adequate effect size was reached for MID of pain VAS and global VAS, with Spearman correlation coefficients of 0.448 and 0.490. This correlates to the core domains identified by OMERACT as the most important areas of PsA to include in clinical trials and observational studies. MID of fatigue and sleep VAS had smaller Spearman correlation coefficients, corresponding to a medium effect size according to Cohen’s rule of thumb. However, all MID estimates were found to be nontrivial (p < 0.01). Perhaps pain and global VAS may be more accurate than fatigue and sleep VAS at detecting small changes in a patient’s perceived sta-

### Table 1. Descriptive characteristics of the study patients.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Mean (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of participants</td>
<td>200</td>
</tr>
<tr>
<td>Sex (% women, men)</td>
<td>58.5, 41.5</td>
</tr>
<tr>
<td>Age, yrs</td>
<td>51.05 (14.08)</td>
</tr>
<tr>
<td>Disease duration, yrs</td>
<td>11.21 (7.88)</td>
</tr>
<tr>
<td>Interval between visits, mo</td>
<td>8.28 (5.94)</td>
</tr>
<tr>
<td>Medication use (% using)</td>
<td>Prednisone: 4.0</td>
</tr>
<tr>
<td></td>
<td>DMARD: 64.0</td>
</tr>
<tr>
<td></td>
<td>Biologics: 18.5</td>
</tr>
</tbody>
</table>

### Table 2. Mean baseline, followup, and change scores of patient-reported outcome measures.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Baseline Mean (SD)</th>
<th>Followup Mean (SD)</th>
<th>Change Mean (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HAQ-DI</td>
<td>0.732 (0.677)</td>
<td>0.711 (0.707)</td>
<td>–0.0211 (0.369)</td>
</tr>
<tr>
<td>Pain VAS</td>
<td>41.45 (27.69)</td>
<td>38.65 (28.84)</td>
<td>–3.29 (26.19)</td>
</tr>
<tr>
<td>Fatigue VAS</td>
<td>40.82 (31.68)</td>
<td>38.30 (30.42)</td>
<td>–2.91 (24.37)</td>
</tr>
<tr>
<td>Sleep VAS</td>
<td>37.99 (32.93)</td>
<td>38.83 (32.32)</td>
<td>0.75 (27.19)</td>
</tr>
<tr>
<td>Global VAS</td>
<td>37.21 (26.63)</td>
<td>35.24 (27.96)</td>
<td>–1.95 (23.23)</td>
</tr>
</tbody>
</table>

### Table 3. Scores in the “better” and “worse” subgroups were used as MID estimates. In particular, better/worse MID [mean (SD; 95% CI)] were: –0.131 (0.411; –0.272, 0.0105) and 0.131 (0.309; 0.0436, 0.219) for HAQ-DI; –9.37 (24.37; –17.74, –1.00) and 13.96 (22.05; 7.62, 20.29) for pain VAS; –8.15 (23.52; –16.35, 0.06) and 3.63 (22.62; –4.30, 11.56) for fatigue VAS; –10.97 (27.94; –21.35, –0.60) and 13.81 (27, 32; 5.88, 21.75) for sleep VAS; and 8.41 (21.17; –15.80, –1.03) and 11.53 (21.03; 5.49, 17.57) for global VAS. Spearman correlation coefficients between the anchor and mean change were 0.374 for HAQ-DI, 0.448 for pain VAS, 0.239 for fatigue VAS, 0.326 for sleep VAS, and 0.490 for global VAS, with p < 0.01 for all measures. Autocorrelations for change between visits were –0.046 for HAQ-DI, –0.136 for pain VAS, –0.048 for fatigue VAS, –0.036 for sleep VAS, and 0.049 for global VAS.

In general, the results of exploratory analyses of gender, DMARD use, and biologic use did not show differing MID values between subgroups (data not shown).

**DISCUSSION**

It is important to determine MID for different patient-reported outcomes in order to ascertain how much of a change in

---

**Table 3.** Scores in the “better” and “worse” subgroups were used as MID estimates. In particular, better/worse MID [mean (SD; 95% CI)] were: –0.131 (0.411; –0.272, 0.0105) and 0.131 (0.309; 0.0436, 0.219) for HAQ-DI; –9.37 (24.37; –17.74, –1.00) and 13.96 (22.05; 7.62, 20.29) for pain VAS; –8.15 (23.52; –16.35, 0.06) and 3.63 (22.62; –4.30, 11.56) for fatigue VAS; –10.97 (27.94; –21.35, –0.60) and 13.81 (27.32; 5.88, 21.75) for sleep VAS; and 8.41 (21.17; –15.80, –1.03) and 11.53 (21.03; 5.49, 17.57) for global VAS. Spearman correlation coefficients between the anchor and mean change were 0.374 for HAQ-DI, 0.448 for pain VAS, 0.239 for fatigue VAS, 0.326 for sleep VAS, and 0.490 for global VAS, with p < 0.01 for all measures. Autocorrelations for change between visits were –0.046 for HAQ-DI, –0.136 for pain VAS, –0.048 for fatigue VAS, –0.036 for sleep VAS, and 0.049 for global VAS.

In general, the results of exploratory analyses of gender, DMARD use, and biologic use did not show differing MID values between subgroups (data not shown).

**DISCUSSION**

It is important to determine MID for different patient-reported outcomes in order to ascertain how much of a change in...
likely that nearly all the patients met CASPAR criteria, as study group (CASPAR) criteria were not used. It is highly investigating MID for different patient-reported outcomes in our clinic, thus the ClASsification of Psoriatic ARthritis these criteria correspond with the way we have diagnosed varying populations of patients with PsA may be warranted. PsA over the last couple of years31. However, the Moll and Wright criteria set was used because it is much simpler and easier to extract retrospectively23.

Wright criteria reported in clinical trials as –0.22 to –0.2428. This was for HAQ-DI in RA at our site calculated as –0.090/0.150 and past, a discordance between MID in our clinic setting and patient-reported outcome in PsA have been calculated. In the study was found to be within 2 standard errors and ranging from 0.049 to –0.136, increasing the reliability of the MID.

This is the first time estimates of MID for any patient-reported outcome in PsA have generally more severe disease requiring active treatment than those who are stable in a clinic setting, and the expectations of patients in clinical trials may be higher. In a clinical trial setting, one could expect MID of these outcomes measured in our study was found to be within 2 standard errors and ranging from 0.049 to –0.136, increasing the reliability of the MID reported.

Patients were diagnosed with PsA by a rheumatologist in our clinic, thus the CIASSification of Psoriatic ARthritis study group (CASPAR) criteria were not used. It is highly likely that nearly all the patients met CASPAR criteria, as these criteria correspond with the way we have diagnosed PsA over the last couple of years31. However, the Moll and Wright criteria set was used because it is much simpler and easier to extract retrospectively23.

There were some limitations in our study. There was a large sample of 200 patients but from 1 rheumatologist only, and large SD were present in the MID. Although the CI were nonoverlapping between groups who rated their overall status as “better” and “worse” at the second visit, there was often overlap between each of these groups with the group that noted no change in their overall status between visits (i.e., “same” group), which may indicate that the better and worse groups are more similar than originally anticipated. This does not invalidate the results. The mean change is how other MID are interpreted, even if they are not mutually exclusive with respect to the CI. Our exploratory analysis was limited by sample size. The patient population was somewhat homogeneous, with most patients being Caucasian, able to read English, and having a long disease duration. Thus, our results may not be generalizable to those with early PsA, where the MID may be larger. Recall bias may have been present when asking patients to determine change of their disease status from one visit to the next, as there was a mean of 8.28 months between visits. The health status anchor (global rating of change) and the multiple patient-reported HRQOL outcome measures have not been validated for use in PsA. Further, in other rheumatologic diseases where a global rating of change has been validated, they are in part linked to changes in some patient-reported HRQOL outcome measures, leading to a degree of circular reasoning. Finally, there are multiple methods to calculate the MID for a given variable, including the use of different anchors, using distributional estimates, and/or triangulation of multiple anchors. Variable followup length may have affected the MID estimates but it is difficult to say if this was the case. This was a real-world study, so patients returned for both scheduled visits and extra visits if they were worse or medication was changed. This is how we returned for both scheduled visits and extra visits if they were worse or medication was changed. This is how we...

**Table 3.** MID estimates for HAQ-DI, pain VAS, fatigue VAS, sleep VAS, and global VAS. Measures are mean (SD) [95% CI] change.

<table>
<thead>
<tr>
<th>Overall Status</th>
<th>HAQ-DI</th>
<th>Pain VAS</th>
<th>Fatigue VAS</th>
<th>Sleep VAS</th>
<th>Global VAS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Much better (n = 22)</td>
<td>–0.362 (0.432)</td>
<td>–29.00 (31.86)</td>
<td>–15.32 (22.54)</td>
<td>–12.09 (28.57)</td>
<td>–27.68 (30.43)</td>
</tr>
<tr>
<td>[–0.554, –0.171]</td>
<td>[–43.13, –14.87]</td>
<td>[–25.31, –5.33]</td>
<td>[–24.76, 0.57]</td>
<td>[–41.17, –14.19]</td>
<td></td>
</tr>
<tr>
<td>Better (n = 35)</td>
<td>–0.131 (0.411)</td>
<td>–9.37 (24.37)</td>
<td>–8.15 (23.52)</td>
<td>–10.97 (29.74)</td>
<td>–8.41 (21.17)</td>
</tr>
<tr>
<td>[–0.272, 0.0105]</td>
<td>[–17.74, –1.00]</td>
<td>[–16.35, 0.06]</td>
<td>[–21.35, –6.00]</td>
<td>[–15.80, –1.03]</td>
<td></td>
</tr>
<tr>
<td>Same (n = 89)</td>
<td>0.00008 (0.292)</td>
<td>–5.03 (20.37)</td>
<td>–2.80 (20.89)</td>
<td>1.19 (22.89)</td>
<td>–1.56 (15.73)</td>
</tr>
<tr>
<td>[–0.0614, 0.0615]</td>
<td>[–9.45, –0.61]</td>
<td>[–7.33, 1.74]</td>
<td>[–3.72, 6.09]</td>
<td>[–4.93, 1.81]</td>
<td></td>
</tr>
<tr>
<td>Worse (n = 50)</td>
<td>0.131 (0.309)</td>
<td>13.96 (22.05)</td>
<td>3.63 (27.615)</td>
<td>13.81 (27.32)</td>
<td>11.53 (21.03)</td>
</tr>
<tr>
<td>[0.0436, 0.219]</td>
<td>[7.62, 20.29]</td>
<td>[–4.30, 11.56]</td>
<td>[8.88, 21.75]</td>
<td>[5.49, 17.57]</td>
<td></td>
</tr>
<tr>
<td>Much worse (n = 4)</td>
<td>0.438 (0.315)</td>
<td>16.50 (16.94)</td>
<td>27.50 (26.64)</td>
<td>4.75 (7.37)</td>
<td>20.75 (20.37)</td>
</tr>
<tr>
<td>[–0.0631, 0.938]</td>
<td>[–10.46, 43.46]</td>
<td>[–14.89, 69.89]</td>
<td>[–6.97, 16.47]</td>
<td>[–11.66, 53.16]</td>
<td></td>
</tr>
</tbody>
</table>

Spearman correlation coefficient* p < 0.01 p < 0.01 p < 0.01 p < 0.01 p < 0.01

* Correlation is between anchor and mean change in score, significant at the level shown. MID: minimally important difference; HAQ-DI: Health Assessment Questionnaire-Disability Index (0–3); VAS: visual analog scale (0–100 mm).
seen in a shorter timeframe because of a medication change, then they could be much better; if they were seen earlier because of a flare, they could be much worse. The overall health question was not specific to skin and psoriasis and can be discordant between the skin and the joints. Future studies of MID in PsA should use multiple estimates of MID by different methods in order to maximize reliability of the results.

This is the first study to describe MID for patient-reported outcomes in PsA. MID for HAQ-DI, pain, and global VAS showed the largest effect size and are seen to be the best predictors of a patient’s perception of overall changes in disease status.

REFERENCES

29. 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.