








## ORIGINAL RESEARCH

# Validity and score interpretation of the 12-item Psoriatic Arthritis Impact of Disease: an analysis of pooled data from two phase 3 trials of bimekizumab in patients with psoriatic arthritis

Laure Gossec <sup>1,2</sup> Ana-Maria Orbai <sup>3</sup> Laura C Coates <sup>4</sup>  
Dafna D Gladman <sup>5</sup> Alexis Ogdie <sup>6</sup> Christopher G Pelligra <sup>7</sup>  
Valérie Ciaravino,<sup>8</sup> Barbara Ink,<sup>9</sup> Vanessa Taieb,<sup>8</sup> Jérémy Lambert,<sup>8</sup>  
Maarten de Wit <sup>10</sup>

**To cite:** Gossec L, Orbai A-M, Coates LC, *et al.* Validity and score interpretation of the 12-item Psoriatic Arthritis Impact of Disease: an analysis of pooled data from two phase 3 trials of bimekizumab in patients with psoriatic arthritis. *RMD Open* 2024;**10**:e003548. doi:10.1136/rmdopen-2023-003548

► Additional supplemental material is published online only. To view, please visit the journal online (<http://dx.doi.org/10.1136/rmdopen-2023-003548>).

Received 28 July 2023

Accepted 29 November 2023



© Author(s) (or their employer(s)) 2024. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

For numbered affiliations see end of article.

### Correspondence to

Professor Laure Gossec;  
laure.gossec@aphp.fr

## ABSTRACT

**Objectives** To investigate psychometric performance of the 12-item Psoriatic Arthritis Impact of Disease (PsAID-12) total and individual item scores in patients with psoriatic arthritis (PsA) and to estimate score change thresholds and scores corresponding to different levels of symptom/impact severity.

**Methods** Data up to week 16 from 1252 patients with active PsA enrolled in two randomised controlled trials of bimekizumab (BE OPTIMAL (NCT03895203) and BE COMPLETE (NCT03896581)) were used to assess construct validity (correlations with other patient-reported outcomes), known-groups validity (based on Minimal Disease Activity index, Disease Activity Index for Psoriatic Arthritis and Psoriatic Arthritis Disease Activity Score), reliability (Cronbach's alpha and intraclass correlation coefficients (ICCs)) and responsiveness (sensitivity to change). Clinically meaningful within-patient improvement thresholds were estimated by anchor-based and distribution-based analyses, and symptom/impact severity thresholds were estimated by receiver operating characteristic curve analyses.

**Results** The mean (SD) PsAID-12 total score at baseline was 4.19 (1.94). PsAID-12 scores demonstrated good convergent validity and good known-groups validity. Internal consistency reliability (Cronbach's alpha 0.95) and test-retest reliability (ICC  $\geq$  0.70) were also good. Responsiveness was acceptable (correlations  $\geq$  0.30 for most scores). Improvement thresholds were estimated at 1.5–2 points for the PsAID-12 total score and 2 or 3 points for item scores. Thresholds for different levels of symptom/impact severity could be derived for most PsAID-12 items.

**Conclusions** The PsAID-12 demonstrated robust psychometric properties in a large sample of patients with active PsA, supporting its use as a fit-for-purpose patient-reported outcome in this

## WHAT IS ALREADY KNOWN ON THIS TOPIC

- ⇒ The 12-item Psoriatic Arthritis Impact of Disease (PsAID-12) is a patient-reported outcome (PRO) measure for assessing the symptoms and impacts of psoriatic arthritis (PsA). It has received provisional endorsement for use in trials from the Group for Research and Assessment of Psoriasis and Psoriatic Arthritis (GRAPPA) and Outcome Measures in Rheumatology (OMERACT).
- ⇒ Several moderate-sized studies have validated psychometric properties of the PsAID-12 total score in patients with PsA. However, additional validation of the PsAID-12 total score and assessment of the psychometric properties of individual PsAID-12 items are needed.
- ⇒ Furthermore, clinically meaningful within-patient improvement thresholds (cut-off values for change) would facilitate interpretation of this PRO measure.

## WHAT THIS STUDY ADDS

- ⇒ Through analysis of data for more than 1200 patients in the BE OPTIMAL and BE COMPLETE randomised controlled trials, this study confirmed the good psychometric properties of the PsAID-12 total score.
- ⇒ Psychometric properties of the individual PsAID-12 items were generally acceptable. In particular, PsAID-12 pain, fatigue and skin problems item scores showed good convergent validity, known-groups validity, test-retest reliability and responsiveness.
- ⇒ Clinically meaningful within-patient improvement thresholds were established using anchor-based methods recommended by the US Food and Drug Administration. These thresholds were 2 or 3 points for individual PsAID-12 items and 1.5–2 points for the PsAID-12 total score.

### HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

- ⇒ This new information on the robust psychometric properties of PsAID-12 item and total scores in a large sample of patients with active PsA demonstrates that the PsAID-12 is fit-for-purpose to capture treatment effects in this patient population. The study provides evidence highlighted by GRAPPA-OMERACT as being necessary for full endorsement of the PsAID-12 as a recommended quality-of-life outcome measure in the PsA core measurement set.
- ⇒ The clinically meaningful within-patient improvement thresholds estimated in this study can be used to evaluate treatment effects in clinical trials and can help physicians to determine whether their patients are responding to treatment.

population. Furthermore, thresholds for score interpretation were derived.

### INTRODUCTION

The physical and psychological manifestations of psoriatic arthritis (PsA) have a major impact on patients' well-being, limiting activities and causing emotional distress.<sup>1</sup> Patients with PsA frequently report pain, fatigue, sleep problems, difficulties at work, withdrawal from social activities, problems interacting with family members and intimacy issues.<sup>2,3</sup>

Many of the symptoms and impacts of PsA are best assessed by patients themselves. The 12-item Psoriatic Arthritis Impact of Disease-12 (PsAID-12) is a patient-reported outcome (PRO) measure for assessing the symptoms and impact of PsA. The PsAID-12 was developed with patients and covers the main aspects important to them.<sup>4</sup> It is provisionally endorsed by the Group for Research and Assessment of Psoriasis and Psoriatic Arthritis (GRAPPA)-Outcome Measures in Rheumatology (OMERACT) community to capture health-related quality of life in PsA clinical trials.<sup>5</sup>

The PsAID-12 total score was previously validated in moderate-scale studies,<sup>4,6–10</sup> which were mostly observational. However, the individual PsAID-12 items have not been subject to psychometric validation, which is needed for them to be used individually. Because they represent different aspects of PsA, and because qualitative research has shown that they are some of the most frequent and disruptive symptoms of the disease,<sup>11</sup> the PsAID-12 items relating to pain, skin problems and fatigue could be considered as priorities.

Thresholds defining clinically meaningful within-patient change in symptoms and different levels of symptom/impact severity—including the level at which patients consider themselves well<sup>12</sup>—are needed to support the interpretation of PRO scores and score changes. Such thresholds are important for fully assessing patients' status and properly interpreting treatment effects in clinical trials. Earlier studies estimated the minimum clinically meaningful within-patient improvement threshold for the PsAID-12 total score.<sup>4,13</sup> However,

the estimates were solely derived using receiver operating characteristic (ROC) curves without reporting a range of estimates based on the anchor used in the analysis, while the US Food and Drug Administration (FDA) recommends methods supplemented by empirical cumulative distribution function and probability density function curves to evaluate a range of thresholds.<sup>14–16</sup>

The aims of this study were to confirm the psychometric performance of the PsAID-12 total score in patients with PsA; to investigate the psychometric performance of individual PsAID-12 items, in particular the pain, fatigue and skin problems items; and to estimate clinically meaningful within-patient change thresholds and disease symptom/impact severity thresholds.

### METHODS

#### Study design

This study was a post hoc analysis of pooled data from two previously reported phase 3 clinical trials of bimekizumab, a monoclonal IgG1 antibody that selectively inhibits IL-17F in addition to IL-17A, in patients with active PsA.<sup>17,18</sup> In the BE OPTIMAL trial (NCT03895203), patients with PsA who were biological disease-modifying antirheumatic drug-naïve were randomised 3:2:1 to subcutaneous bimekizumab (160 mg every 4 weeks (Q4W)), placebo or an active reference arm (adalimumab 40 mg every 2 weeks).<sup>17</sup> In the BE COMPLETE trial (NCT03896581), patients with PsA and a prior inadequate response or intolerance to tumour necrosis factor- $\alpha$  inhibitor treatment were randomised 2:1 to bimekizumab (160 mg Q4W) or placebo.<sup>18</sup> For both trials, blinded data from the 16-week, double-blind, placebo-controlled treatment period were used. Data for patients who received placebo and active treatment were pooled and analysed together. Patients with major depression were ineligible to participate in the trials.

#### PRO measures and other assessments

The PRO measure of interest was the PsAID-12, formed of 12 questions which are each assessed on a 0–10 numeric rating scale. A higher score indicates a worse impact of PsA.<sup>19</sup> The PsAID-12 was completed in the BE OPTIMAL and BE COMPLETE trials at baseline, week 4, week 12 (BE COMPLETE only) and week 16. It was administered using an electronic device, which did not allow for partially completed questionnaires. Other assessments, done at various time points including baseline and week 16, included American College of Rheumatology (ACR) response,<sup>20,21</sup> Health Assessment Questionnaire-Disability Index (HAQ-DI),<sup>22</sup> Patient's Assessment of Arthritis Pain,<sup>21</sup> Patient's Global Assessment of PsA (PGA-PsA),<sup>21</sup> Psoriatic Arthritis Quality of Life (PsAQoL),<sup>23</sup> Functional Assessment of Chronic Illness Therapy-Fatigue (FACIT-Fatigue),<sup>24</sup> Short Form 36-item Health Survey (SF-36),<sup>25</sup> Disease Activity Index for Psoriatic Arthritis (DAPSA),<sup>26</sup> Psoriatic Arthritis Disease Activity Score (PASDAS)<sup>27</sup> and Minimal Disease Activity (MDA) index.<sup>28–30</sup>

### Statistical analyses

Psychometric analyses were conducted using data from observed cases for all randomised patients who had at least one non-missing PsAID-12 assessment in the period from baseline to week 16. Treatment allocation was not considered in any of the analyses. Descriptive statistics were calculated for PsAID-12 scores and scores for other outcome measures. PsAID-12 scores were further analysed for normality of distribution using the Shapiro-Wilk test.

For PsAID-12 item and total scores, the available data rate was calculated as the number of patients with a non-missing score at a given assessment divided by the number of randomised patients who were expected to complete the questionnaire (ie, were still receiving study drug) or had a non-missing score at that assessment. The completion rate was calculated as the number of patients with a non-missing score at a given assessment divided by the total number of randomised patients.

### Construct validity and reliability

To assess convergent and divergent validity, Spearman rank correlation coefficients were computed for the associations between PsAID-12 scores and scores for selected outcome measures (PsaQoL total score, FACIT-Fatigue, PGA-PsA, and SF-36 physical functioning, role-physical, bodily pain, general health, vitality, social functioning, role-emotional, mental health, physical component summary (PCS), and mental component summary (MCS)). Correlations were interpreted as  $|r| < 0.3$ =weak,  $0.3 \leq |r| < 0.5$ =moderate and  $|r| \geq 0.5$ =strong.<sup>31</sup> We hypothesised that, compared with other outcome measures, PsAID-12 pain would be more strongly correlated with SF-36 bodily pain and that PsAID-12 fatigue would be more strongly correlated with FACIT-Fatigue and SF-36 vitality. As skin problems were not assessed directly by the other PROs, it was hypothesised that this PsAID-12 item would generally have weaker correlations than the other PsAID-12 items. PsAID-12 total score was hypothesised to be more strongly correlated with PGA-PsA, PsaQoL total score and SF-36 general health than with other outcome measures.

Known-groups validity was assessed by analysis of variance (ANOVA) comparing mean PsAID-12 scores between groups of patients categorised based on MDA index (MDA, very low disease activity (VLDA) and non-MDA), DAPSA ( $\leq 4$ =remission,  $>4$  to  $14$ =low disease activity,  $>14$  to  $28$ =moderate disease activity and  $>28$ =high disease activity)<sup>26</sup> and PASDAS ( $\leq 1.9$ =remission,  $>1.9$  to  $<3.2$ =low disease activity,  $3.2$  to  $<5.4$ =moderate disease activity and  $\geq 5.4$ =high disease activity).<sup>32</sup>

To assess internal consistency reliability, item-total correlations were calculated as the Spearman rank correlation coefficient for a given PsAID-12 item score with the PsAID-12 total score derived without that item (ie, the corrected total score).  $|r| \geq 0.30$  was considered evidence of acceptable internal consistency reliability.<sup>33</sup> Internal consistency reliability was further assessed by

calculating Cronbach's alpha after deletion of individual items.<sup>34</sup> A standardised alpha coefficient  $\geq 0.70$  was regarded as acceptable.<sup>35</sup>

Test-retest reliability was assessed in stable patients (those with the same DAPSA category at week 12 and week 16) in BE COMPLETE (the BE OPTIMAL trial did not include a PsAID-12 assessment at week 12). The intra-class correlation coefficient (ICC) was calculated using two-way mixed-effect ANOVA models that included time (week) as a fixed effect. An ICC value  $>0.70$  was considered acceptable.<sup>36</sup>

### Responsiveness

To assess potential anchors for the responsiveness (sensitivity to change) analysis, Spearman correlation coefficients were calculated for the associations between changes from baseline to week 16 in PsAID-12 scores and changes from baseline to week 16 in other outcomes (PGA-PsA, DAPSA and PASDAS scores), as well as ACR response at week 16 (absolute value since ACR response is already a change).  $|r| \geq 0.3$  was considered an acceptable association for including the anchor in the responsiveness analysis.<sup>37,38</sup> The above composite anchors were chosen as they include measures assessing the key features of PsA and are widely used and validated in this disease.<sup>39</sup> The anchors chosen refer to disease activity, not at the symptoms levels, however, thus leading to a grouping of different concepts.

### Clinically meaningful within-patient change thresholds

In accordance with FDA guidance,<sup>14</sup> clinically meaningful within-patient change thresholds were derived by a mixed approach combining anchor-based and distribution-based analyses. In the anchor-based analysis, only outcome measures with  $|r| \geq 0.3$  in the responsiveness analysis were considered as potential anchors. For the included anchors (ACR response and PASDAS), patients were divided into different response groups, and descriptive statistics were calculated for PsAID-12 score changes from baseline to week 16 within these response groups. The descriptive statistics (particularly the mean and median PsAID-12 score changes) were used as a basis for clinically meaningful within-patient change thresholds. Effect sizes were calculated as the mean change from baseline to week 16 divided by the overall baseline SD. In addition, empirical cumulative distribution function and probability density function curves of changes in PsAID-12 scores from baseline to week 16 were used in the selection of the final clinically meaningful within-patient change thresholds.

In the distribution-based analysis, SE of measurement (SEM) and 0.5 SD were calculated using baseline data.<sup>35,40</sup> The estimates from the anchor- and distribution-based analyses were triangulated to determine a threshold for meaningful within-patient change. Ranges for the thresholds were provided based on the point estimates from the anchor-based and distribution-based estimates, with greater weight given to the anchor-based estimates

than the distribution-based estimates, as recommended in FDA guidance.<sup>14</sup>

### Patient symptom/impact severity thresholds

For each PsAID-12 score, thresholds for different levels of symptom/impact severity (ie, to interpret absolute PsAID-12 score at a given time) were derived using pooled data from the baseline, week 4 and week 16 assessments. First, patients' disease activity was separately categorised as 'remission', 'low disease activity', 'moderate disease activity' or 'high disease activity' based on published thresholds for the DAPSA and PASDAS.<sup>26–32</sup> Then, PsAID-12 scores were assessed for their ability to classify patients in each of the different DAPSA and PASDAS disease activity groups in ROC analyses. For example, to determine the threshold to distinguish patients with no symptoms/impact from patients with low symptom/impact severity on the PsAID-12, the optimal PsAID-12 item score was determined to correctly classify patients into the remission group versus the combined low/moderate/high disease activity group defined by DAPSA. This was repeated for each of the thresholds (ie, combined remission/low disease activity vs combined moderate/high disease activity group by DAPSA to distinguish low symptoms/impact severity from moderate symptoms/impact severity on the PsAID-12; combined remission/low disease activity/moderate disease activity group vs high disease activity group by DAPSA to distinguish moderate symptoms/impact severity from high symptoms/impact severity on the PsAID-12). The optimal symptom/impact severity threshold was the PsAID-12 score that maximised Youden's index (sensitivity+specificity–1).<sup>41</sup> An area under the curve (AUC) value of  $\geq 0.70$  was interpreted as evidence of satisfactory accuracy in differentiating patients.<sup>42</sup> The above analysis that used severity groups on the DAPSA was repeated using severity groups on the PASDAS.

### Patient and public involvement

Patients were not directly involved in the design or conduct of this study, which used data from previous trials. However, interpretation of the data and writing of this manuscript benefited from input from an author and patient research-partner (MdW).

## RESULTS

### Patient characteristics

A total of 1252 patients completed at least one PsAID-12 assessment and were included in the analysis. Patients had a mean age of 49.3 years at baseline (table 1). Just over half (53.0%) were female. Disease activity was high. However, most patients (84.2%) had a Psoriasis Area and Severity Index score of  $<10$  and thus had mild or moderate psoriasis. Nearly three-quarters of patients (73.4%) had previously received one or more conventional disease-modifying antirheumatic drugs.

### Available data and completion rates

The available data rate for the PsAID-12 was 99.9% at baseline, 99.4% at week 4 and 99.8% at week 16. The completion rate was 99.9% at baseline, 98.3% at week 4 and 96.2% at week 16.

### Distributions of PsAID-12 item scores

The mean (SD) PsAID-12 total score at baseline, on a 0–10 scale, was 4.19 (1.94) (online supplemental table 1). Mean (SD) baseline scores for PsAID-12 pain, fatigue and skin problems were 5.56 (2.26), 4.83 (2.53) and 4.68 (2.80), respectively. Compared with baseline, mean (SD) scores for PsAID-12 total score (2.65 (2.02)), pain (3.76 (2.51)), fatigue (3.32 (2.56)) and skin problems (2.60 (2.60)) were lower at week 16.

Floor and ceiling effects were not observed for PsAID-12 total score, pain, fatigue or skin problems at baseline or week 16. By contrast, strong floor effects were observed for anxiety/fear/uncertainty (scored as 0 by 32.6% of patients at baseline and 47.1% at week 16), embarrassment/shame (31.3%/52.2%), social participation (30.6%/46.9%) and depression (56.9%/65.1%) (online supplemental table 2).

### Convergent and divergent validity

For PsAID-12 pain, fatigue, skin problems and total score, correlations with other outcome measures were generally stronger at week 16 than at baseline (table 2 and online supplemental tables 3,4). PsAID-12 pain showed strong correlations with SF-36 bodily pain and moderate to strong correlations with most other outcome measures at baseline and week 16 (table 2). The exceptions were SF-36 role-emotional (weak correlations at baseline and week 16), SF-36 mental health (weak correlation at baseline) and SF-36 MCS (weak correlations at baseline and week 16). PsAID-12 pain was more strongly correlated with SF-36 bodily pain than with most other outcome measures.

PsAID-12 fatigue showed moderate to strong correlations with all outcome measures at baseline and week 16 and was more strongly correlated with FACIT-Fatigue and SF-36 vitality than with most other outcome measures.

Generally, correlations for PsAID-12 skin problems were weaker than those for other PsAID-12 items, including PsAID-12 pain and fatigue. At baseline, PsAID-12 skin problems showed weak correlations with FACIT-Fatigue and all SF-36 outcomes. At week 16, PsAID-12 skin problems had moderate to strong correlations with all outcomes except for SF-36 role-emotional, SF-36 mental health and SF-36 MCS. Correlations with PsAQoL total score were moderate at both baseline and week 16.

The PsAID-12 total score showed moderate to strong correlations with all outcome measures at baseline and week 16. Among the strongest correlations were those with PsAQoL total score, FACIT-Fatigue, PGA-PsA, SF-36 physical functioning, SF-36 role-physical, SF-36 bodily pain and SF-36 PCS.

**Table 1** Patient demographics and baseline disease characteristics for 1252 patients with active PsA

Characteristic	BE OPTIMAL (bDMARD-naïve) (N=852)	BE COMPLETE (TNF- $\alpha$ inhibitor-IR) (N=400)	Overall (N=1252)
Age (years), mean (SD)	48.7 (12.3)	50.5 (12.6)	49.3 (12.4)
BMI (kg/m <sup>2</sup> ), n (%)			
<25	233 (27.3)	83 (20.8)	316 (25.2)
25 to <30	282 (33.1)	133 (33.3)	415 (33.1)
≥30	337 (39.6)	184 (46.0)	521 (41.6)
Gender, n (%)			
Female	453 (53.2)	210 (52.5)	663 (53.0)
Male	399 (46.8)	190 (47.5)	589 (47.0)
Race, n (%)			
White	813 (95.8)	384 (96.0)	1197 (95.8)
Black or African American	2 (0.2)	0 (0.0)	2 (0.2)
Asian	28 (3.3)	13 (3.3)	41 (3.3)
American Indian or Alaska Native	1 (0.1)	0 (0.0)	1 (0.1)
Other or mixed	5 (0.6)	3 (0.8)	8 (0.6)
Missing	3 (0.4)	0 (0.0)	3 (0.2)
Geographical region, n (%)			
Asia	88 (10.3)	13 (3.3)	101 (8.1)
Eastern Europe	534 (62.7)	250 (62.5)	784 (62.6)
North America	103 (12.1)	109 (27.3)	212 (16.9)
Western Europe	127 (14.9)	28 (7.0)	155 (12.4)
Tender joint count, mean (SD)	17.0 (12.2)	18.7 (13.8)	17.5 (12.8)
Swollen joint count, mean (SD)	9.2 (6.7)	9.9 (7.7)	9.4 (7.1)
Body surface affected by psoriasis, n (%)			
<3%	427 (50.1)	136 (34.0)	563 (45.0)
≥3% to 10%	278 (32.6)	172 (43.0)	450 (35.9)
>10%	147 (17.3)	92 (23.0)	239 (19.1)
PASI score, n (%)			
<10	739 (86.7)	315 (78.8)	1054 (84.2)
10 to 20	91 (10.7)	68 (17.0)	159 (12.7)
>20	22 (2.6)	17 (4.3)	39 (3.1)
Prior TNF- $\alpha$ inhibitor exposure, n (%)			
TNF- $\alpha$ inhibitor intolerance	0 (0.0)	49 (12.3)	49 (3.9)
Inadequate response to 1 TNF- $\alpha$ inhibitor	0 (0.0)	306 (76.5)	306 (24.4)
Inadequate response to 2 TNF- $\alpha$ inhibitors	0 (0.0)	45 (11.3)	45 (3.6)
None	852 (100)	0 (0.0)	852 (68.1)
Prior conventional DMARDs, n (%)			
0	185 (21.7)	148 (37.0)	333 (26.6)
1	545 (64.0)	218 (54.5)	763 (60.9)
≥2	122 (14.3)	34 (8.5)	156 (12.5)

bDMARD, biological disease-modifying antirheumatic drug; BMI, body mass index; IR, inadequate response; PASI, Psoriasis Area and Severity Index; PsA, psoriatic arthritis; TNF, tumour necrosis factor.

**Table 2** Correlations between PsAID-12 pain, fatigue, skin problems, and total score and scores for other patient-reported outcomes

Outcome	Visit	Spearman rank correlation*†			
		PsAID-12 pain	PsAID-12 fatigue	PsAID-12 skin problems	PsAID-12 total score
PsAQoL total score	Baseline	0.48	<b>0.62</b>	0.30	<b>0.68</b>
	Week 16	<b>0.63</b>	<b>0.71</b>	0.42	<b>0.75</b>
FACIT-fatigue	Baseline	0.46	<b>0.71</b>	0.27	<b>0.66</b>
	Week 16	<b>0.58</b>	<b>0.77</b>	0.42	<b>0.74</b>
PGA-PsA	Baseline	<b>0.65</b>	<b>0.51</b>	0.48	<b>0.64</b>
	Week 16	<b>0.81</b>	<b>0.65</b>	<b>0.59</b>	<b>0.79</b>
SF-36 physical functioning	Baseline	<b>0.54</b>	<b>0.53</b>	0.25	<b>0.60</b>
	Week 16	<b>0.67</b>	<b>0.65</b>	0.40	<b>0.72</b>
SF-36 role-physical	Baseline	<b>0.53</b>	<b>0.55</b>	0.23	<b>0.61</b>
	Week 16	<b>0.65</b>	<b>0.68</b>	0.41	<b>0.74</b>
SF-36 bodily pain	Baseline	<b>0.69</b>	<b>0.58</b>	0.24	<b>0.66</b>
	Week 16	<b>0.81</b>	<b>0.69</b>	0.46	<b>0.80</b>
SF-36 general health	Baseline	0.38	0.41	0.22	0.46
	Week 16	<b>0.50</b>	<b>0.53</b>	0.35	<b>0.58</b>
SF-36 vitality	Baseline	0.40	<b>0.61</b>	0.22	<b>0.57</b>
	Week 16	<b>0.54</b>	<b>0.72</b>	0.37	<b>0.67</b>
SF-36 PCS	Baseline	<b>0.62</b>	<b>0.55</b>	0.25	<b>0.62</b>
	Week 16	<b>0.76</b>	<b>0.69</b>	0.44	<b>0.78</b>

Convergent and divergent validity were assessed at baseline (n=1251) and week 16 (n=1204).

Strong correlations ( $\geq 0.5$ ) are shown in bold.

\*Absolute values are presented.

†p<0.001 for all correlations.

FACIT-Fatigue, Functional Assessment of Chronic Illness Therapy – Fatigue; PCS, physical component summary; PGA-PsA, Patient's Global Assessment of Psoriatic Arthritis; PsAID-12, Psoriatic Arthritis Impact of Disease-12; PsAQoL, Psoriatic Arthritis Quality of Life; SF-36, 36-item Short Form Health Survey.

For the other PsAID-12 items, most correlations with other outcome measures were moderate or strong, especially at week 16 (online supplemental table 4). Of note, PsAID-12 embarrassment/shame had weaker correlation with other outcome measures than other PsAID-12 items.

### Known-groups validity

PsAID-12 pain, fatigue, skin problems and total score showed known-groups validity based on MDA index and DAPSA and PASDAS scores at week 16 (figure 1 and online supplemental table 5). Mean and median scores were consistently higher (worse) for patients with higher disease activity than for those with lower disease activity or remission. Known-groups validity was also demonstrated for the other PsAID-12 items based on week 16 data (online supplemental table 6) and for all PsAID-12 items and the PsAID-12 total score based on baseline data (online supplemental tables 7,8). All distributions of PsAID-12 scores significantly differed by known group, as shown by the p values from the Kruskal-Wallis test (p<0.001).

### Internal consistency reliability

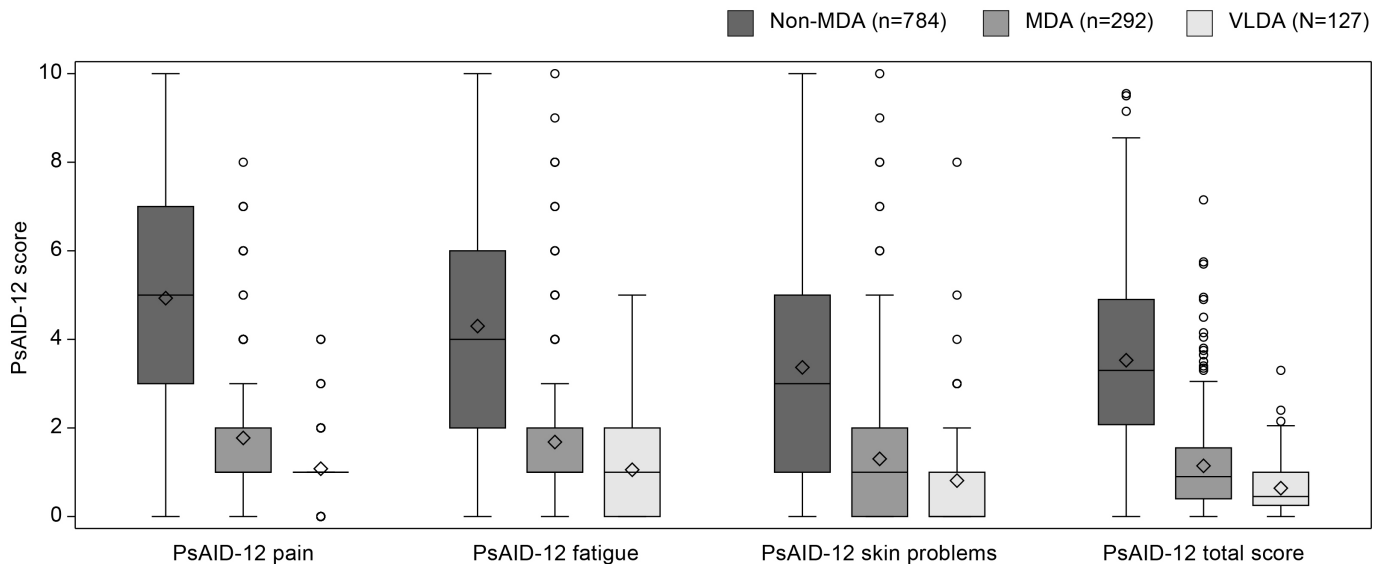
Item-total correlations at week 16 ranged from 0.51 to 0.90—above the prespecified threshold of 0.30 (table 3). This indicated a high degree of consistency among the PsAID-12 items. Internal consistency reliability was confirmed by an overall standardised Cronbach's alpha of 0.95 for the PsAID-12 total score at week 16 (table 3). This was substantially higher than the prespecified threshold of 0.70. Removing individual PsAID-12 items had little or no effect on Cronbach's alpha, suggesting that there was no item redundancy. Similar results were obtained at baseline (online supplemental table 9).

### Test-retest reliability

PsAID-12 pain, fatigue, skin problems and total score all showed acceptable test-retest reliability (ICC  $\geq 0.70$ ) (table 4). Test-retest reliability was also acceptable for the other PsAID-12 items (online supplemental table 10).

### Responsiveness

For PsAID-12 pain, fatigue, skin problems and total score, correlations of score changes from baseline to week 16 with ACR response were all  $\leq -0.3$ , and



**Figure 1** Box plots of PsAID-12 scores for known groups based on disease activity. Data are shown as the mean (diamonds), median (horizontal lines in boxes), first and third quartiles (lower and upper edges of boxes), and minimum and maximum (excluding outliers: lower and upper external horizontal bars). Outliers are represented by circles. A lower score indicates a less severe symptom or impact. MDA, minimal disease activity; PsAID-12, Psoriatic Arthritis Impact of Disease-12; VLDA, very low disease activity.

correlations with changes from baseline to week 16 in DAPSA, PASDAS, and PGA-PsA were all  $\geq 0.3$ , indicating acceptable responsiveness (all  $|r| \geq 0.3$ ; table 5). PsAID-12 work/leisure activities, functional capacity, discomfort, sleep disturbance, and coping also showed acceptable responsiveness based on all four anchors ( $|r| \geq 0.3$ ). PsAID-12 social participation showed low responsiveness based on ACR response ( $|r| < 0.3$ ), but acceptable responsiveness based on DAPSA, PASDAS, and PGA-PsA ( $|r| \geq 0.3$ ). PsAID-12

embarrassment/shame showed low responsiveness based on ACR response and DAPSA but acceptable responsiveness based on PASDAS and PGA-PsA. PsAID-12 anxiety/fear/uncertainty and depression showed low responsiveness based on all four anchors. Effect sizes for changes from baseline to week 16 in PsAID-12 scores by ACR response ( $\geq 70\%$ ;  $\geq 50\%$  to  $< 70\%$ ;  $\geq 20\%$  to  $< 50\%$ ;  $< 20\%$ ) and PASDAS improvement categories (three levels, two levels and one level of improvement, no change and one level of

**Table 3** Internal consistency reliability of the PsAID-12, as assessed using week 16 data

PsAID-12 score	Spearman rank correlation: item-corrected total score*	Cronbach's alpha: standardised items†
PsAID-12 total score	–	0.95
PsAID-12 pain	0.85	0.95
PsAID-12 fatigue	0.85	0.95
PsAID-12 skin problems	0.62	0.95
PsAID-12 work/leisure activities	0.89	0.94
PsAID-12 functional capacity	0.90	0.94
PsAID-12 discomfort	0.86	0.94
PsAID-12 sleep disturbance	0.76	0.95
PsAID-12 coping	0.84	0.95
PsAID-12 anxiety/fear/uncertainty	0.58	0.95
PsAID-12 embarrassment/shame	0.57	0.95
PsAID-12 social participation	0.71	0.95
PsAID-12 depression	0.51	0.95

\* $p < 0.001$  for all correlations.

†Cronbach's alpha was calculated with and without removal of individual PsAID-12 items.

PsAID-12, Psoriatic Arthritis Impact of Disease-12.

**Table 4** Test–retest reliability (n=251): pain, fatigue, skin problems and PsAID-12 total score

PsAID-12 score	Mean (95% CI) PsAID-12 score			ICC
	Week 12	Week 16	Change from week 12 to week 16	
PsAID-12 pain	4.31 (3.99, 4.64)	4.35 (4.01, 4.69)	0.04 (–0.12, 0.20)	0.88
PsAID-12 fatigue	3.89 (3.56, 4.22)	3.81 (3.48, 4.15)	–0.08 (–0.27, 0.11)	0.83
PsAID-12 skin problems	3.29 (2.94, 3.65)	3.20 (2.85, 3.56)	–0.09 (–0.34, 0.16)	0.75
PsAID-12 total score	3.18 (2.90, 3.45)	3.16 (2.88, 3.45)	–0.01 (–0.14, 0.12)	0.89

The analysis used data for patients in BE COMPLETE with the same DAPSA category at week 12 and week 16. DAPSA, Disease Activity Index for Psoriatic Arthritis; ICC, intraclass correlation coefficient; PsAID-12, Psoriatic Arthritis Impact of Disease-12.

worsening) are presented in online supplemental tables 11,12. The higher the level of ACR response or disease activity improvement assessed with the PASDAS, the larger the effect sizes for the changes in PsAID-12 scores; while the effect sizes were small in the group of patients with an ACR response <20% or who had no change in disease activity assessed with the PASDAS.

#### Clinically meaningful within-patient improvement thresholds

Based on the clear separation for each response group within each of the selected anchors as plotted in the empirical cumulative distribution function curves (figure 2 and online supplemental figure 1), probability density function curves (online supplemental figures 2,3), and the generally higher correlations for PASDAS than those for DAPSA in the responsiveness analysis,

ACR response and PASDAS were chosen as appropriate anchors for deriving clinically meaningful within-patient improvement thresholds. PGA-PsA was not used because clinically meaningful within-patient change has not yet been clearly established. Clinically meaningful within-patient worsening thresholds were not derived due to the small number of patients who experienced worsening on the PASDAS (n=33).

Median changes from baseline to week 16 in patients with an ACR response of 20% to <50% at week 16 were –2 for pain, fatigue and skin problems, and –1.95 for PsAID-12 total score (online supplemental table 13). Mean changes were –2.11 for pain, –1.67 for fatigue, –2.57 for skin problems and –1.86 for PsAID-12 total score. Median changes from baseline to week 16 in patients with a 1-level improvement in PASDAS from

**Table 5** Responsiveness analysis of correlations between PsAID-12 score changes from baseline to week 16 and changes for other outcomes during the same time interval, as well as ACR response at week 16

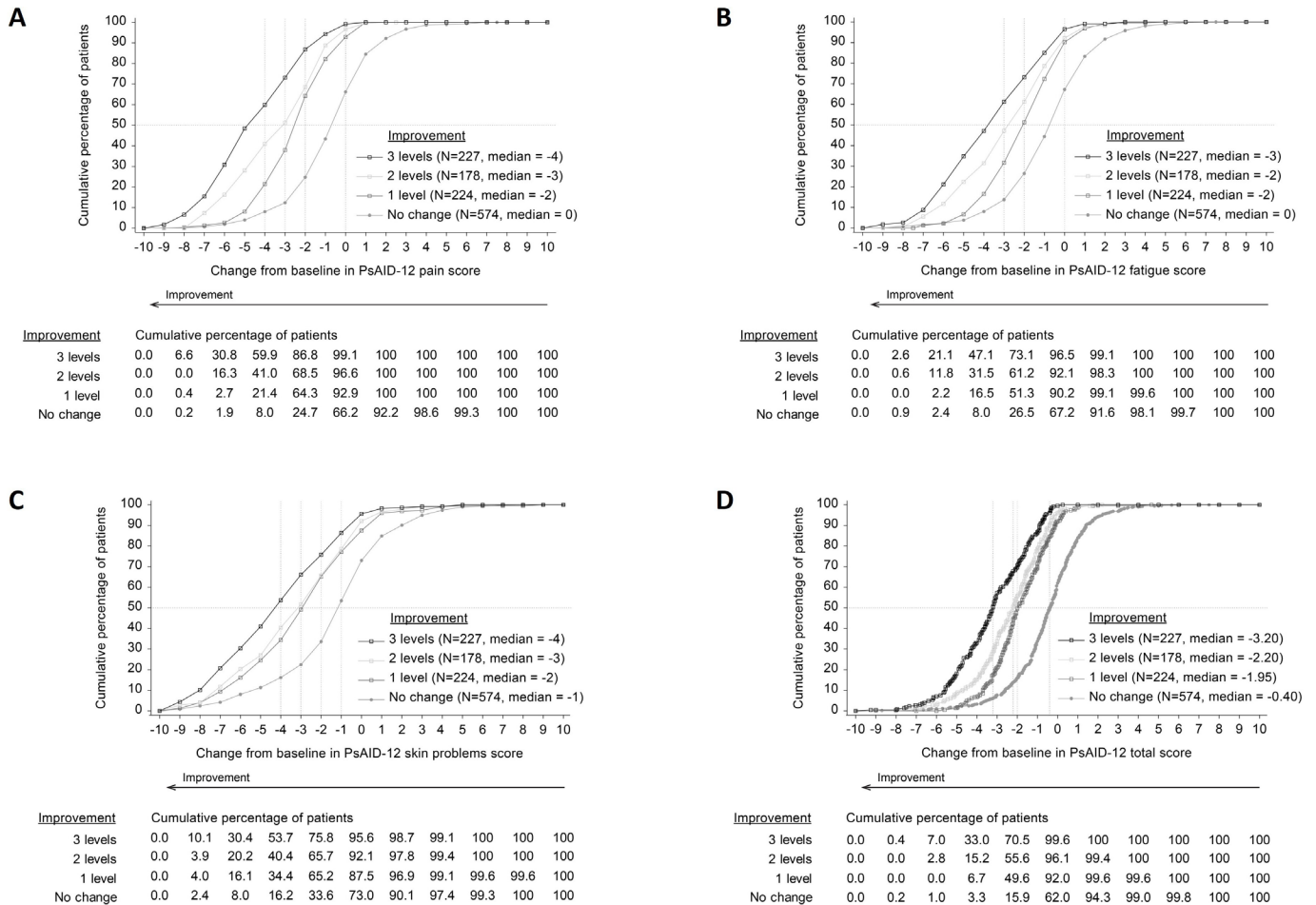
PsAID-12 score	Spearman rank correlation*			
	ACR response† (n=1203)	DAPSA (n=1196)	PASDAS (n=1159)	PGA-PsA (n=1203)
PsAID-12 pain	<b>–0.59</b>	<b>0.58</b>	<b>0.66</b>	<b>0.65</b>
PsAID-12 fatigue	<b>–0.46</b>	<b>0.50</b>	<b>0.54</b>	<b>0.53</b>
PsAID-12 skin problems	<b>–0.39</b>	<b>0.39</b>	<b>0.51</b>	<b>0.50</b>
PsAID-12 work/leisure activities	<b>–0.45</b>	<b>0.46</b>	<b>0.54</b>	<b>0.53</b>
PsAID-12 functional capacity	<b>–0.50</b>	<b>0.51</b>	<b>0.59</b>	<b>0.59</b>
PsAID-12 discomfort	<b>–0.46</b>	<b>0.46</b>	<b>0.53</b>	<b>0.53</b>
PsAID-12 sleep disturbance	<b>–0.34</b>	<b>0.40</b>	<b>0.40</b>	<b>0.41</b>
PsAID-12 coping	<b>–0.42</b>	<b>0.43</b>	<b>0.48</b>	<b>0.48</b>
PsAID-12 anxiety/fear/uncertainty	–0.23	0.21	0.25	0.24
PsAID-12 embarrassment/shame	–0.28	0.24	<b>0.35</b>	<b>0.36</b>
PsAID-12 social participation	–0.25	<b>0.32</b>	<b>0.32</b>	<b>0.36</b>
PsAID-12 depression	–0.18	0.20	0.21	0.20
PsAID-12 total score	<b>–0.56</b>	<b>0.57</b>	<b>0.65</b>	<b>0.66</b>

Correlations indicating acceptable responsiveness ( $|r| \geq 0.3$ ) are shown in bold.

\* $p < 0.001$  for all correlations.

†Categorical: absolute value at week 16 defined in categories ACR <20%, 20% to <50%, 50% to <70%,  $\geq 70\%$ .

ACR, American College of Rheumatology; DAPSA, Disease Activity Index for Psoriatic Arthritis; PASDAS, Psoriatic Arthritis Disease Activity Score; PGA-PsA, Patient's Global Assessment of Psoriatic Arthritis; PsAID-12, Psoriatic Arthritis Impact of Disease-12.



**Figure 2** Empirical cumulative distribution function curves of changes in PsAID-12 scores from baseline to week 16 by ACR response category. (A) PsAID-12 pain. (B) PsAID-12 fatigue. (C) PsAID-12 skin problems. (D) PsAID-12 total score. A three-level improvement was defined as an ACR response of  $\geq 70\%$ , a two-level improvement as an ACR response of 50% to  $<70\%$ , a one-level improvement as an ACR response of 20% to  $<50\%$ , and no change as an ACR response of  $<20\%$ . ACR, American College of Rheumatology; PsAID-12, Psoriatic Arthritis Impact of Disease-12.

baseline to week 16 were  $-2$  for pain,  $-1$  for fatigue,  $-2$  for skin problems and  $-1.35$  for PsAID-12 total score. Mean changes were  $-1.91$  for pain,  $-1.50$  for fatigue,  $-2.17$  for skin problems and  $-1.62$  for PsAID-12 total score.

In a distribution-based analysis, SEM was 0.82 for pain, 1.07 for fatigue, 1.37 for skin problems and 0.68 for PsAID-12 total score; 0.5 SD was 1.13 for pain, 1.26 for fatigue, 1.40 for skin problems and 0.97 for PsAID-12 total score.

Triangulation of the various estimates indicated that the clinically meaningful within-patient improvement thresholds should be a 2-point or 3-point decrease for PsAID-12 pain, fatigue and skin problems and a 1.5–2 point decrease for the PsAID-12 total score, with a higher threshold indicating greater stringency. Similar analyses for PsAID-12 work/leisure activities, functional capacity, discomfort, sleep disturbance and coping gave clinically meaningful within-patient improvement thresholds of 2 or 3 points. Improvement thresholds were not estimated for the anxiety/fear/uncertainty, embarrassment/shame and depression items.

### Patient symptom/impact severity thresholds for PsAID-12 scores

Derived from ROC analyses (online supplemental table 14), the threshold values shown in table 6 distinguish different levels of symptom/impact severity from the patient perspective. For PsAID-12 pain, a score of  $\leq 2$  was associated with remission/no symptoms/no impact, 3 low symptom/impact severity, 4 moderate symptom/impact severity and  $\geq 5$  high symptom/impact severity. The corresponding values for PsAID-12 fatigue and skin problems were 0 or 1, 2, 3 or 4 and  $\geq 5$ . For PsAID-12 total score, a score of  $\leq 1.15$  denotes remission/no symptoms/no impact,  $>1.15$  to 1.95 low symptom/impact severity,  $>1.95$  to 3.60 moderate symptom/impact severity and  $>3.60$  high symptom/impact severity. Symptom/impact severity thresholds were not derived for PsAID-12 anxiety/fear/uncertainty, embarrassment/shame and depression because AUC values from the ROC analyses were less than the prespecified threshold of 0.70.

**Table 6** PsAID-12 item and total scores for different disease activity states

	Remission/no symptoms/no impact	Low symptom/impact severity	Moderate symptom/impact severity	High symptom/impact severity
PsAID-12 pain	≤2	3	4	≥5
PsAID-12 fatigue	0 or 1	2	3 or 4	≥5
PsAID-12 skin problems	0 or 1	2	3 or 4	≥5
PsAID-12 work/leisure activities	0 or 1	2	3 or 4	≥5
PsAID-12 functional capacity	0 or 1	2	3 or 4	≥5
PsAID-12 discomfort	0 or 1	2	3 or 4	≥5
PsAID-12 sleep disturbance		← 0 or 1 →	2 or 3	≥4
PsAID-12 coping	0 or 1	2	3 or 4	≥5
PsAID-12 social participation		← 0 or 1 →	2	≥3
PsAID-12 total score	≤1.15	>1.15 to 1.95	>1.95 to 3.60	>3.60

Each PsAID-12 item is scored on a 0–10 scale. The PsAID-12 total score ranges from 0 to 10. Thresholds were derived from an ROC analysis in which potential cutoffs were assessed for their ability to correctly classify patients on the DAPSA and PASDAS.

DAPSA, Disease Activity Index for Psoriatic Arthritis; PASDAS, Psoriatic Arthritis Disease Activity Score; PsAID-12, Psoriatic Arthritis Impact of Disease-12; ROC, receiver operating characteristic.

## DISCUSSION

PRO measures are valuable tools for assessing symptoms and impacts; like other outcome measures, they need to be validated to ensure scores reflect what is intended to be measured, and threshold values are very useful to aid their interpretation. In this analysis of data from more than 1200 patients with active PsA, conducted in accordance with FDA guidance, measurement properties of the PsAID-12 total score were shown to be robust. Moreover, we were able to validate PsAID-12 items, and to estimate threshold values corresponding to clinically meaningful within-patient improvements for most PsAID item scores and the PsAID-12 total score. These ‘change thresholds’ can be used to interpret changes in PsAID-12 scores over time, for a given individual. Furthermore, we derived thresholds for PsAID-12 scores to define different levels of symptom/impact severity from the patient perspective, which correspond to different levels of disease activity on composite scores. These ‘state thresholds’ can be used to interpret absolute PsAID-12 scores at a given time point. Our findings confirm the validity of the PsAID-12 total score, and in particular its responsiveness, and are expected to facilitate its interpretation. They also open the possibility to use individual PsAID-12 items to facilitate patient management by maximising information use from this measure. Importantly, this study provides the requisite randomised controlled trial (RCT) evidence and threshold derivation for the PsAID-12 to progress from provisional to full endorsement by GRAPPA-OMERACT as a recommended quality-of-life outcome measure in the PsA core measurement set to be used in observational and interventional clinical trials.<sup>5</sup>

Distributional properties of the PsAID-12 were as expected, with most items showing no floor effects. PsAID-12 item and total scores demonstrated good

convergent, divergent and known-groups validity and moderate-to-good test–retest reliability. Internal-consistency reliability analyses indicated strong consistency among the PsAID-12 items and no item redundancy. Responsiveness was established for the PsAID-12 total score and individual PsAID-12 items, except for the anxiety/fear/uncertainty, embarrassment/shame, social participation and depression items.

The results of this study are consistent with a preliminary psychometric evaluation of the PsAID-12 total score using data from 474 patients with PsA.<sup>4</sup> Values for construct validity (Spearman correlations with other instruments), internal-consistency reliability (Cronbach’s alpha: 0.95 in this study, 0.93–0.94 in the preliminary validation study) and test–retest reliability (ICC: 0.89 in this study, 0.95 in the preliminary validation study) were similar in the two studies. Moreover, responsiveness was acceptable in both studies. Importantly, this study confirms the responsiveness of the PsAID-12 and its items, when the concept of responsiveness of PRO measures has led to debate in the past.<sup>43</sup> In an analysis based on prospectively collected data from 129 patients with PsA attending a specialist rheumatology hospital in the UK, the PsAID-12 total score had a Cronbach’s alpha of 0.95 and an ICC of 0.91.<sup>13</sup> As in this study, the PsAID-12 total score was strongly correlated with PsAQoL, FACIT-Fatigue and PGA scores. Additional validation studies have reported similar findings.<sup>6–10</sup>

A 2-point decrease in PsAID-12 pain, fatigue or skin problems score was identified as the lower bound for a clinically meaningful within-patient improvement, with a 3-point decrease indicating a marked within-patient improvement (large to very large effect size). These results are consistent with the responder definition for the Fatigue numeric rating scale derived from the clinical trial programme of ixekizumab in PsA, which

demonstrated a 3-point decrease as a large improvement and a meaningful score change estimated between 2 and 4 points.<sup>44</sup> For the PsAID-12 total score, the corresponding score changes were 1.5-point and 2-point decreases. Threshold estimates were based on mean and median PsAID-12 score changes in groups of patients who experienced an ACR response of 20% to <50% or a one-level improvement on the PASDAS. These two anchors had shown acceptable responsiveness and demonstrated separation between response groups in empirical cumulative distribution function and probability density function curves of change in PsAID-12 scores. The threshold estimates were supported by distribution-based estimates, which were lower than all anchor-based estimates.

The threshold values for the PsAID-12 total score are lower than the estimate of 3 points from the previous psychometric evaluation by Gossec *et al.*<sup>4</sup> This difference is likely due to differences in the study design and populations, and in the methodology and anchors used. Indeed, in the clinical study data sets, no patient-derived anchor, such as a patient global impression of severity or of change for which a clinically meaningful change threshold has been established, was available; composite outcome measures such as the PASDAS and DAPSA, with evidence for what constitutes a clinically meaningful within-patient change, were used as anchors. This is a limitation to the current study since the anchors used as references are not based on patient perceptions, but rather on disease activity on composite scores. Thus, the cut-off values proposed herein offer a way to superimpose different but overlapping concepts—the concepts of disease activity, and of patient-perceived disease impact. Furthermore, using composite measures which have been well-validated and are widely used, as anchors, offers an original approach to the issue often found when defining cut-off values, which is how to validate the reference standards used as anchors. The timing of the assessments for the clinically meaningful within-patient improvement analysis was also different for the two studies: 16 weeks in the present analysis vs 10–16 weeks in the previous study. Similarly, our symptom/impact severity thresholds are lower than similar thresholds reported from a cross-sectional analysis of 144 patients with PsA attending two tertiary rheumatology centres in Italy,<sup>7</sup> although this difference may reflect the use of a different methodological approach and a different patient population.

The floor effects and psychometric performance limitations observed for PsAID-12 anxiety/fear/uncertainty, embarrassment/shame, social participation and depression were likely due to formal exclusion of patients with depression from the BE OPTIMAL and BE COMPLETE trials. It should be noted that three of these four PsAID-12 items (embarrassment/shame, social participation and depression) were excluded from an alternative 9-item version of the PsAID because their inclusion did not improve psychometric performance of the PsAID.<sup>4</sup> Although these items were ranked as less important than pain, fatigue and skin problems, they were included in

the PsAID-12 because they are important for a minority of people with PsA<sup>4 45</sup> and should be evaluated and addressed in clinical practice.

This study has strengths and limitations. The analysis of convergent and divergent analysis did not include a dedicated skin-related outcome measure, which limited our ability to assess convergent validity of the PsAID-12 skin problems item. Moreover, as noted above, the estimation of PsAID-12 symptom/impact severity thresholds did not include MDA index or another measure that includes skin symptoms, but instead used the DAPSA and PASDAS. However, the significance of this limitation when using the estimated thresholds for patients with PsA may depend on the patients' particular disease characteristics. A further limitation is the use of data from phase 3 trial populations who were required to meet certain eligibility criteria. This skewed the sample to patients with active moderate-to-severe PsA and may limit extrapolation of the findings to the broader population of patients with PsA. Conversely, the multinational nature of the BE OPTIMAL and BE COMPLETE trials is a strength of the study since PsAID-12 validation can be considered in different languages. However, no cross-country comparisons were made. Additionally, data from RCTs with both active and placebo arms allowed for wider variation in PsAID-12 score changes and disease activity to be captured in the analyses, thus potentially resulting in more accurate estimates than those that might have been obtained from a non-interventional study.

In conclusion, the PsAID-12 total score and individual PsAID items capturing disease concepts important to patients with PsA demonstrated robust psychometric properties in a large sample of the target patient population. The results support use of the PsAID-12 as a fit-for-purpose PRO to inform assessment of treatment effects in patients with active PsA. Future studies should evaluate the utility of individual PsAID-12 items for monitoring and managing patients with PsA and should explore PsAID-12 symptom/impact severity thresholds maximising sensitivity (instead of Youden's index), which may be more useful in clinical practice.

#### Author affiliations

<sup>1</sup>INSERM, Institut Pierre Louis d'Epidémiologie et de Santé Publique, Sorbonne Université, Paris, France

<sup>2</sup>Rheumatology Department, AP-HP, Pitié-Salpêtrière Hospital, Paris, France

<sup>3</sup>Division of Rheumatology, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA

<sup>4</sup>Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, University of Oxford, Oxford, UK

<sup>5</sup>Schroeder Arthritis Institute, Krembil Research Institute, Toronto Western Hospital, University Health Network, University of Toronto, Toronto, Ontario, Canada

<sup>6</sup>Perelman School of Medicine at the University of Pennsylvania, Philadelphia, Pennsylvania, USA

<sup>7</sup>Evidera, Medellín, Colombia

<sup>8</sup>UCB Pharma, Colomnes, France

<sup>9</sup>UCB Pharma, Slough, UK

<sup>10</sup>Patient Research Partner, Stichting Tools, Amsterdam, Netherlands

**Correction notice** This article has been corrected since it was first published online. In this analysis, Disease Activity Index for Psoriatic Arthritis (DAPSA) disease states were used as a potential anchor in the psychometric analyses. The original DAPSA scores and number of patients in each DAPSA disease state had been calculated incorrectly, and correct values have since been generated to enable a correction of the work reported here. The error impacted the test-retest reliability analysis reported in Table 4 and Supplementary Table S10, the responsiveness analysis reported in Table 5, the known-groups validity analysis reported in Supplementary Tables S5–8, the analysis of the identified symptom/impact severity thresholds for 12-item Psoriatic Arthritis Impact of Disease (PsAID-12) reported in Supplementary Table S14, and corresponding text in the manuscript. All conclusions regarding the psychometric performance of the PsAID-12 total and individual domain scores remain unchanged with the corrected DAPSA data analysis. The authors note that DAPSA was not selected for use as an anchor to derive the clinically meaningful within-patient improvement thresholds for the PsAID-12 scores, either in the originally published analysis or the corrected analysis we have now reported. This was due to the generally smaller correlations in the responsiveness analysis with DAPSA, than with American College of Rheumatology (ACR) or Psoriatic Arthritis Disease Activity Score (PASDAS). The improvement thresholds identified in the analysis for the PsAID-12 total and individual domain scores are unchanged, and the correction of the DAPSA data has not impacted the conclusions of the analysis.

**X** Laura C Coates @drlauracoates

**Acknowledgements** The authors thank the patients and the investigators and their teams who took part in this study. Portions of this work were previously presented at ISPOR 2023 (7 May 2023–10 May 2023, Boston, Massachusetts, USA) and at EULAR 2023 (31 May–3 June, Milan, Italy). The authors acknowledge Stephen Gilliver, PhD, of Evidera, Sweden for providing medical writing support based on the authors' input and direction; Bethan Taylor, BA, Costello Medical, UK, for publication coordination; Dr Martin Bauer, MD, FNWC, UCB Pharma, Monheim am Rhein, Germany for his work in the BKZ in PsA program; and Heather Edens, PhD, UCB Pharma, Smyrna, GA, USA for publication coordination and editorial assistance. Medical writing support, publication coordination, and editorial assistance were provided in accordance with Good Publication Practice (GPP) 2022 guidelines (<https://www.ismpp.org/gpp-2022>).

**Contributors** Substantial contributions to study conception and design: LG, CGP, JL, VC, BI and VT; substantial contributions to analysis and interpretation of the data: LG, LCC, AO, A-MO, MdW, CGP, JL, VC, BI, VT and DDG; drafting the article or revising it critically for important intellectual content and final approval of the version of the article to be published: LG, LCC, AO, A-MO, MdW, CGP, JL, VC, BI, VT and DDG; guarantor: LG

**Funding** This study, as well as medical writing support, publication coordination, editorial assistance, and journal submission fees, were funded by UCB Pharma, Brussels, Belgium.

**Competing interests** LG: Grants from AbbVie, Biogen, Eli Lilly, Novartis, Sandoz and UCB Pharma; consulting fees from AbbVie, Amgen, BMS, Celltrion, Galapagos, Janssen, Eli Lilly, MSD, Novartis, Pfizer, Sandoz and UCB Pharma. A-MO: Research grants to Johns Hopkins University from AbbVie, Amgen and Janssen; consulting fees from BMS, Janssen, Sanofi and UCB Pharma. LCC: Grants/research support from AbbVie, Amgen, Celgene, Eli Lilly, Gilead, Janssen, Novartis, Pfizer and UCB Pharma; consultant for AbbVie, Amgen, BMS, Boehringer Ingelheim, Celgene, Domain, Eli Lilly, Galapagos, Gilead, Janssen, Moonlake Pharma, Novartis, Pfizer, and UCB Pharma; speaking fees from AbbVie, Amgen, Biogen, Celgene, Eli Lilly, Galapagos, Gilead, GSK, Janssen, medac, Novartis, Pfizer, and UCB Pharma. DG: Consultant and/or received grant support from AbbVie, Amgen, BMS, Celgene, Eli Lilly, Gilead, Galapagos, Janssen, Novartis, Pfizer, and UCB Pharma. AO: Grant/research support from AbbVie, Amgen, Janssen, Novartis and Pfizer, consultant of AbbVie, Amgen, BMS, Celgene, CorEvitas, Eli Lilly, GSK, Gilead, Janssen, Novartis, Pfizer, Takeda and UCB Pharma. CGP: Employed by Evidera, a part of Thermo Fisher Scientific, which received funding for this research from UCB Pharma. VC, VT and JL: Employees and stockholders of UCB Pharma. BI: Employee of UCB Pharma; shareholder of AbbVie, GlaxoSmithKline and UCB Pharma. MdW: Over the last five years Stichting Tools has received fees for lectures or consultancy provided by MdW from Celgene, Eli Lilly, Janssen-Cilag, Pfizer and UCB Pharma.

**Patient consent for publication** Not applicable.

**Ethics approval** This study involves human participants and these studies were done in accordance with the Declaration of Helsinki and the International Conference on Harmonisation on Guidance for Good Clinical Practice. Ethics approval was obtained from the relevant institutional review boards at participating sites. Participants gave informed consent to participate in the study before taking part.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Data availability statement** Data are available on reasonable request.

**Supplemental material** This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

**Open access** This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

#### ORCID iDs

Laure Gossec <http://orcid.org/0000-0002-4528-310X>  
 Ana-Maria Orbai <http://orcid.org/0000-0001-8644-8567>  
 Laura C Coates <http://orcid.org/0000-0002-4756-663X>  
 Dafna D Gladman <http://orcid.org/0000-0002-9074-0592>  
 Alexis Ogdie <http://orcid.org/0000-0002-4639-0775>  
 Christopher G Pelligra <http://orcid.org/0000-0002-5255-2777>  
 Maarten de Wit <http://orcid.org/0000-0002-8428-6354>

#### REFERENCES

- Walsh JA, Ogdie A, Michaud K, *et al*. Impact of key manifestations of Psoriatic arthritis on patient quality of life, functional status, and work productivity: findings from a real-world study in the United States and Europe. *Joint Bone Spine* 2023;90:105534.
- Ritchlin CT, Colbert RA, Gladman DD. Psoriatic arthritis. *N Engl J Med* 2017;376:2095–6.
- Coates LC, Orbai A-M, Azevedo VF, *et al*. Results of a global, patient-based survey assessing the impact of Psoriatic arthritis discussed in the context of the Psoriatic arthritis impact of disease (Psaid) questionnaire. *Health Qual Life Outcomes* 2020;18:173.
- Gossec L, de Wit M, Kiltz U, *et al*. A patient-derived and patient-reported outcome measure for assessing Psoriatic arthritis: elaboration and preliminary validation of the Psoriatic arthritis impact of disease (Psaid) questionnaire, a 13-country EULAR initiative. *Ann Rheum Dis* 2014;73:1012–9.
- Orbai A-M, Holland R, Leung YY, *et al*. Psaid12 provisionally endorsed at OMERACT 2018 as core outcome measure to assess Psoriatic arthritis-specific health-related quality of life in clinical trials. *J Rheumatol* 2019;46:990–5.
- Salaffi F, Di Carlo M, Carotti M, *et al*. The Psoriatic arthritis impact of disease 12-item questionnaire: equivalence, reliability, validity, and feasibility of the touch-screen administration versus the paper-and-pencil version. *Ther Clin Risk Manag* 2016;12:631–42.
- Di Carlo M, Becciolini A, Lato V, *et al*. The 12-item Psoriatic arthritis impact of disease questionnaire: construct validity, reliability, and Interpretability in a clinical setting. *J Rheumatol* 2017;44:279–85.
- Queiro R, Cañete JD, Montilla C, *et al*. Minimal disease activity and impact of disease in Psoriatic arthritis: a Spanish cross-sectional multicenter study. *Arthritis Res Ther* 2017;19:72.
- Walsh JA, Scher JU, Reddy SM, *et al*. Longitudinal construct validity of the Psaid individual items: can we eliminate other questionnaires if using the Psaid? *Arthritis Rheumatol* 2018;70:suppl.
- da Cruz Ribeiro e Souza E, da Silva Carneiro SC, Yazbek MA, *et al*. Validation and clinical Interpretability of Psaid - Psoriatic arthritis impact of disease. *Adv Rheumatol* 2020;60.
- Ogdie A, Michaud K, Nowak M, *et al*. Patient's experience of Psoriatic arthritis: a conceptual model based on qualitative interviews. *RMD Open* 2020;6:e001321.
- Lubrano E, Scriffignano S, Azuaga AB, *et al*. Assessment of the patient acceptable symptom state (PASS) in Psoriatic arthritis: association with disease activity and quality of life indices. *RMD Open* 2020;6:e001170.
- Holland R, Tillett W, Korendowych E, *et al*. Validation of the Psoriatic arthritis impact of disease (Psaid) questionnaire and its potential as a single-item outcome measure in clinical practice. *Ann Rheum Dis* 2018;77:343–7.
- FDA. Guidance for industry. patient-reported outcome measures: use in medical product development to support labeling claims. 2009. Available: <https://www.fda.gov/media/77832/download> [Accessed 10 Nov 2022].

- 15 FDA. Patient-focused drug development guidance public workshop. methods to identify what is important to patients & select, develop or modify fit-for-purpose clinical outcomes assessments. 2018. Available: <https://www.fda.gov/media/116277/download>
- 16 FDA. Patient-focused drug development guidance public workshop. incorporating clinical outcome assessments into endpoints for regulatory decision-making. 2019. Available: <https://www.fda.gov/media/132505/download>
- 17 McInnes IB, Asahina A, Coates LC, *et al.* Bimekizumab in patients with Psoriatic arthritis, naive to biologic treatment: a randomised, double-blind, placebo-controlled. *The Lancet* 2023;401:25–37.
- 18 Merola JF, Landewé R, McInnes IB, *et al.* Bimekizumab in patients with active Psoriatic arthritis and previous inadequate response or intolerance to tumour necrosis factor-A inhibitors: a randomised, double-blind. *The Lancet* 2023;401:38–48.
- 19 The EULAR Psoriatic Arthritis Impact of Disease: PsAID12 for clinical practice, Available: <http://pitie-salpetriere.aphp.fr/psaid/questionnaire.php> [Accessed 24 Apr 2023].
- 20 Felson DT, Anderson JJ, Lange ML, *et al.* Should improvement in rheumatoid arthritis clinical trials be defined as fifty percent or seventy percent improvement in core set measures, rather than twenty percent *Arthritis Rheum* 1998;41:1564–70.
- 21 Felson DT, Anderson JJ, Boers M, *et al.* The American college of rheumatology preliminary core set of disease activity measures for rheumatoid arthritis clinical trials. *Arthritis & Rheumatism* 1993;36:729–40.
- 22 Johnson SR, Hawker GA, Davis AM. The health assessment questionnaire disability index and scleroderma health assessment questionnaire in scleroderma trials: an evaluation of their measurement properties. *Arthritis Rheum* 2005;53:256–62.
- 23 McKenna SP, Doward LC, Whalley D, *et al.* Development of the Psaqol: a quality of life instrument specific to Psoriatic arthritis. *Ann Rheum Dis* 2004;63:162–9.
- 24 Cella D, Wilson H, Shalhoub H, *et al.* Content validity and Psychometric evaluation of functional assessment of chronic illness therapy-fatigue in patients with Psoriatic arthritis. *J Patient Rep Outcomes* 2019;3:32.
- 25 Mark EM. *User's Manual for the SF-36v2 Health Survey* 3rd edition. Lincoln, RI: QualityMetric, 2011.
- 26 Schoels MM, Aletaha D, Alasti F, *et al.* Disease activity in Psoriatic arthritis (PSA): defining remission and treatment success using the DAPSA score. *Ann Rheum Dis* 2016;75:811–8.
- 27 Helliwell PS, FitzGerald O, Fransen J, *et al.* The development of candidate composite disease activity and responder indices for Psoriatic arthritis (GRACE project). *Ann Rheum Dis* 2013;72:986–91.
- 28 Mease PJ. Measures of Psoriatic arthritis: tender and swollen joint assessment, psoriasis area and severity index (PASI), nail psoriasis severity index (NAPSI), modified nail psoriasis severity index (mNAPSI), Mander/Newcastle Enthesitis index (MEI), Leeds Enthesitis index (LEI), Spondyloarthritis research consortium of Canada (SPARCC), Maastricht Ankylosing Spondylitis Enthesis score (MASES), Leeds Dactylitis index (LDI), patient global for Psoriatic arthritis, Dermatology life quality index (DLQI), Psoriatic arthritis quality of life (Psaqol), functional assessment of chronic illness therapy-fatigue (FACIT-F), Psoriatic arthritis response criteria (Psarc), Psoriatic arthritis joint activity index (Psajai), disease activity in Psoriatic arthritis (DAPSA), and composite Psoriatic disease activity index (CPDAI). *Arthritis Care Res (Hoboken)* 2011;63 Suppl 11:S64–85.
- 29 Coates LC, Fransen J, Helliwell PS. Defining minimal disease activity in Psoriatic arthritis: a proposed objective target for treatment. *Ann Rheum Dis* 2010;69:48–53.
- 30 Hagège B, Tan E, Gayraud M, *et al.* Remission and low disease activity in Psoriatic arthritis publications: a systematic literature review with meta-analysis. *Rheumatology (Oxford)* 2020;59:1818–25.
- 31 Cohen J. *Statistical Power Analysis for the Behavioral Sciences* 2nd edition. Hillsdale, NJ: Lawrence Erlbaum Associates, 1988.
- 32 Coates LC, Gladman DD, Nash P, *et al.* Secukinumab provides sustained PASDAS-defined remission in Psoriatic arthritis and improves health-related quality of life in patients achieving remission: 2-year results from the phase III FUTURE 2 study. *Arthritis Res Ther* 2018;20:272.
- 33 Nunnally JC, Bernstein IH. *Psychometric theory* 3rd edition. New York: McGraw-Hill, 1994.
- 34 Cronbach LJ. Coefficient alpha and the internal structure of tests. *Psychometrika* 1951;16:297–334.
- 35 Aaronson N, Alonso J, Burnam A, *et al.* Assessing health status and quality-of-life instruments: attributes and review criteria. *Qual Life Res* 2002;11:193–205.
- 36 Cappelleri JC. *Patient-reported outcomes: measurement, implementation and interpretation*. Boca Raton, FL: CRC Press/ Taylor & Francis Group, 2014.
- 37 Hays RD, Farivar SS, Liu H. Approaches and recommendations for estimating minimally important differences for health-related quality of life measures. *COPD* 2005;2:63–7.
- 38 Revicki D, Hays RD, Cella D, *et al.* Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. *J Clin Epidemiol* 2008;61:102–9.
- 39 Schoels M. Psoriatic arthritis indices. *Clin Exp Rheumatol* 2014;32(5 Suppl 85):S–109.
- 40 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.
- 41 Fluss R, Faraggi D, Reiser B. Estimation of the Youden index and its associated cutoff point. *Biom J* 2005;47:458–72.
- 42 Hosmer DW, Lemeshow S. *Applied logistic regression*. In: . 2nd edition. New York: John Wiley & Sons, 13 September 2000.
- 43 Orbai A-M, Ogdie A. Patient-reported outcomes in Psoriatic arthritis. *Rheum Dis Clin North Am* 2016;42:265–83.
- 44 Gladman D, Nash P, Goto H, *et al.* Fatigue Numeric rating scale validity, discrimination and responder definition in patients with Psoriatic arthritis. *RMD Open* 2020;6:e000928.
- 45 de Wit MPT, Kvien TK, Gossec L. Patient participation as an integral part of patient-reported outcomes development ensures the representation of the patient voice: a case study from the field of rheumatology. *RMD Open* 2015;1:e000129.