Comparing the Patient-Reported Physical Function Outcome Measures in a Real-Life International Cohort of Patients With Psoriatic Arthritis

Ying Ying Leung, Ana-Maria Orbai, Maarten Wit, Andra Balanescu, Emmanuelle Dernis, Martin Soubrier, Lihi Eder, Josef S. Smolen, Laura C. Coates, Laure Gossec

HHS Public Access

Author manuscript

Arthritis Care Res (Hoboken). Author manuscript; available in PMC 2022 April 01.

Published in final edited form as:

Arthritis Care Res (Hoboken). 2021 April ; 73(4): 593–602. doi:10.1002/acr.24139.

Comparing the patient reported physical function outcome measures in a real-life international cohort of patients with psoriatic arthritis.

Ying Ying Leung, MB ChB, MD[1], **Ana-Maria Orbai, MD, MHS**[2], **Maarten de Wit, PhD**[3], **Andra Balanescu, MD, PhD**[4], **Emmanuelle Dernis, MD**[5], **Martin Soubrier, MD, PhD**[6], **Lihi Eder, MD, PhD**[7], **Josef S Smolen, MD**[8], **Laura C Coates, MB ChB, PhD**[9], **Laure Gossec, MD, PhD**[10,11], **The ReFlap Study Group**

[1.]Singapore General Hospital, Duke-NUS Medical School, Singapore [2.]John Hopkins University, Division of Rheumatology, Baltimore, MD, USA [3.]Patient Research Partner, Netherlands [4.]Sf Maria Hospital, University of Medicine and Pharmacy Carol Davila, Bucharest, Romania [5.]Le Mans Central Hospital, Le Mans, France [6.]Gabriel Montpied Hospital, Clermont Ferrand, France [7.]Women's College Hospital, University of Toronto, Toronto, ON, Canada [8.]Division of Rheumatology, Department of Medicine, Medical University of Vienna, Vienna, Austria [9.]Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, University of Oxford, Oxford, UK [10.]Sorbonne Université, INSERM, Institut Pierre Louis d'Epidémiologie et de Santé Publique, Paris France [11.]Pitié Salpêtrière hospital, AP-HP, Rheumatology department, Paris, France

Abstract

Objectives.—We evaluated the psychometric properties of three patient-reported outcomes (PROMs) to assess the physical function in psoriatic arthritis (PsA).

Methods.—Data available for Health Assessment Questionnaire-disability index (HAQ-DI), Physical component summary score of SF-12 (PCS12) and functional capacity of Psoriatic Arthritic Impact of Disease Instrument (PsAID-FC) from a longitudinal study in 14 countries of consecutive adults with definite PsA with ≥ 2 years of duration. The score distribution, construct validity, responsiveness and thresholds of meaning of the PROMs were evaluated.

Results.—At baseline, 414 subjects (52% male) were analysed. The mean (SD) age and duration of illness were 52.4 (12.5) and 10.9 (8.1) years. Ceiling effects were noted in 31% and 21% of patients for HAQ-DI and PsAID-FC; floor effects were minimal. All three PROMs met *a priori* hypotheses for construct validity. After a median (IQR) follow-up of 4.1 (2.7) months in 350 patients, 27%, 54% and 18% of patients reported themselves improved, not changed and worse, respectively. Change scores were statistically different for groups for worsening versus no-change

Correspondence to: Ying-Ying Leung, MB ChB, MD; Department of Rheumatology and Immunology, Singapore General Hospital, The Academia, level 4, 20 College Road, Singapore 169856, Contact No.: +65 63265276, Fax no.: +65 62203321, katyccc@hotmail.com.

Declaration

All authors have declared no conflict of interest.

for all PROMs. PsAID-FC was more sensitive to change than the other two PROMs. Comparing groups with worsening condition to no-change, the standardized response mean square ratios (SRM^2) were for HAQ-DI: 29.9, PCS12: 16.7 and PsAID-FC: 40.1, respectively.

Conclusions.—HAQ-DI, PCS12 and PsAID-FC are valid measures of function for PsA. PsAID-FC, a single question, performed similarly to the other PROMs and may be an additional option to measure PsA-specific physical function.

Keywords

Psoriatic arthritis; Physical function; Outcome measures; construct; responsiveness

INTRODUCTION

Psoriatic arthritis is a chronic inflammatory disease with diverse manifestations, including peripheral joints inflammation, dactylitis, enthesitis, spine inflammation, skin psoriasis and nail lesions. It has tremendous impact on patients' lives affecting multiple aspects (1, 2). Inflammation of joints leads to pain and loss of function, and structural damage resulting from PsA has been well recognized (3, 4). Many PsA patients suffer significant joint damage and disability over time (5, 6). For these reasons, physical function is an important outcome in PsA, and is one of the core domains to be monitored in every randomized controlled trial and longitudinal observational study for PsA (7, 8).

Several patient-reported outcomes (PROMs) that assess physical function in PsA (9, 10), have been used in clinical trials (11), although none were developed specifically for PsA. The Health Assessment Questionnaire – Disability index (HAQ-DI) has been the most commonly used PROM, followed by the physical component summary (PCS) and the physical functioning (PF) domain of Medical Outcome Survey 36-Item Short Form Instrument (SF-36). Apart from two studies that evaluated the construct validity of HAQ-DI and SF-36 PF (12, 13), the validity and responsiveness of different PROMs for the assessment of physical function in PsA have not been compared. The SF-12 is a short version of SF-36 and was recently proposed to more feasibly replace SF-36 in calculation of PsA composite indices, giving similar results as the original formula (14)

Physical function is also an important aspect of health-related quality of life (HRQoL). HRQoL is a multi-dimensional concept influenced by individuals' experiences of their illness, treatment, interacting with individuals' beliefs, expectations, culture and environment (15, 16). HRQoL is also one of the core domains to be measured for PsA (8). The PsA Impact of Disease (PsAID) is a multi-faceted instrument developed from the perspective of PsA patients to assess both the physical and psychological impact specifically for PsA (17) The domains of importance to PsA were derived from PsA patients and validated across 13 European countries. The Outcome Measure in Rheumatology (OMERACT) has recently provisionally endorsed PsAID as a measurement of health-related quality of life domain for PsA (18). The PsAID includes a numeric rating scale (NRS) for physical function (PsAID-Functional Capacity, FC). The psychometric properties of PsAID-FC for the assessment of physical function in PsA have not been evaluated. Individual

components of the PsAID have been previously suggested for use as single measures of PsA impact domains (19).

In this study, we aim to compare the score distribution, construct validity, known group validity (distinguishing patients with/ without remission), responsiveness and thresholds of meaning of three PROMs, HAQ-DI, PCS12 and PsAID-FC, for the assessment of physical function in PsA.

Methods and Materials

Participants

We used data from the Remission/Flare in PsA (ReFlaP) study (NCT03119805), which was a prospective longitudinal observational study in 14 countries of consecutive adults with physician diagnosed PsA with ≥ 2 years of disease duration (20). Ethics approval was obtained in each country or centre, and all patients signed informed consents prior to participation. Investigators were advised to consider the Classification Criteria for Psoriatic Arthritis (CASPAR) criteria for classification of PsA. Patients without definite PsA or < 2 years of disease duration, and patients who did not speak or read the local language were excluded. Patients were interviewed twice at baseline and a follow-up time point at 1-6 months apart, according to usual practice. The study did not include a specific intervention.

Data collection

Patient demographic variables including age, gender, work status, level of education, date of onset of PsA, and the current treatment (Yes/No to methotrexate, leflunomide, sulfasalazine, D-penicillamine, etanercept, adalimumab, infliximab, golimumab, certolizumab, secukinumab, ustekinumab, oral glucocorticoids, and others) were collected. Comorbidities were collected using the Functional Comorbidity Index (21). Physical examination included assessment of 66/68 swollen and tender joint count, Leeds enthesitis count (6-sites) and body surface area of psoriasis.

PROMs for physical function

Patients filled in their own language the following PROMs that assess physical function in paper and pencil format.

The Health Assessment Questionnaire – Disability Index (HAQ-DI)

The HAQ-DI was developed for rheumatoid arthritis (RA) (22), and adapted for use in PsA. It has 20 items assessing eight domains: dressing, rising, eating, walking, hygiene, reach, grip, and usual activities, with scores range from 0 none to 3 maximum disability. The threshold considered to represent minimally clinically important difference (MCID) for improvement in PsA RCTs is a score of 0.35 (23).

The Medical Outcome Survey 12-Item Short Form Instrument (SF-12)

The SF-12 is a generic, multipurpose survey with 12 questions selected from the SF-36 which results in two weighted summary scores of mental and physical (MCS and PCS) to represent overall HRQoL (24). The physical functioning summary score (PCS12) was

scored according to the method by the original authors (25). The SF-12 has not been formally validated in a PsA population and no MCID threshold has been defined for PCS12 in PsA (9).

The PsA Impact of Disease -12 (PsAID-12) Functional Capacity

The PsAID-12 (17) is a composite instrument developed to assess both the physical and psychological impact of PsA (17). The PsAID assesses health impacts attributed to PsA in the past one week. Within the PsAID-12, physical function was assessed in one item (PsAID-FC) as follows: *"Circle the number that best describes the difficulty you had in doing daily physical activities due to your psoriatic arthritis during the last week"*. The response is recorded on a 11-point NRS with 0 - no difficulty to 10 - extreme difficulty. No MCID threshold has been defined for the PsAID-FC.

Pain and global PROMs

Three PROMs for patient global assessment of arthritis (PGA-arthritis), pain and patient global assessment of skin (PGA-skin) were recorded on 11-point NRS with 0 – very good to 10 - very bad.

Global disease status change

At follow up visits, patients were asked: *"Think about all the ways your psoriatic arthritis has affected you during the last 48 hours. Compared to your last assessment with your rheumatologist, how did you feel during the last 48 hours?(improved/ no change/worse)".*

Patient reported disease status

At both visits, patients were asked about their own perception of PsA remission: *"At this time, is your psoriatic arthritis in remission, if this means: you feel your disease is as good as gone?"* (for Remission, REM) and *"At this time, are you in low disease activity, if this means: your disease is in low activity but it's not as good as gone?'* (for Low disease activity, LDA)*."* The wording for these items was developed with input from patient research partners (20). Patients were also asked if they were at a Patient Acceptable Symptom State (PASS) (26): *"If you were to remain for the next few months as you were during the last 48 hours, would this be acceptable or unacceptable for you? (acceptable/ unacceptable)".*

Composites measures of disease activity

Treatment targets for PsA have been proposed through a consensus exercise as Very Low/ Minimal Disease Activity or Disease Activity Index for PSoriatic Arthritis (DAPSA) remission/low disease activity (21, 27). For REM/LDA, two methods of definition were used: 1) DAPSA) cut-offs of ≤4/≤14 (28, 29) and 2) MDA)/ very low disease activity (VLDA) (30).

Psychometric evaluations

We evaluated the 3 PROMs for consistency in assessment of physical function according to the instrument selection algorithm outlined in the OMERACT Filter 2.1 (31, 32).

Score distribution.—The three physical function PROMs (HAQ-DI, PCS12 and PsAID-FC), were described for score distribution and proportion of missing data of the cohort. Floor effect (defined as worse possible score for physical capacity) and ceiling effect (defined as best possible score for physical capacity) were reported for each PROM individually.

Construct validity.—We assessed Spearman's rho correlations of PF PROMs having *a priori* hypotheses (Table 1). Spearman's rho <0.3, 0.3-0.5, 0.5-0.7, and >0.7 were considered very weak, weak, moderate and strong, respectively (33). We hypothesized physical function PROMs will correlate strongly with each other, and less strong (moderately to strongly) with disease activity, PGA-arthritis and the 68 tender joint count. We hypothesized weak-moderate correlations with the 66 swollen joint count, and very weak with PGA-skin. The expected direction and magnitude of correlation are summarized in Supplementary table 1. We also evaluated whether the PROMs for physical function may distinguish groups that are known to be different. The known groups evaluated were patient defined REM/LDA in comparison with the group at higher disease activity levels.

Longitudinal construct validity.—This psychometric property is traditionally known as responsiveness. We first evaluated the correlations between change scores of physical function PROMs with other PROMs. The correlations were expected to be less strong compared with those of construct validity, but with similar pattern (Supplementary Table 1). We evaluated the change of score in HAQ-DI, PCS12 and PsAID-FC, anchoring with patient defined change of condition at the follow-up visit (improved/no change/worse). We calculated the effect size (ES) by Cohen's d (mean change in scores / pooled standard deviation, SD) (34, 35) and standardized response mean (SRM) (mean change in scores / SD of the change scores)(35). The effect sizes are interpreted as 'trivial' (ES <0.20), 'small' (ES ≥0.20 and <0.50), 'moderate' (ES ≥0.50 and <0.80), or 'large' (ES ≥0.80) (36). We also calculated the relative effectiveness by dividing the square of the SRM of improved/worse group by the SRM of placebo group (SRM^2 ratio).

Threshold of meaning.—We evaluated several thresholds of meaning for interpretation. We evaluated the cut-offs for the three PROMs to predict the following outcomes with the external anchor:

1. Minimally clinically important difference (MCID) for improvement/ worsening, anchored with the group of patients endorsing *"improved/ worse"* to the change in condition question at follow-up visit. The mean change of scores from baseline to follow up for each PROM was taken as the MCID (37, 38).
2. Remission: VLDA and DAPSA REM.
3. At least achieving low disease activity (inclusive of patients in REM): MDA and DAPSA LDA.
4. Patient Acceptable Disease status (PASS), anchored with the group of patients endorsing "acceptable" for the PASS question.

For threshold of REM/LDA/PASS, the cut-off scores of the PROMs were established by two methods, first using the $25^{th}/75^{th}$ centile of scores (39); second is the Receiver Operator Curve (ROC) and the cut-off correspondent to the Youden's J index (40) for the group identified by the external anchors as mentioned. For MCID for worsening, the change scores of the PROMs from baseline to follow-up visits were used, whereas for other outcomes, data for baseline were calculated.

For all analysis, patients with missing data for a PROM were excluded from analysis of that particular PROM. Statistical analyses were conducted using IBM SPSS Statistics version 25 (41). The effect size analyses were conducted using Medcalc Statistical Software version 19.1 (42). All reported p values were two-sided, and $p < 0.05$ was considered statistically significant.

RESULTS

Of 466 patients recruited, 12 were ineligible (no confirmed PsA diagnosis, n=11; age below 18, n=1). Out of these 454 eligible patients, 394 patients were followed up after a median (interquartile range, IQR) of 4.1 (3.0, 5.7) months. Missing data for HAQ-DI, PCS12 and PsAID-FC at baseline were 1.3%, 7.5% and 0.2%, respectively. The respective figures for missing data at follow-up were 0.8%, 4.1% and 2.0%.

For all other analysis with baseline data, we included 414 patients who had complete data for all three PROMs at baseline. For analysis of the follow-up data, we included 350 patients who had complete data for all three PROMs at both time points. The characteristics of patients included in the analysis are summarized in Table 1. Half of these patients were men, mean (SD) age and disease duration at baseline were 52.4 (12.5) and 10.9 (8.1) years respectively. Patients came from 14 countries across four continents. Patients had moderate disease affecting joints, and mild disease affecting skin. At baseline, 40% and 58% fulfilled low disease activity criteria (including VLDA or remission) by MDA and DAPSA, respectively. 63% were taking conventional disease modifying anti-rheumatic drugs (DMARDs), and 61% were taking biological DMARDs.

Score distribution.

Score distribution of the three PROMs for physical function is summarized in Table 2. Ceiling effects were noted in a third of patients for HAQ-DI, and 20% for PCS12. Internal consistency measured by Cronbach's alpha for HAQ-DI was 0.92. Cronbach's alpha was not applicable for PCS12 and the single item PsAID-FC.

Construct validity.

The PROMs for physical function met *a priori* hypotheses (Table 3). As we hypothesized, the PROMs correlated strongly with each other. As expected, we found the highest correlation (rho >0.9) between PsAID-FC and PsAID-12 as the former is a component of the latter. The physical function PROMs correlated strongly with each other (rho > 0.7); and moderately to strongly with patient global assessments for arthritis (rho 0.61 to 0.78), pain (rho 0.61 to 0.77); moderately with tender joint count (rho 0.39 to 0.51) and DAPSA (rho 0.55 to 0.72); weakly with swollen joint count (rho 0.19 to 0.32); and very weakly with

patient global assessment for skin (rho 0.24 to 0.36). Average scores for all three physical function PROMs distinguished groups of patients in REM, LDA and PASS versus not achieving the respective status (all $p < 0.01$) (Supplementary Table 2).

Longitudinal construct validity.

The correlations between change scores of physical function PROMs and other measures were consistent with the hypothesized magnitude and direction (Supplementary Table 3). Similarly, the change scores for physical function PROMs strongly correlated with each other, moderately with disease activity indices (joint count, pain, patient global assessments for arthritis), and weakly with patient global assessment for skin.

Among the 350 patients who had follow-up and complete dataset for all three PROMs at both time points, 27%, 54% and 18% reported their condition had improved, did not change or get worse, respectively. Compared to the change scores of the group who reported no change, the change scores for all three PROMs in the worse group were statistically significantly different, but not for the change scores in the improved group (Table 4). All three PROMs for physical function were more sensitive for worsening than improvement. Moderate effect sizes were seen in all three PROMs in measurement of worsening. The SRM for worsening for HAQ-DI, PCS12 and PsAID-FC were 0.37, −0.45 and 0.38 respectively. Although the effect sizes estimations (Cohen's d, and SRM) were similar across the physical function PROMs, the relative effectiveness (SRM^2 ratio) was higher for PsAID-FC than the other two generic PROMs for physical function for worsening (Table 4).

Threshold of meaning.

The MCID for improvement/ worsening for HAQ-DI were −0.16 (SD: 0.87) for improvement, and 0.30 (SD: 0.81) for worsening; for PCS12 were 0.84 (SD: 14.2) for improvement, and −6.05 (SD: 13.4) for worsening; for PsAID-FC were −0.56 (SD: 4.08) for improvement, and 1.54 (SD: 4.01) for worsening (Table 4). The cut-offs for physical function PROMs that define patient-defined REM/LDA/PASS are presented in Table 5. Further information on REM/LDA by composites is given in Supplementary Table 4. The thresholds to define PASS and LDA were similar, and more stringent for REM. The respective cut-offs for patient defined REM/LDA were: 0.5-0.63/ 0.75 for HAQ-DI, 43.5/ 36.7-38.0 for PCS12, and 2.0/ 3.0-4.0 for PsAID-FC respectively.

DISCUSSION

Data from the current study supports the similar construct validity according to *a priori* hypothesis for PsAID-FC compared to HAQ-DI and PCS12 in PsA. The three PROMs distinguished groups hypothesized to be different: patient defined REM/LDA/PASS versus those not achieving these states. All three PROMs were more sensitive to worsening than improvement. The PsAID-FC was more sensitive to change compared with the other two generic PROMs. The cut-off threshold for REM/LDA/PASS were shown as reference. Among the three PROMs, PsAID-FC had the least missing data and supported the feasibility of PsAID-FC in clinical practice. PCS12 has the most missing data, while having no celing and floor effects. Compared to HAQ-DI, PsAID-FC had less ceiling effect in measurement.

Both physical function and HRQoL are within the core domain set to be measured in all clinical trials and observational studies (8). The World Health Organization (WHO) defines health as a "state of complete physical, mental, and social wellbeing, and not merely the absence of disease" (43), physical function has always been an integral component of HRQoL despite not encompassing the totality of it. Physical function in PsA has been assessed in observational studies and clinical trials using different measures. The HAQ-DI and SF-36 have been the most commonly used measures for physical function in clinical trials. The construct validity of HAQ-DI have been demonstrated in PsA (12, 13, 38, 44, 45). The HAQ-DI is the most widely used PROM for Physical function in clinical trials, and generally shows responsiveness to change and good discrimination between active treatments and placebo (11). The HAQ-DI was originally developed for RA (22) and later adapted to be a generic measurement of physical function for arthritis. Although PsA patients share some features with RA, there are differences in manifestations including more involvement of lower limbs, spine, enthesitis, dactylitis and skin that may not be captured by HAQ-DI. The concern for "domain match" of HAQ-DI to physical function in PsA has been raised. In a face-to-face discussion exercise followed by Delphi voting (both PsA experts and patient research partners), 56% of patients voted "uncertain" for use of HAQ-DI in PsA (46). Besides, the HAQ-DI has more pronounced ceiling effect in PsA patients than RA. In a study using Rasch model analysis, the ceiling effect of HAQ-DI was 30.4% in PsA compared to 6.9% in RA (12). In another study that compared several outcome measures for physical function in PsA, the ceiling effect of HAQ-DI in PsA stands high at 24% compared with 7.5% for the physical functioning domain of SF-36 (13). This ceiling effect makes HAQ-DI inappropriate for modern PsA clinical studies where physical function is being examined at a higher level of functioning than before targeted therapeutics were available (47).

The PF domain of SF-36 may have better construct validity including less ceiling effect, better fit to the Rasch model with higher item separation, longer measurement span and better measurement distribution (12, 13). However, to use the PF domain of SF-36, the full 36-items must be administered, making it difficult to be feasible in clinical practice. The SF-12 is a shorter version developed using normative data of SF-36 in the United States, to reproduce the two summary scores (MCS and PCS) (24). The ability to distinguish between different disease groups was less precise (25, 48), and has lower responsiveness compared to SF-36 (48). The SF-12 has not been formally validated in a PsA population (49). In addition, the PCS12 has a complicated HRQoL concept than purely physical function. It could be suitable for large epidemiology studies for comparing disease groups with population standards (norm) rather than in observational trials where improvement or worsening are more relevant. In the current study, we provided evidence to support the construct validity of PCS12 in PsA. The absence of ceiling and floor effects was desirable. However, the PCS12 had the most missing data in the whole cohort, as high as 7.5% at baseline. which would have an impact on its feasibility. Besides, the PCS12 was the least responsive, particularly for an improvement in PsA.

It has been envisioned that the PsAID-12 could serve as a multi-facet measurement of disease impact for PsA inclusive of the physical function aspect with great feasibility in observational trials or daily clinical practice (18). Although physical function is evaluated in

a single item in PsAID-12, the wording for assessment of physical function was framed in such a way that it is attributed to PsA per se. The data from this study provides evidence for the single item PsAID-FC in the measurement of physical function in PsA in setting of real-life cohort studies. All three PROMs were more sensitive to worsening than improvement. For instance, the MCID for HAQ-DI for improvement/ worsening were −0.16/ 0.3 which were consistent with previously reported in cohort study setting (38). A MCID for HAQ-DI for improvement of 0.3-0.35 was derived from previous trial settings (23). Although it is well known that sensitivity to change for PROMs can be different in different directions (32), an explanation for this perhaps lies in relative stable condition of the study patients (56% in LDA by DAPSA) with scores of physical function PROMs skewed towards minimal physical functional impairment end with higher ceiling effects. This limited the PROMs to show further sensitivity to change towards the good physical function end, and in fact change scores were not statistically significant at group level between patients who reported improvement versus stable status. The MCID for improvement may be more appropriately derived for its intended use from clinical trial settings where patients with active disease were expected to improve with a certain treatment. However, the MCID for deterioration derived from this cohort study setting would be a good guidance for stable patients who have deteriorated if there is a change score of +0.3 in HAQ-DI, −6.1 for PCS12 and +1.5 for PsAID-FC. Out of the three PROMs, the PsAID-FC showed slightly better sensitivity to change by the SRM^2 ratio comparing either improvement / worsening group to the group of patients endorsing no change to their condition. This is perhaps because of its disease specificity, where patients were asked to attribute their physical capacity due to PsA. Several methods of effect sizes estimation have been used in the current study and has resulted in similar estimation. The Guyatt's responsiveness have been proposed as a good alternative method for effect size estimation (35). When we calculated the Guyatt's effect size, similar results were obtained (data not shown).

One of the strengths of our study is the adherence to guideline set forth by the OMERACT in outcome measure evaluation (31, 32). It is the first study to compare the longitudinal validity and responsiveness of PROMs for physical function within the same cohort study design that involved large sample of PsA patients from 4 continents. We also provided the threshold or cut-off of meaning anchored by outcomes that are patient-defined. These included the REM/LDA/PASS. The wordings of these external anchors were developed with input from PsA patients, and therefore are relevant to patients. In the present study, the cutoffs changed slightly (within a small range) according to the different standard used. Although the definite cut-off threshold of meaning may require validation from other cohorts, it makes reasonable sense to recommend taking the tighter REM/LDA cut-offs of ≤ 0.5/ ≤0.75 for HAQ-DI; ≥44.0/ ≥38.0 for PCS12; and ≤2.0/ ≤3.0 for PsAID-FC respectively. There are a few limitations for interpretation of the current study. First, the patients in the study were recruited from highly specialized tertiary centers. They have definite PsA of more than 2-year duration. This may limit the generalizability to patients with milder or early disease. Data from PROMs administered in local languages was combined, although the regionally validated versions were encouraged. Translation of PROMs if necessary have been performed using standardized guideline (50) and protocol. The responsiveness for improvement of all three PROMs were relatively small compared to data derived from RCTs

(11, 23), but rather comparable to that derived from a longitudinal cohort study (38). This reflect the real situation in longitudinal cohort studies where most patients have been stabilized and not expected to improve over time. The MCID thresholds for improvement need to be interpreted with caution due to the high floor effect particularly true for HAQ-DI and PCS12.

CONCLUSION

In summary, we report results from an international real-life setting study, on the construct validity, responsiveness and thresholds of meaning of HAQ-DI, PCS12 and PsAID-FC to assess physical function in PsA. The PsAID-FC had less ceiling effect compared to HAQ-DI, exhibited similar construct validity to the other two generic instruments, and was slightly more responsive to change. PsAID-FC can be used to measure physical function in PsA at the same time as measuring life impact through administration of the PsAID questionnaire.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Funding

This study was funded by Pfizer through an unrestricted investigator-initiated grant.

Disclosures

YYL is funded by the Clinician Scientist award of the National Medical Research Council, Singapore (NMRC/CSA-INV/0022/2017). The views expressed are those of the author(s) and not necessarily those of the NMRC. AMO is a Jerome L. Greene Foundation Scholar and is supported in part by a research grant from the National Institute of Arthritis and Musculoskeletal and Skin Diseases (NIAMS) of the National Institutes of Health (NIH) under award number P30-AR070254. LCC is funded by a National Institute for Health Research Clinician Scientist award. The research was supported by the National Institute for Health Research (NIHR) Oxford Biomedical Research Centre (BRC). The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR, or the Department of Health.

REFERENCE

1. Strand V, Sharp V, Koenig AS, Park G, Shi Y, Wang B, et al. Comparison of health-related quality of life in rheumatoid arthritis, psoriatic arthritis and psoriasis and effects of etanercept treatment. Ann Rheum Dis. 2012;71(7):1143–50. [PubMed: 22258482]
2. Leung YY, Ho KW, Zhu TY, Tam LS, Kun EW, Li EK. Testing scaling assumptions, reliability and validity of medical outcomes study short-form 36 health survey in psoriatic arthritis. Rheumatology (Oxford, England). 2010;49(8):1495–501.
3. Gladman DD, Farewell VT, Nadeau C. Clinical indicators of progression in psoriatic arthritis: multivariate relative risk model. J Rheumatol. 1995;22(4):675–9. [PubMed: 7791162]
4. Gladman DD, Antoni C, Mease P, Clegg DO, Nash P. Psoriatic arthritis: epidemiology, clinical features, course, and outcome. Ann Rheum Dis. 2005;64 Suppl 2:ii14–7. [PubMed: 15708927]
5. Husted JA, Tom BD, Farewell VT, Schentag CT, Gladman DD. Description and prediction of physical functional disability in psoriatic arthritis: a longitudinal analysis using a Markov model approach. Arthritis Rheum. 2005;53(3):404–9. [PubMed: 15934101]
6. Leung YY, Ho KW, Li EK, Li M, Kwok LW, Wong PC, et al. Predictors of functional deterioration in Chinese patients with psoriatic arthritis: a longitudinal study. BMC Musculoskelet Disord. 2014;15:284. [PubMed: 25160684]

7. Gladman DD, Mease PJ, Strand V, Healy P, Helliwell PS, Fitzgerald O, et al. Consensus on a core set of domains for psoriatic arthritis. J Rheumatol. 2007;34(5):1167–70. [PubMed: 17477480]
8. Orbai AM, de Wit M, Mease P, Shea JA, Gossec L, Leung YY, et al. International patient and physician consensus on a psoriatic arthritis core outcome set for clinical trials. Ann Rheum Dis. 2017;76(4):673–80. [PubMed: 27613807]
9. Hojgaard P, Klokker L, Orbai AM, Holmsted K, Bartels EM, Leung YY, et al. A systematic review of measurement properties of patient reported outcome measures in psoriatic arthritis: A GRAPPA-OMERACT initiative. Semin Arthritis Rheum. 2018;47(5):654–65. [PubMed: 29037523]
10. Kalyoncu U, Ogdie A, Campbell W, Bingham CO 3rd, de Wit M, Gladman DD, et al. Systematic literature review of domains assessed in psoriatic arthritis to inform the update of the psoriatic arthritis core domain set. RMD Open. 2016;2(1):e000217. [PubMed: 26966554]
11. Mease P, Strand V, Gladman D. Functional impairment measurement in psoriatic arthritis: Importance and challenges. Semin Arthritis Rheum. 2018;48(3):436–48. [PubMed: 30029795]
12. Taylor WJ, McPherson KM. Using Rasch analysis to compare the psychometric properties of the Short Form 36 physical function score and the Health Assessment Questionnaire disability index in patients with psoriatic arthritis and rheumatoid arthritis. Arthritis Rheum. 2007;57(5):723–9. [PubMed: 17530670]
13. Leung YY, Tam LS, Kun EW, Ho KW, Li EK. Comparison of 4 functional indexes in psoriatic arthritis with axial or peripheral disease subgroups using Rasch analyses. The Journal of rheumatology. 2008;35(8):1613–21. [PubMed: 18597399]
14. Perruccio AV, Got M, Li S, Ye Y, Gladman DD, Chandran V. The development of a modified Psoriatic Arthritis Disease Activity Score (mPASDAS) using SF-12 as a measure of quality of life. Arthritis Care Res (Hoboken). 2019.
15. Testa MA, Simonson DC. Assessment of quality-of-life outcomes. The New England journal of medicine. 1996;334(13):835–40. [PubMed: 8596551]
16. Sajid MS, Tonsi A, Baig MK. Health-related quality of life measurement. International journal of health care quality assurance. 2008;21(4):365–73. [PubMed: 18785462]
17. Gossec L, de Wit M, Kiltz U, Braun J, Kalyoncu U, Scrivo R, 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(6):1012–9. [PubMed: 24790067]
18. Orbai AM, Holland R, Leung YY, Tillett W, Goel N, Christensen R, 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. 2018.
19. Holland R, Tillett W, Korendowych E, Cavill C, Waldron N, Brooke M, 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(3):343–7. [PubMed: 29146740]
20. Gorlier C, Orbai AM, Puyraimond-Zemmour D, Coates LC, Kiltz U, Leung YY, et al. Comparing patient-perceived and physician-perceived remission and low disease activity in psoriatic arthritis: an analysis of 410 patients from 14 countries. Ann Rheum Dis. 2019;78(2):201–8. [PubMed: 30442648]
21. Groll DL, To T, Bombardier C, Wright JG. The development of a comorbidity index with physical function as the outcome. J Clin Epidemiol. 2005;58(6):595–602. [PubMed: 15878473]
22. Fries JF, Spitz P, Kraines RG, Holman HR. Measurement of patient outcome in arthritis. Arthritis Rheum. 1980;23(2):137–45. [PubMed: 7362664]
23. Mease PJ, Woolley JM, Bitman B, Wang BC, Globe DR, Singh A. Minimally important difference of Health Assessment Questionnaire in psoriatic arthritis: relating thresholds of improvement in functional ability to patient-rated importance and satisfaction. J Rheumatol. 2011;38(11):2461–5. [PubMed: 21885498]
24. Ware J Jr., Kosinski M, Keller SD. A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. Med Care. 1996;34(3):220–33. [PubMed: 8628042]

25. Ware JE, Kosinski M, Keller SD. SF12: How to Score the SF-12 Physical and Mental Health Summary Scales. Second Edition ed. Boston, MA: The Health Institute, New England Medical Center; 1995.
26. Maksymowych WP, Richardson R, Mallon C, van der Heijde D, Boonen A. Evaluation and validation of the patient acceptable symptom state (PASS) in patients with ankylosing spondylitis. Arthritis Rheum. 2007;57(1):133–9. [PubMed: 17266072]
27. Smolen JS, Schols M, Braun J, Dougados M, FitzGerald O, Gladman DD, et al. Treating axial spondyloarthritis and peripheral spondyloarthritis, especially psoriatic arthritis, to target: 2017 update of recommendations by an international task force. Ann Rheum Dis. 2018;77(1):3–17. [PubMed: 28684559]
28. Schoels M, Aletaha D, Funovits J, Kavanaugh A, Baker D, Smolen JS. Application of the DAREA/DAPSA score for assessment of disease activity in psoriatic arthritis. Ann Rheum Dis. 2010;69(8):1441–7. [PubMed: 20525844]
29. Schoels MM, Aletaha D, Alasti F, Smolen JS. Disease activity in psoriatic arthritis (PsA): defining remission and treatment success using the DAPSA score. Ann Rheum Dis. 2016;75(5):811–8. [PubMed: 26269398]
30. Coates LC, Fransen J, Helliwell PS. Defining minimal disease activity in psoriatic arthritis: a proposed objective target for treatment. Ann Rheum Dis. 2010;69(1):48–53. [PubMed: 19147615]
31. Beaton DE, Maxwell LJ, Shea BJ, Wells GA, Boers M, Grosskleg S, et al. Instrument Selection Using the OMERACT Filter 2.1: The OMERACT Methodology. J Rheumatol. 2019.
32. Boers M, Kirwan JR, Tugwell P, Beaton D, Bingham III CO, Conaghan PG, et al. The OMERACT Handbook. 2017.
33. Hinkle DE, Wiersma W, Jurs SG. Applied Statistics for the Behavioral Sciences 5th ed. Boston; New York Houghton Mifflin Company, cop. ; 2003.
34. Kazis LE, Anderson JJ, Meenan RF. Effect sizes for interpreting changes in health status. Med Care. 1989;27(3 Suppl):S178–89. [PubMed: 2646488]
35. Husted JA, Cook RJ, Farewell VT, Gladman DD. Methods for assessing responsiveness: a critical review and recommendations. J Clin Epidemiol. 2000;53(5):459–68. [PubMed: 10812317]
36. Cohen J Statistical power analysis for the behavioural sciences. Rev. ed. ed: New York: Academic Press; 1977.
37. 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(2):102–9. [PubMed: 18177782]
38. 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(5):1024–8. [PubMed: 20231193]
39. Tubach F, Ravaud P, Baron G, Falissard B, Logeart I, Bellamy N, et al. Evaluation of clinically relevant changes in patient reported outcomes in knee and hip osteoarthritis: the minimal clinically important improvement. Ann Rheum Dis. 2005;64(1):29–33. [PubMed: 15208174]
40. Terwee CB, Mokkink LB, Knol DL, Ostelo RW, Bouter LM, de Vet HC. Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist. Qual Life Res. 2012;21(4):651–7. [PubMed: 21732199]
41. IBM SPSS Statistics for Windows. Armonk, NY: IBM Corp.; 2017.
42. MedCalc Statistical Software MedCalc Software bv, Ostend, Belgium.
43. Organization WH. Health Related Quality of Life (HRQOL). Preamble to the Constitution of WHO as adopted by the International Health Conference, New York, 19 June - 22 July 1946. 1946.
44. Blackmore MG, Gladman DD, Husted J, Long JA, Farewell VT. Measuring health status in psoriatic arthritis: the Health Assessment Questionnaire and its modification. J Rheumatol. 1995;22(5):886–93. [PubMed: 8587077]
45. Husted JA, Gladman DD, Cook RJ, Farewell VT. Responsiveness of health status instruments to changes in articular status and perceived health in patients with psoriatic arthritis. J Rheumatol. 1998;25(11):2146–55. [PubMed: 9818657]
46. Holland R, Tillett W, Ogdie A, Leung YY, Gladman DD, Callis Duffin K, et al. Content and Face Validity and Feasibility of 5 Candidate Instruments for Psoriatic Arthritis Randomized Controlled

Trials: The PsA OMERACT Core Set Workshop at the GRAPPA 2017 Annual Meeting. J Rheumatol Suppl. 2018;94:17–25. [PubMed: 29858348]

47. Allard A, Antony A, Shaddick G, Jadon DR, Cavill C, Robinson G, et al. Trajectory of radiographic change over a decade: the effect of transition from conventional synthetic disease-modifying antirheumatic drugs to anti-tumour necrosis factor in patients with psoriatic arthritis. Rheumatology (Oxford, England). 2019;58(2):269–73.
48. Rubenach S, Shadbolt B, McCallum J, Nakamura T. Assessing health-related quality of life following myocardial infarction: is the SF-12 useful? J Clin Epidemiol. 2002;55(3):306–9. [PubMed: 11864802]
49. Hojgaard P, Klokker L, Orbai AM, Holmsted K, Bartels EM, Leung YY, et al. A systematic review of measurement properties of patient reported outcome measures in psoriatic arthritis: A GRAPPA-OMERACT initiative. Semin Arthritis Rheum. 2017.
50. Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report measures. Spine. 2000;25(24):3186–91. [PubMed: 11124735]

SIGNIFICANCE AND INNOVATIONS

- We demonstrated the construct validity, responsiveness and thresholds of meaning of HAQ-DI, PCS12 and PsAID-FC to assess physical function in PsA.
- The PsAID-FC had less ceiling effect compared to HAQ-DI.
- PsAID-FC has better responsiveness than the other two generic PROMs.
- PsAID-FC can be used to measure physical function in PsA at the same time as measuring life impact through administration of the PsAID questionnaire.

Table 1.

Baseline characteristics of the psoriatic arthritis cohort, ReflaP study

	All eligible patients	Patients with complete dataset	
	Baseline (n= 454)	**Baseline (n = 414)**	**Follow-up (n= 350)**
Male (%)	50.4	51.5	51.0
Age, years¶	52.2 (12.6)	52.4 (12.5)	53.1 (12.4)
Duration of PsA, years¶	10.9 (8.4)	10.9 (8.1)	10.8 (8.8)
Continent (%)			
Asia	6.4	7.0	8.0
Europe	68.7	66.2	65.1
N. America	18.1	19.6	19.1
S. America	6.8	7.2	7.7
Education, years¶	12.3 (4.4)	12.3 (4.5)	12.2 (4.5)
Paid work (%)	56.2	56.6	54.6
Swollen joint count (0-66) ¶	2.20 (6.94)	1.30 (2.92)	1.26 (2.83)
Tender joint count (0-68) ¶	4.68 (9.46)	3.96 (7.80)	3.70 (6.85)
Body Surface area affected by psoriasis n (%)			
None	158 (36.1)	143 (34.9)	130 (37.1)
1-5%	236 (53.9)	221 (53.6)	176 (50.3)
6-20%	35 (8.0)	29 (7.1)	31 (8.9)
>20%	9 (2.1)	6 (1.5)	3 (0.9)
Patient global assessment for arthritis (0-10) ¶	4.20 (2.75)	4.06 (2.71)	3.99 (2.72)
Pain (0-10) ¶	4.12 (2.80)	4.02 (2.75)	3.83 (2.76)
Patient global assessment for skin (0-10) ¶	3.02 (2.87)	2.88 (2.75)	2.55 (2.56)
MDA (%)	37.7	40.4	39.4
DAPSA LDA (%)	56.4	58.1	55.7
Current cDMARDs (%)	62.4	62.9	59.2
Current bDMARDs (%)	61.1	61.3	61.2

¶mean (SD); PsA: psoriatic arthritis; MDA: minimal disease activity; DAPSA: Disease Activity Index for PSoriatic Arthritis; LDA: low disease activity criteria by Disease Activity Index for PSoriatic Arthritis; bDMARDs: biologics disease modifying anti-rheumatic drugs; cDMARDs: conventional disease modifying anti-rheumatic drugs.

Table 2.

Physical function PROMs score distribution

		n	Mean (SD)	Median (IQR)	Floor effect (%)	Ceiling effect (%)
HAQ-DI	Baseline	414	0.64 (0.68)	0.50 (0.00, 1.13)	0	31.2
	Follow-up	350	0.63 (0.64)	0.50 (0.00, 1.13)	0	28.6
PCS12	Baseline	414	41.4 (11.0)	42.1 (32.9, 51.0)	0	0
	Follow-up	350	41.2 (10.4)	42.2 (33.1, 50.3)	0	0
PsAID-FC	Baseline	414	3.54 (2.99)	3.00 (1.00, 6.00)	2.2	20.6
	Follow-up	350	3.57 (3.05)	3.00 (1.00, 6.00)	2.6	22.0
PsAID-12	Baseline	414	3.27 (2.38)	2.80 (1.29, 4.91)	0	2.2
	Follow-up	350	3.15 (2.33)	2.78 (1.20, 5.01)	0	3.7

SD: standard deviation; IQR: interquartile range; HAQ-DI: Health Assessment Questionnaire-Disability index; PCS12: Physical Functioning of Medical Outcome Short Form 12; PsAID-FC: Psoriatic Arthritic Impact of Disease Instrument (PsAID12)-Functional Capacity; PGA - joint: patient global assessment of joint condition; PGA – skin: patient global assessment of skin condition; DAPSA: Disease Activity Index for PSoriatic Arthritis; PsAID-12: 12-item Psoriatic Arthritic Impact of Disease Instrument.

Floor and ceiling effects were defined as scores representing maximal and minimal level of physical capacity by the PROM.

Arthritis Care Res (Hoboken). Author manuscript; available in PMC 2022 April 01.

Table 3.

Spearman's rho correlations between physical function PROMs with other measures

			Spearman's Rho		
		N	**HAQ-DI**	**PCS12**	**PsAID-FC**
HAQ-DI	Baseline	414	-	−0.788**	0.711**
	Follow-up	350	-	−0.751**	0.725**
PCS12	Baseline	414	−0.788**	-	−0.751**
	Follow-up	350	−0.751**	-	−0.754**
PsAID-FC	Baseline	414	0.711**	−0.751**	-
	Follow-up	350	0.725**	−0.754**	-
PGA - arthritis	Baseline	414	0.608**	−0.665**	0.784**
	Follow-up	350	0.609**	−0.666**	0.747**
Pain	Baseline	414	0.605**	−0.651**	0.757**
	Follow-up	350	0.612**	−0.647**	0.771**
PGA - skin	Baseline	414	0.228**	−0.241**	0.364**
	Follow-up	350	0.287**	−0.284**	0.350**
66 swollen joint count	Baseline	414	0.316**	−0.291**	0.292**
	Follow-up	350	0.256**	−0.191**	0.272**
68 tender joint count	Baseline	414	0.481**	−0.469**	0.472**
	Follow-up	350	0.446**	−0.391**	0.511**
DAPSA	Baseline	414	0.593**	−0.600**	0.687**
	Follow-up	350	0.603**	−0.554**	0.716**
PSAID12	Baseline	414	0.707**	−0.731**	0.902**
	Follow-up	350	0.724**	−0.721**	0.912**

** p<0.001; N = sample size; Rho = Spearman's correlation coefficient; HAQ-DI: Health Assessment Questionnaire-Disability index; PCS12: Physical Functioning of Medical Outcome Short Form 12; PsAID-FC: Psoriatic Arthritic Impact of Disease Instrument (PsAID12)-Functional Capacity; PGA - joint: patient global assessment of joint condition; PGA – skin: patient global assessment of skin condition; DAPSA: Disease Activity Index for PSoriatic Arthritis; PsAID-12: 12-item Psoriatic Arthritic Impact of Disease Instrument.

Table 4.

Responsiveness of the PROMs for physical function

	Baseline score ¶	Change score ¶	Responsiveness		
			Cohen's d (95% CI)	SRM (95% CI)	SRM^2 ratio ¥
HAQ-DI (0-3)					
Improved (n=95)	0.78 (0.75)	−0.16 (0.87)	−0.24 (−0.48, 0.04)	−0.19 (−0.38, 0.04)	7.37
No change (n=189)	0.59 (0.66)	−0.05 (0.81)	−0.08 (−0.27, 0.08)	−0.07 (−0.21, 0.07)	
Worse (n=64)	0.65 (0.64)	0.30 (0.81) *	0.46 (0.12, 0.78)	0.37 (0.10, 0.61)	27.9
PCS12 (0-100)					
Improved (n=95)	41.0 (10.5)	0.84 (14.2)	0.08 (−0.23, 0.34)	0.06 (−0.17, 0.24)	0.30
No change (n=189)	41.4 (11.5)	1.54 (13.5)	0.14 (−0.06, 0.31)	0.11 (−0.04, 0.25)	
Worse (n=64)	41.4 (10.7)	−6.05 (13.4) *	−0.57 (−0.92, −0.24)	−0.45 (−0.74, −0.19)	16.7
PsAID-FC (0-10)					
Improved (n=95)	3.66 (2.95)	−0.56 (4.08)	−0.19 (−0.46, 0.10)	−0.14 (−0.33, 0.08)	5.4
No change (n=189)	3.34 (2.97)	−0.23 (3.83)	−0.08 (−0.26, 0.11)	−0.06 (−0.20, 0.09)	
Worse (n=64)	4.05 (3.08)	1.54 (4.01) *	0.51 (0.16, 0.87)	0.38 (0.12, 0.66)	40.1

Cohen's d = mean change in scores / pooled standard deviation; SRM = mean change in scores / SD of the change scores.

¶ mean (SD);

¥ comparing to the no change group;

* p-values < 0.05 by Kruskal-Wallis test;

CI: confidence interval; SRM: standardized response mean; HAQ-DI: Health Assessment Questionnaire-Disability index; PCS12: Physical Functioning of Medical Outcome Short Form-12; PsAID-FC: Psoriatic Arthritic Impact of Disease Instrument (PsAID12)-Functional Capacity; SD = standard deviation; MDA = Minimal disease activity; SRM = standardized response means.

Table 5.

Thresholds of meaning for each physical function PROMs

		Patient defined REM (n=86)	**Patient defined LDA (n=245)**	**PASS (n=280)**
HAQ-DI	Median	0.00	0.25	0.13
	75th centile of scores ¶	0.50	0.75	0.63
	Youden's J index cut-off from ROC ¥	0.63	0.75	0.63
	sensitivity/specificity/AUC	0.88/ 0.47/ 0.71	0.79/ 0.55/ 0.69	0.76/ 0.72/ 0.81
PCS12	Median	50.1	46.4	46.9
	75th centile of scores ¶	43.5	36.7	38.3
	Youden's J index cut-off from ROC ¥	43.5	38.0	39.6
	sensitivity/specificity/AUC	0.76/ 0.61/ 0.71	0.73/ 0.59/ 0.68	0.73/ 0.78/ 0.81
PsAID-FC	Median	1.00	2.00	2.00
	75th centile of scores ¶	2.00	4.00	4.00
	Youden's J index cut-off from ROC ¥	2.00	3.00	3.00
	sensitivity/specificity/AUC	0.79/ 0.64/ 0.76	0.70/ 0.68/ 0.74	0.73/ 0.84/ 0.85

¶: the 75th centile of scores in the group as defined by the external anchor;

¥: The correspondent cut-off from the Youden's J index from the Receiver Operator Curve.

PASS: patient acceptable symptom state; MDA: minimal disease activity; LDA: low disease activity; DAPSA: Disease Activity Index for PSoriatic Arthritis; VLDA: very low disease activity; REM: remission; LDA: low disease activity; HAQ-DI: Health Assessment Questionnaire-Disability index; PCS12: Physical Functioning of Medical Outcome Short Form 12; PsAID-FC: Psoriatic Arthritic Impact of Disease Instrument (PsAID12)-Functional Capacity; AUC: area under curve of the Receiver Operator Curve.