

ORIGINAL RESEARCH

Assessing internal construct validity of DAPSA and DAPSA28 in psoriatic arthritis: a European observational study using confirmatory factor analysis and additional psychometric testing

Jacob Brauner Jørgensen ^{1,2}, Karl Bang Christensen,³ Brigitte Michelsen,^{1,4,5} Anne Gitte Loft,^{6,7} Simon Horskjær Rasmussen,^{1,8} Jette Heberg,¹ Merete Lund Hetland ^{1,9}, Jakub Zavada,¹⁰ Karel Pavelka,^{10,11} Florenzo Iannone ¹², Fabrizio Conti,¹³ Joana Borges,¹⁴ Catalin Codreanu,¹⁵ Corina Mogosan,¹⁵ Bente Glintborg ^{1,9,16}, Adrian Ciurea ¹⁷, Michael J Nissen ¹⁸, Dan Nordström,¹⁹ Laura Kuusalo,²⁰ Karin Laas,^{20,21} Sigrid Vorobjov,²² Bjorn Gudbjornsson ^{23,24}, Thorvardur Jon Love,^{25,26} Mikkel Østergaard ^{1,9}, Lykke Midtbøll Ørnberg ¹

To cite: Jørgensen JB, Christensen KB, Michelsen B, *et al.* Assessing internal construct validity of DAPSA and DAPSA28 in psoriatic arthritis: a European observational study using confirmatory factor analysis and additional psychometric testing. *RMD Open* 2025;11:e006104. doi:10.1136/rmdopen-2025-006104

► Additional supplemental material is published online only. To view, please visit the journal online (<https://doi.org/10.1136/rmdopen-2025-006104>).

MØ and LMØ contributed equally.

This work was originally presented at EULAR 2025 (Barcelona, Spain).³³

Received 7 July 2025
Accepted 15 November 2025



© Author(s) (or their employer(s)) 2025. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ Group.

For numbered affiliations see end of article.

Correspondence to

Dr Lykke Midtbøll Ørnberg; lykke.midtboell.oernberg@regionh.dk

ABSTRACT

Objectives The Disease Activity index for Psoriatic Arthritis (DAPSA) was developed to assess disease activity in patients with psoriatic arthritis (PsA). A modified version, DAPSA28, uses 28 joints instead of 66/68. This study evaluated key psychometric properties of DAPSA and DAPSA28.

Methods Data from 1865 patients with PsA in the European Spondyloarthritis (EuroSpA) Research Collaboration Network, having DAPSA and DAPSA28 scores at baseline and follow-up, were analysed. Tests included assessment of internal construct validity by scree plots, confirmatory factor analysis (CFA) and structural equation modelling (SEM), supplemented by tests of differential item functioning (DIF) and evaluation of internal consistency reliability by Cronbach's α (CA). A subset of 625 patients was used for most analyses, except descriptive statistics, correlation matrix and CA.

Results One-dimensional CFA models for DAPSA and DAPSA28 showed acceptable model fit at baseline (root mean square error of approximation, RMSEA: 0.020, 0.034). However, model fit at 6 months follow-up was poor (RMSEA: 0.057, 0.063). SEM combining baseline and follow-up data could not identify an acceptable model fit. DIF was found for sex and country. CA indicated acceptable internal consistency (DAPSA: 0.65; DAPSA28: 0.63). Heterogeneity across countries was observed.

Conclusions Overall, the model fit was acceptable across model fit statistics, supporting internal construct validity, but some evidence of misfit at country level was disclosed. Our findings support acceptable internal consistency reliability, but DIF was found for sex and country. Based on mixed results of model fit and DIF, further investigation of these and other PsA disease activity measures is warranted.

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ Previous research in psoriatic arthritis (PsA) demonstrates that Disease Activity index for Psoriatic Arthritis (DAPSA) and DAPSA28 correlate well with other disease activity measures, such as the Psoriatic Arthritis Disease Activity Score (PASDAS) for PsA and the Disease Activity Score 28 (DAS28), and are sensitive to change, supporting their external construct and criterion validity.

WHAT THIS STUDY ADDS

⇒ This study evaluated whether DAPSA and DAPSA28 accurately measure the latent construct of PsA disease activity, the differential item functioning and Cronbach's α . Confirmatory factor analysis indicated acceptable internal construct validity at baseline and at 6 months follow-up, while structural equation models across both time points did not demonstrate acceptable fit. Both indices conformed to the standard tests of internal consistency based on Cronbach's α . Thus, internal construct validity and internal consistency reliability were supported, and differential item functioning was identified.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ Further studies comparing DAPSA and DAPSA28 and other existing PsA and inflammatory arthritis disease activity measures recommended by key international bodies, such as American College of Rheumatology (ACR), European Alliance of Associations for Rheumatology (EULAR), Group for Research and Assessment of Psoriasis and Psoriatic Arthritis / Outcome Measures in Rheumatology (GRAPPA/OMERACT), are encouraged. Assessment of internal construct validity of several indices and country-wise comparison would be useful.

INTRODUCTION

Psoriatic arthritis (PsA) is a chronic inflammatory joint disease with manifestations in peripheral joints, the axial skeleton, skin, nails, entheses and digits. Patients present with pain and stiffness in affected joints, functional disability and fatigue.¹ PsA is also associated with comorbidities, including cardiovascular disease, diabetes mellitus, metabolic syndrome, obesity, inflammatory bowel disease and depression.²

To effectively monitor disease activity and treatment response in PsA, regular monitoring using validated assessment tools is recommended.³ One such outcome measure is the Disease Activity index for Psoriatic Arthritis (DAPSA). DAPSA was designed in 2010 based on testing the Disease Activity index for the assessment of Reactive Arthritis (DAREA) designed for reactive arthritis assessment⁴ on PsA subsamples of 180 and 99 patients.⁴ It has since been used globally to assess PsA disease activity.⁵ The DAPSA score includes the sum of the count of tender (0–68) and swollen (0–66) joints, C reactive protein (CRP) levels (mg/dL) and patient-reported global assessments of disease activity (0–10) and pain (0–10). In 2018, a 28-joint version, DAPSA28, was introduced for cases where the full 66/68 joint count is unavailable, with joint counts multiplied by a conversion factor of 1.6.⁶

External construct validity is tested by relations between a given index and other indices, sometimes based on correlation. Existing evidence supports DAPSA's external construct validity through correlations with other disease activity measures (Psoriatic Arthritis Disease Activity Score for PsA, clinical DAPSA, Clinical Disease Activity Index for rheumatoid arthritis, Simplified Disease Activity Index for rheumatoid arthritis, Disease Activity Score 28 CRP for rheumatoid arthritis, and DAPSA28 for PsA), concurrent validity, discriminant validity, test–retest and principal component analysis.^{4–9} Similarly, DAPSA28 has been validated through tests of external construct validity (same indices as with DAPSA), sensitivity to change and discriminant power.⁶ DAPSA is recommended for outpatient clinical treatments^{10 11} and DAPSA28 has been suggested for use in research when 66/68 joint count is not available,⁶ but can they be considered valid disease activity measures? Even though several aspects have been tested, this has not been thoroughly investigated. This could mean that what we measure when using these indices is not valid, leading to misjudgements and possibly not the best treatment of patients or the best assessments in research. Therefore, we investigate their psychometric properties.

Psychometric validation evaluates the measurement properties of indices, such as outcome measures, using statistical methods. Doing this based on structural equation models (SEM¹²) and confirmatory factor analysis (CFA¹³) models is considered standard practice for exploring properties of an instrument that indirectly measures a variable that cannot be directly observed. These approaches are considered assessments of internal

construct validity, that is, assessments of the construct itself unrelated to other constructs. A variable that cannot be directly observed is often called a latent variable, and PsA disease activity is an example of this. The SEM and CFA approaches test whether combinations of observable variables (such as CRP, joint counts and patient global assessment included in DAPSA) can accurately measure the latent construct of PsA disease activity. Modern test theory like SEM and CFA tests the fit between observable variables and the model proposed by index developers, that is, the DAPSA and DAPSA28. These statistical models test ideal measurement requirements of an index against observed data.

While the external construct validity for DAPSA and DAPSA28 is well-established, fewer studies have investigated their internal construct validity, particularly through modern test theory. Therefore, the primary aim of this study was to assess the internal construct validity of DAPSA and DAPSA28 using CFA and SEM, alongside assessments of additional psychometric properties.

METHODS

Material

The EuroSpA Research Collaboration Network is a scientific consortium comprising 17 European clinical arthritis registries that collect data on patients with SpA, including axial SpA (axSpA) and PsA.¹⁴ Each registry gathers comprehensive clinical data pertinent to the management of patients with PsA (www.eurospa.eu). This study included a cohort of 1863 adult patients (>18 years) diagnosed with PsA (1964–2021) who initiated their first biologic or targeted synthetic disease-modifying antirheumatic drug (b/tsDMARD) between 1 January 2015 and 16 July 2022. This start date of initial treatment is considered the baseline date in this study. They had available DAPSA and DAPSA28 scores at a baseline and 6-month follow-up visit, allowing for analysis of time-dependent construct validity. Patients missing either baseline or follow-up DAPSA or DAPSA28 assessments were excluded. The dataset included demographic and clinical variables, including sex, country, country's gross domestic product (GDP) per capita,¹⁵ body mass index, disease duration, DAPSA and DAPSA28 scores, and individual DAPSA/DAPSA28 items: 28 and 68 Tender Joint Counts (TJC), 28 and 66 Swollen Joint Counts (SJC), CRP, Patient Global Score (PGS), Patient Pain Score (PPS) and Physician Global Score.

Psychometric tests and statistics

We evaluated the psychometric validity and reliability of DAPSA and DAPSA28 following OMERACT guidelines¹⁶ and the American Psychological Association Standards for psychometric tests.¹⁷ This evaluation of construct validity had six steps, as outlined below.

Relationship between items

First, using data from all 1863 patients, we performed polychoric rank correlation analysis for items of different

scales between each of the DAPSA and DAPSA28 items at baseline and follow-up, and physician global score, as well as pairwise scatter plots and histograms for each item to map the overall pattern of relationships between items. Correlations were presented as a heatmap. Correlations above 0.90 were considered strong, suggesting measurement of related traits.¹⁸

Dimensionality

To ensure that the identified models were not biased by variation in patient numbers between countries, 125 patients were randomly selected from each of the five countries who had the highest number of patients. This resulted in a total of 625 patients which allows for the analysis of scree plot, CFA, SEM and differential item functioning (DIF) assessments. Scree plots were used to evaluate whether most part of the variances in the DAPSA and DAPSA28 scores were caused by a single one-dimensional underlying variable, that is, PsA disease activity. Scree plots including full sample and resampled subsamples were assessed to test for sampling sensitivity and to evaluate dimensionality visually, by identifying a point at which the proportion of variance explained by each subsequent dimension (principal component) drops off, often referred to as a ‘knee’ or ‘elbow’ effect because an angle is seen in the plotted line. A significant drop between the first and second dimensions suggests that a second dimension should not be included. Dimensionality can also be based on an eigenvalue 1 cut-off, meaning that dimensions with a higher score than 1 are supported in the data. Both approaches were considered in this analysis. Hypothetically, if two dimensions are identified, these could be an ‘experienced disease’ based on patient pain and global assessment and a ‘physical impact’ based on the rest of the items, depending on how the items loaded on the dimensions. A specified multidimensional alternative could not be tested because DAPSA and DAPSA28 include only five items, whereas six are required for a multidimensional model (three per dimension). However, failure of a one-dimensional model to fit the data would indicate multidimensionality.

Internal construct validity

Evaluation of polychoric correlation scree plots informed the choice of a CFA model based on the dimensionality assessment, and the fit of this model was tested using the χ^2 goodness-of-fit statistic that directly checks how well the model matches the data, known as a test of exact fit. For exact fit, χ^2 , df and significance by p value were reported. χ^2 is high for poor fit and low for good fit, df allows comparison of constraints between each of the models, and p values above 0.05 suggest a good fit. Larger sample sizes will very often lead to an analysis rejecting that the models fit observed data and for this reason the χ^2 goodness-of-fit analysis was supplemented by two indices of close fit: the comparative fit index (CFI;

acceptable model fit if above 0.96) and the root mean square error of approximation (RMSEA; acceptable model fit if below 0.06).¹⁹ When one of these measures was acceptable and the other was not, RMSEA was chosen to judge model fit, as it is considered more suitable for this purpose.²⁰ For each item in each model, the eigenvalue was also reported. High eigenvalues over 1 between the manifest item (such as the joint count) and the latent construct (such as DAPSA28) suggest a stronger relation, whereas lower values suggest a weaker relation. A SEM using combined data from baseline and 6 months’ follow-up was used to test factor structure patterns over time. Two-tier models were used since residual variance for an item was assumed to be correlated across the two time points.

The fit of CFA models and of SEMs was tested in separate analyses for each of the countries with more than 125 respondents and in multiple groups analyses combining data from these countries assuming configural invariance.²¹ Briefly, configural invariance means that the same factor structure, for example, a single one-dimensional latent variable, is assumed in each of the countries with more than 125 respondents, but that no equality constraints are placed on the parameters. Thus, the analysis combines evidence of model fit across countries, but it does not test if the parameters are the same.

Differential item functioning

To determine if different subgroups of patients would have different probabilities of DAPSA and DAPSA28 disease activity scores, DIF was applied. DIF tests whether patient characteristics influence the likelihood of obtaining high or low scores. DIF was stratified by national registries/countries, GDP levels and sex. Significance was assessed using the Mantel-Haenszel test, with p values less than 0.05 indicating the presence of DIF.

Internal consistency reliability

For each index (DAPSA and DAPSA28) internal consistency reliability was estimated using Cronbach’s coefficient α (CA), which implies that test–retest reliability would be the same or higher.²² CA was considered acceptable above 0.6.²³

Sensitivity analysis

CFA, SEM and DIF were analysed for each country and with and without removal of DAPSA and DAPSA28 change-score outliers.

RESULTS

The study included 1863 patients, of which 55% were women. Patients had a median disease duration of 4 years (table 1). Five countries (Czech Republic, Denmark, Italy, Portugal and Romania) had data from 125 patients or more. DAPSA and DAPSA28 scores for each of these five countries are shown in online supplemental table S1. In 5% of patients, the 28 TJC was 0, while the 68 TJC was

Table 1 Descriptive statistics of cohort, demographics, DAPSA and DAPSA28 scores and item scores at baseline and 6 months follow-up

	Cohort (n=1863)	
	Available data, n (%)	Median (IQR) or n (%)*
Demographics		
BMI	1584 (85)	27.5 (24.2, 31.2)
Country		
Czech Republic	–	786 (42%)
Italy	–	499 (27%)
Portugal	–	217 (12%)
Romania	–	159 (8.5%)
Denmark	–	129 (6.9%)
Switzerland	–	40 (2.1%)
Finland	–	23 (1.2%)
Estonia	–	6 (0.3%)
Iceland	–	4 (0.2%)
Sex		
Female	–	1023 (55%)
Male	–	840 (45%)
Age at baseline/treatment initiation	1546 (83)	51 (42, 58)
Treatment		
Tumour necrosis factor inhibitor	–	1464 (79%)
Interleukin 17 inhibitor	–	277 (15%)
Phosphodiesterase-4 inhibitor	–	72 (3.9%)
Interleukin 12/23 inhibitor	–	39 (2.1%)
Janus Kinase inhibitor	–	11 (0.6%)
Disease duration in years	1546 (83) (100)	4 (1, 9)
Year of diagnoses		
–1999	–	35 (2.3%)
1990–1999	–	76 (4.9%)
2000–2009	–	305 (20%)
2010–2019	–	1043 (67%)
2020–	–	87 (5.6%)
Clinical measures, baseline		
Tender Joint Count, 68	1863 (100)	9 (3, 16)
Tender Joint Count, 28	1863 (100)	6 (2, 11)
Swollen Joint Count, 66	1863 (100)	5 (1, 9)
Swollen Joint Count, 28	1863 (100)	3 (0, 7)
CRP, mg/L	1863 (100)	8 (3, 19)
Patient Global Score 0–100	1863 (100)	70 (50, 80)
Patient Pain Score 0–100	1863 (100)	70 (50, 80)
Physician Global Score 0–100	1817 (100)	52 (30, 70)
Clinical measures, 6 months follow-up		
Tender Joint Count, 68	1863 (100)	1 (0, 3)
Tender Joint Count, 28	1863 (100)	0 (0, 2)
Swollen Joint Count, 66	1863 (100)	0 (0, 1)
Swollen Joint Count, 28	1863 (100)	0 (0, 1)
CRP, mg/L	1863 (100)	2.3 (1.0, 5.4)

Continued

Table 1 Continued

	Cohort (n=1863)	
	Available data, n (%)	Median (IQR) or n (%)*
Patient Global Score, 0–100	1863 (100)	30 (10, 50)
Patient Pain Score, 0–100	1863 (100)	30 (10, 50)
Physician Global Score, 0–100	1782 (96)	12 (5, 30)
Composite measures		
DAPSA68, baseline	1863 (100)	29 (18, 41)
DAPSA28, baseline	1863 (100)	29 (18, 44)
DAPSA68, 6 months	1863 (100)	9 (4, 15)
DAPSA28, 6 months	1863 (100)	9 (4, 15)

*Median (IQR=25% percentile–75% percentile); n (%). BMI, body mass index; CRP, C reactive protein; DAPSA, Disease Activity Index for Psoriatic Arthritis.

>0. Similarly, 4% of patients had swollen joints assessed by the 66 joint count, with no joints registered as swollen with the 28 swollen joint count. Thus, some articular involvement captured by DAPSA was not captured by DAPSA28.

DAPSA and DAPSA28 validity
Relationship between items

Inspection of marginal item distribution histograms and pairwise scatter plots revealed skewed distributions and non-linear inter-item relations (online supplemental figure S2). At baseline, a positive correlation was generally observed between DAPSA and DAPSA28 items and sum scores (see online supplemental figure S2 and figure 1 below). Item correlations ranged from –0.14 (for SJC at baseline and PGS at follow-up) to 0.99 (DAPSA and DAPSA28 at follow-up) for DAPSA and DAPSA28 items and sum scores. The

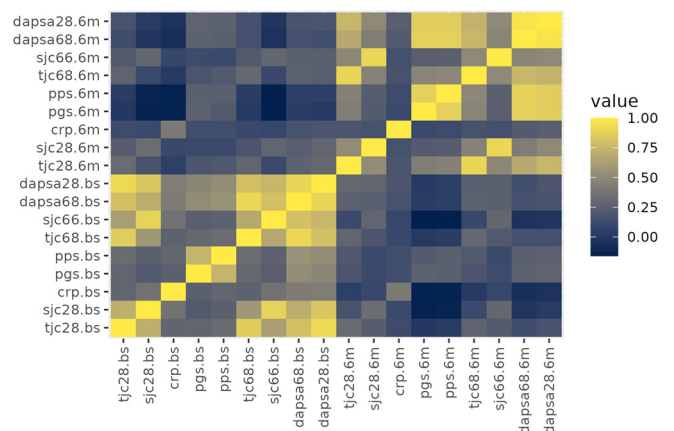


Figure 1 Polychoric rank correlation heat map for DAPSA and DAPSA28 total scores and standardised DAPSA and DAPSA28 individual items at baseline and follow-up. Each item has a suffix for baseline ('bs') or for 6 months follow-up (6m). For tjc, sjc and total scores, names have '28' for 28 joints and '66' or '68' for 66/68 joints. CRP, C reactive protein; DAPSA, Disease Activity Index for Psoriatic Arthritis; PGS, Patient Global Score; PPS, Patient Pain Score; SJC, Swollen Joint Count; TJC, Tender Joint Count.

RMD Open: first published as 10.1136/rmdopen-2025-006104 on 23 December 2025. Downloaded from https://rmdopen.bmj.com on 19 January 2026 by guest. Protected by copyright, including for uses related to text and data mining, AI training, and similar technologies.

correlation between 68 TJC (TJC68) and 28 TJC (TJC28) at baseline was 0.89 and the correlation between swollen joint count 66 (SJC66) and swollen joint count 28 (SWC28) was 0.91. The correlation between DAPSA and DAPSA28 total score at baseline was 0.94, indicating a high degree of correlation. For all DAPSA and DAPSA28 items and total scores at baseline, correlations ranged from 0.22 to 0.92. The correlation between DAPSA and DAPSA28 at 6 months' follow-up was 0.99, that is, higher than between DAPSA and DAPSA28 at baseline. The correlations for each of the DAPSA items measured at baseline vs the same measure at follow-up were between 0.02 (PGS) and 0.39 (TJC).

Dimensionality

Figure 2 displays scree plots for DAPSA and DAPSA28 at baseline and 6 months follow-up for the randomly selected 625 patients. Online supplemental figure S3 presents scree plots for each of the five countries with more than 125 patients. The difference between the line with a dot for each component illustrates eigenvalues and the solid line cut-off at eigenvalue 1 illustrates multidimensionality in baseline data. Based on an eigenvalue cut-off of 1, the scree plots suggest two dimensions, and based on the knee, they suggest that one dimension is present. Since testing of multidimensionality was not possible due to the number of items, models were analysed as one-dimensional.

Internal construct validity

The DAPSA one factor single group CFA model at baseline showed eigenvalues ranging from 0.51 (CRP) to 0.63 (TJC), with item residual variances between 0.61 (TJC) and 0.74 (CRP) (figure 3a). The DAPSA28 CFA model at baseline was similar, with eigenvalues ranging

from 0.47 (PGS) to 0.96 (TJC) and item residual variances between 0.08 (TJC) and 0.96 (CRP) (figure 3b). At 6 months follow-up compared with the baseline CFA, there was a decreased and negative (−0.01) relation between tender and swollen joint count in DAPSA and DAPSA28, and eigenvalues varied more. The SEM model (online supplemental figure S4a) between DAPSA at baseline and DAPSA at 6 months' follow-up showed similar eigenvalues as the baseline CFA models and residual variances between baseline and follow-up, except lower residual values for PGS and PPS at 6 months follow-up. Latent factor covariance between DAPSA at baseline and follow-up was −0.43 and −0.48, that is, low, suggesting that patients with high disease activity at baseline had higher improvement. The DAPSA28 SEM model (online supplemental figure S4b) between baseline and follow-up also found eigenvalues and residual variance like the DAPSA28 CFA model. Covariance between the latent factor for DAPSA28 at baseline and follow-up was the same as for DAPSA.

Fit indices

For DAPSA and DAPSA28, CFA models testing configural invariance—which checks whether the index functions the same way across groups—showed good fit at baseline, although the p-value for DAPSA28 was borderline significant. The p values at 6 months follow-up were below 0.05 and χ^2 values higher than at baseline, suggesting poor fit (see table 2). Acceptable values for the indices of close fit were seen for DAPSA, while for DAPSA28 one of the indicators was just outside the range of acceptable values. For the SEM models, the χ^2 test and the RMSEA rejected model fit (values 0.079 and 0.082).

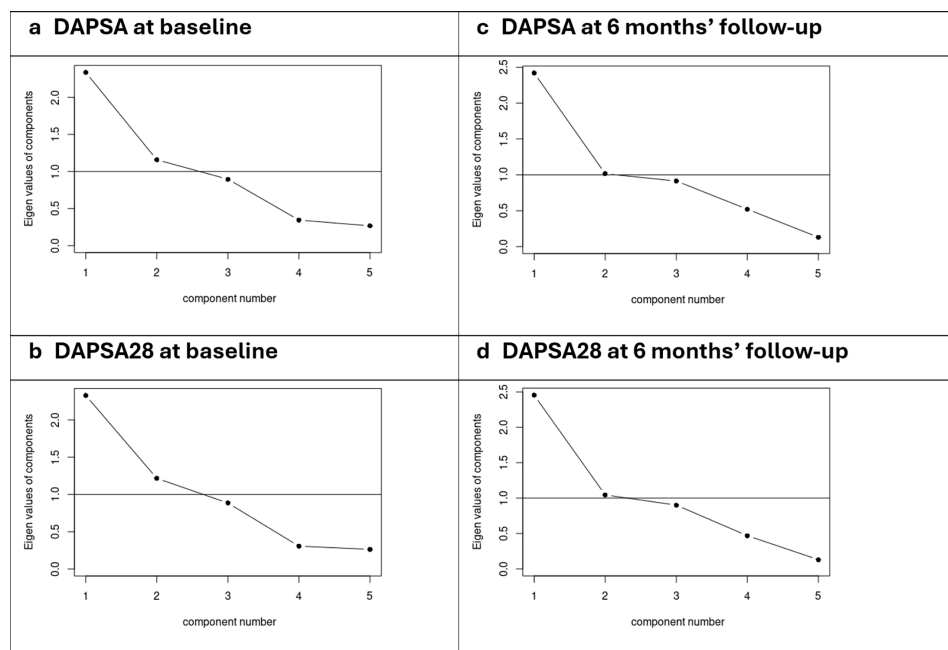


Figure 2 Scree plots for DAPSA and DAPSA28 at baseline and follow-up. The dots are the eigenvalues for each of the components, the solid line is the cut-off at eigenvalue 1. DAPSA, Disease Activity index for Psoriatic Arthritis.

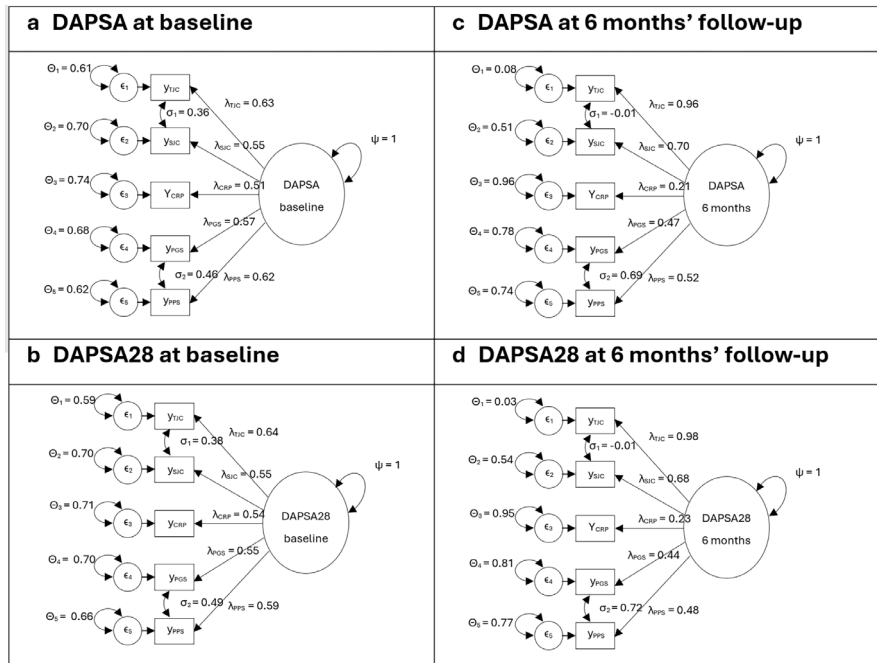


Figure 3 Baseline DAPSA and DAPSA28 confirmatory factor analysis model. λ (lambda) indicates factor loadings, Θ (theta) indicates residual variances, σ (sigma) indicates covariance between error terms, ψ (psi) indicates factor variance and φ (phi) indicates latent variable covariance. CRP, C reactive protein; DAPSA, Disease Activity index for Psoriatic Arthritis; PGS, Patient Global Score; PPS, Patient Pain Score; SJC, Swollen Joint Count; TJC, Tender Joint Count.

Differential item functioning

The analyses of DAPSA DIF based on baseline scores revealed highly significant evidence of DIF with respect to country (see online supplemental table S6). The strongest evidence was for the item TJC. The DIF pattern for DAPSA28 was comparable to the pattern for DAPSA regarding how items scores differed systematically across countries. There was evidence of DIF with respect to sex for PPS only and little evidence of DIF with respect to GDP.

DAPSA and DAPSA28 reliability

The internal consistency reliability as assessed by CA for DAPSA and DAPSA28 was 0.65 and 0.63 at baseline. At 6 months follow-up, CA for DAPSA and DAPSA28 was 0.66 and 0.64.

Sensitivity analysis

For both DAPSA and DAPSA28 country-level CFA at baseline for each of the five countries with 125 patients, model fit varied with some models having acceptable RMSEA model fit (online supplemental table S5). Sensitivity testing on a subsample excluding DAPSA and DAPSA28 change-score outliers showed no change in model fit acceptability.

DISCUSSION

In this study, DAPSA and DAPSA28 were investigated by polychoric rank correlations, dimensionality by scree-plots, CFA models, SEM models, CFA and SEM model fit, DIF and internal consistency reliability by CA. Overall,

Table 2 DAPSA and DAPSA28 fit of configural invariance multigroup CFA models at baseline and 6 months follow-up and SEM model fit

Model	Exact fit			Close fit	
	χ^2	df	P value	RMSEA	CFI
(a) DAPSA CFA at baseline	21.0	17	0.225	0.020	1.000
(b) DAPSA28 CFA at baseline	30.5	18	0.033	0.034	1.000
(c) DAPSA CFA at 6 months follow-up	43.6	15	0.000	0.057	1.000
(d) DAPSA28 CFA at 6 months follow-up	59.7	18	0.000	0.063	1.000
(e) DAPSA SEM	736.8	150	0.000	0.079	0.980
(f) DAPSA28 SEM	768.3	150	0.000	0.082	0.983

CFA, Confirmatory factor analysis; CFI, Comparative Fit Index; DAPSA, Disease Activity index for Psoriatic Arthritis; RMSEA, root mean square error of approximation; SEM, structural equation modelling.

the findings support acceptable internal construct validity and internal consistency reliability; however, DIF was observed for sex and country.

Several earlier studies have tested the relationships between both DAPSA or DAPSA28 and other similar and related disease activity measures, although kappa agreement of 0.32 between DAPSA and Psoriatic Arthritis Disease Activity Score for PsA (PASDAS) was the only identified result based on PsA disease activity.⁹ In previous research, correlation coefficients at baseline of 0.60–0.99 were found between DAPSA/DAPSA28 and other disease activity measures.^{6 8 24 25} Still, there is limited evidence regarding correlations between DAPSA and DAPSA28 individual items. Also, no previous studies exist regarding DAPSA and DAPSA28 internal construct validity evidence based on exploration of latent traits. A previous study on DAPSA by Nell-Duxneuner *et al*⁷ (DAPSA referred to as DAREA in the publication) evaluated dimensionality reduction by principal component analysis in 105 patients. Nell-Duxneuner *et al* found most of the variance in DAPSA to be explained by four components, indicating that it captures multiple aspects of disease activity. Nell-Duxneuner *et al* identified patient global and pain assessment on one component, swollen and tender joints on another, CRP on a third, and skin assessment (not available in DAPSA) on a fourth insignificant component. Our study only identified one or two dimensions in DAPSA and DAPSA28 depending on scree plot interpretations, whereas the study by Nell-Duxneuner *et al* found three. In this study, the one-dimensional model fit the data well, supporting the one-dimensionality of the latent construct of disease activity in PsA. Nell-Duxneuner *et al* found that DAPSA was the disease activity measure that had the best fit with these four components compared with DAS28, Simple Disease Activity Index (SDAI) and Clinical Disease Activity Index (CDAI).

No previous tests of CA were identified, but a 2-week test–retest was assessed previously.⁸ CA provides a lower bound test–retest estimate, meaning that test–retest would not be lower than these estimates. Therefore, the CA results for DAPSA of 0.65 and 0.66 at baseline and follow-up identified in this study match the test–retest of 0.92 identified in previous research.

We conducted CFA, SEM and DIF analyses using subsamples from countries with more than 125 cases, totalling 625 patients. Selecting a subsample of 125 patients at random from each of these countries was applied since the varying sample size from the countries could otherwise affect the CFA, SEM and DIF models. No significant differences in characteristics were identified between the full 1863 and the 625 randomly selected patients. To ensure meaningful comparisons across countries, these models were adjusted based on configural invariance, which checks whether the overall factor structure is consistent across groups. This adjustment allowed us to combine the model fit results across countries, providing a more robust evaluation of the models' performance at the multinational level.

Polychoric rank correlations were used to account for the ordinal nature of items due to non-normal distributions. Both DAPSA and DAPSA28 showed an overall positive correlation pattern; however, non-linearity and skewness in several items may obscure latent factor identification. As composite scores of disease activity, DAPSA and DAPSA28 include measures from diverse sources with varying distributional characteristics. Therefore, lower correlations between items like CRP and others were expected.

Neither DAPSA nor DAPSA28 was originally conceptualised as multidimensional measures,^{4 6} that is, they assumed a single underlying factor in the concept of PsA disease activity. Yet, the DAPSA dimensionality has been discussed previously. For instance, DAPSA and DAPSA28 could consist of a patient-assessed component and inflammation/physician component.²⁶ Similar discussions of dual targets (symptoms and biological) exist for rheumatoid arthritis.^{27 28} It is important to note that PsA disease dimensionality is different from the PsA disease activity dimensionality. For instance, radiographic progression and quality of life impairment are typically not seen as a part of disease activity assessment. Existing research⁷ has proposed that patient pain and global assessment could represent a separate factor, meaning that the construct PsA disease activity is two-dimensional. However, with only two items, modelling a distinct sub-factor is not statistically viable, supporting the use of one-dimensional modelling. If, for instance, DAPSA and DAPSA28 included an additional patient assessed item, ie, six items in total, this could allow for two-dimensional modelling to support this hypothesis.

The CFA models included correlations between the two patient-reported items (pain and global assessment) and between the joint counts (swollen and tender). This testing aimed to identify if high correlations between these items would suggest overlapping traits, even if they are intended to assess different aspects. Since no correlations above 0.90 were found, the results suggest these items do not measure the same traits and should not be excluded.

In both SEM models for DAPSA and DAPSA28, a negative association between the latent construct of DAPSA at 6 months follow-up between SJClen and TJC was identified (–0.28 and –0.22). Although smaller, a negative association (–0.01) was also seen in the CFA models at 6 months follow-up. This was not caused by scores being lower at follow-up due to effective treatment since the model controls for this. Instead, it means that items were affected differently over the 6 months of treatment. Country-wise models with samples of 125 patients each revealed good fit for some countries and not accepted fit for others. This result should be interpreted with caution since 125 patients are considered a small cohort for CFA analyses.

For DAPSA, DIF analysis revealed DIF for country and GDP, and for DAPSA28 the analysis revealed DIF for all tested aspects, that is, country, GDP and sex.

The findings of DIF, especially from country but also sex in DAPSA and DAPSA28, raise awareness of differences between different groups of patients and their chance of high and low DAPSA and DAPSA28 scores. These findings may impact clinical decision-making, since sex-related DIF could reflect differences in the perception of pain and assessment of tender joints, as sex has been identified as associated with both in previous studies.^{29 30} This highlights the need to further examine sex-specific DAPSA and DAPSA28 disease activity thresholds and the influence of sex on the psychometric properties of these indices. Different inclusion criteria for outpatient treatment across countries could influence the identified DIF for country. This was the first study where sex DIF in DAPSA and DAPSA28 outcomes was explored, and it was expected to identify some DIF items on DAPSA and DAPSA28 as treatment outcomes are known to differ between male and female patients with PsA³¹; yet DIF based on sex was only identified for DAPSA28. The DIF results found in four countries are in line with the differences identified in CFA and SEM country-wise models.

Strengths of this study include a large sample size of 1865 patients and 625 for advanced analysis and data from nine European countries, although for most analyses, 625 patients were included (125 patients selected at random from five countries). The choice of sampling 125 patients per country was made to balance representativeness across countries and test conditions. Including two measurement points per patient allowed for testing DAPSA and DAPSA28 over time. This provides a more solid basis for generalisation of the results than analysis based on data from one country and one time point alone. Tests of sensitivity revealed significant variation by country but no impact from the removal of outliers. There is no perfect way to model CFA and SEM to explore internal construct validity. Therefore, these different approaches were applied. Country-wise modelling suggests that DAPSA and DAPSA28 may not measure the exact same construct in each country, which means comparison of scores within a country would be more reliable than between countries.

Some psychometric aspects could not be tested based on the available data. Inter-rater and intrarater reliability assessment was not possible since physician ID was not included. An anchor-based minimal clinically important difference (MCID) could not be tested because the data did not include patient assessment of improvement. This is an important limitation of this study since MCID for improvement as well as worsening based on DAPSA and DAPSA28 has not been tested previously. Also, the data did not include other PsA specific disease activity measures, so we could not investigate relations between DAPSA and DAPSA28 vs indices such as the multidimensional PASDAS or the Composite Psoriatic Disease Activity Index (CPDAI).

In line with Nunnally et al. 1967²³, a CA of 0.6 was applied as an early-stage threshold. However, the absence of prior testing and the lack of a universally accepted standard³² can be considered a limitation of the study.

Improved validity and internal consistency reliability with different or moderated indices could enhance treatment decisions, allowing for greater precision in patient care. This, in turn, could benefit clinical trial designs by potentially reducing the required sample size, thus saving economic resources. However, it is important to recognise that some of the tested aspects have not been previously explored, and further testing in other cohorts would be valuable.

In conclusion, the overall model fit was acceptable across model fit statistics, but some evidence of misfit at country level was disclosed. In addition, our findings support acceptable internal consistency reliability, while DIF was found for sex and country. Thus, our study supports the use of DAPSA and DAPSA28 for assessment of PsA disease activity. Still, it would be relevant to explore internal construct validity of other existing or future PsA disease activity measures, hoping to identify an outcome measure with higher internal construct validity.

Author affiliations

¹Copenhagen Center for Arthritis Research, Center for Rheumatology and Spine Diseases, Rigshospitalet, Glostrup, Denmark

²Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark

³Department of Public Health, University of Copenhagen, Copenhagen, Denmark

⁴Center for treatment of Rheumatic and Musculoskeletal Diseases (REMEDY), Diakonhjemmet Hospital, Oslo, Norway

⁵Research Unit, Sørlandet Sykehus HF, Kristiansand, Norway

⁶Department of Rheumatology and DANBIO Aarhus University Hospital, Aarhus, Denmark

⁷Department of Clinical Medicine, Aarhus University, Aarhus, Denmark

⁸Copenhagen Research Center for Biological and Precision Psychiatry, Mental Health Centre Copenhagen, Copenhagen University Hospital, Copenhagen, Denmark

⁹Department of Clinical Medicine, University of Copenhagen, Copenhagen, Denmark

¹⁰Institute of Rheumatology, Prague, Czech Republic

¹¹Department of Rheumatology, First Faculty of Medicine, Charles University, Prague, Czech Republic

¹²Rheumatology Unit, DiMePREJ, University of Bari, Bari, Italy

¹³Department of Clinical Internal, Anesthesiological and Cardiovascular Sciences, Sapienza University of Rome, Rome, Italy

¹⁴Instituto Português de Reumatologia, Lisbon, Portugal

¹⁵Center for Rheumatic Diseases, University of Medicine and Pharmacy, Bucharest, Romania

¹⁶DANBIO Rigshospitalet, Glostrup, Denmark

¹⁷Department of Rheumatology, University Hospital Zurich, University of Zurich, Zürich, Switzerland

¹⁸Department of Rheumatology, Geneva University Hospitals, Geneva, Switzerland

¹⁹Departments of Medicine and Rheumatology, Helsinki University Central Hospital, Helsinki, Finland

²⁰Department of Rheumatology, Division of Medicine, Turku University Central Hospital, Turku, Finland

²¹Department of Internal Medicine, University of Turku, Turku, Finland

²²Department of Rheumatology, East Tallinn Central Hospital, Tallinn, Estonia

²³Centre for Rheumatology Research, Landspítali University Hospital, Reykjavík, Iceland

²⁴Faculty of Medicine, University of Iceland, Reykjavík, Iceland

²⁵University of Iceland, Faculty of Medicine, Reykjavik, Iceland

²⁶Landspítali University Hospital, Reykjavik, Iceland

Acknowledgements The EuroSpA collaboration has been supported by Novartis Pharma AG since 2017 and UCB Biopharma SRL since 2022.

Contributors JBJ (guarantor), KBC, LMØ and MØ contributed to the study design. JBJ contributed to the analysis. SHR contributed with data curation. JBJ drafted the article. All authors contributed with critical revision of the manuscript, approved the final version of the manuscript and took responsibility for the manuscript content.

Funding No financial sponsors had any influence on the data collection, statistical analyses, abstract/manuscript preparation, or decision to submit.

Disclaimer No financial sponsors had any influence on the data collection, statistical analyses, abstract/manuscript preparation or decision to submit.

Competing interests JBJ: Novartis, KBC: None declared, BM: Novartis, Pfizer, AGL: Novartis, Janssen, UCB, SHR: Novartis, JH: Novartis, UCB, MLH: AbbVie, Biogen, BMS, Celltrion, DANBIO, DRQ, Eli Lilly, Janssen Biologics B.V, Lundbeck Fonden, Medac, MSD, Novartis, Nordforsk, Pfizer, Roche, Sandoz, Samsung Biopies, JZ: AbbVie, AstraZeneca, Egis, Eli-Lilly, Glaxo, Novartis, Sandoz, Sanofi, Sobi, Swix-biopharma, UCB, KP: AbbVie, Amgen, Celltrion, Eli Lilly, Fresenius, Pfizer, MSD, Novartis, Pfizer, UCB, Viatrix, FI: Abbvie, AlfaSigma, Amgen, AstraZeneca, BMS, Galapagos, Janssen, Lilly, MSD, Novartis, Pfizer, UCB, FC: Abbvie, AlfaSigma, AstraZeneca, GSK, Lilly, Novartis, Pfizer, UCB, JB: None declared, CC: AbbVie, Amgen, AstraZeneca, Boehringer Ingelheim, Ewopharma, Lilly, Novartis, Pfizer, Sandoz, Sobi, CM: Pfizer, UCB, BG: Abbvie, BMS, Pfizer, Sandoz, AC: None declared, MJN: AbbVie, Amgen, Eli Lilly, Janssen, Novartis, Pfizer, DN: BMS, Lilly, MSD, Novartis, Pfizer, UCB, LK: AbbVie, Amgen, Boehringer Ingelheim, Eli Lilly, Fresenius Kabi, Galapagos, Johnson & Johnson, Lilly, Medac, Pfizer, UCB, KL: AbbVie, Johnson and Johnson, Novartis, Pfizer, SV: None declared, BG: None declared, TJL: None declared, MØ: AbbVie, Amgen, BMS, Boehringer Ingelheim, Celgene, Eli Lilly, Galapagos, Gilead, Hospira, Janssen, MEDAC, Merck, Novartis, Novo, Orion, Pfizer, Regeneron, Roche, Sandoz, Sanofi, UCB, LMØ: Novartis.

Patient consent for publication Not applicable.

Ethics approval Each participating registry had obtained the necessary approval from the relevant national data protection agency and ethical committee before transferring their data for this study.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available on reasonable request. The data in this article was collected in the individual registries and made available for secondary use through the EuroSpA Research Collaboration Network [https://eurospa.eu/#registries] Relevant patient level data may be made available on reasonable request to the corresponding author but will require approval from all contributing registries.

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: <https://creativecommons.org/licenses/by-nc/4.0/>.

ORCID iDs

Jacob Brauner Jørgensen <https://orcid.org/0000-0002-6533-6336>

Merete Lund Hetland <https://orcid.org/0000-0003-4229-6818>

Florenzo Iannone <https://orcid.org/0000-0003-0474-5344>

Bente Glinthorg <https://orcid.org/0000-0002-8931-8482>

Adrian Ciurea <https://orcid.org/0000-0002-7870-7132>

Michael J Nissen <https://orcid.org/0000-0002-6326-1764>

Bjorn Gudbjornsson <https://orcid.org/0000-0003-4631-6505>

Mikkel Østergaard <https://orcid.org/0000-0003-3690-467X>

Lykke Midtbøll Ørnberg <https://orcid.org/0000-0002-7832-6831>

REFERENCES

- Eder L, Haddad A, Rosen CF, *et al*. The Incidence and Risk Factors for Psoriatic Arthritis in Patients With Psoriasis: A Prospective Cohort Study. *Arthritis Rheumatol* 2016;68:915–23.
- Haddad A, Zisman D. Comorbidities in Patients with Psoriatic Arthritis. *Rambam Maimonides Med J* 2017;8:e0004.
- Sandoval DC, Fernández-Avila DG. Assessment tools in psoriatic arthritis: A review. *Rev Colomb Reumatol* 2023;30:S75–86.
- Schoels M, Aletaha D, Funovits J, *et al*. Application of the DAREA/ DAPSA score for assessment of disease activity in psoriatic arthritis. *Ann Rheum Dis* 2010;69:1441–7.
- Eberl G, Studnicka-Benke A, Hitzelhammer H, *et al*. Development of a disease activity index for the assessment of reactive arthritis (DAREA). *Rheumatology (Oxford)* 2000;39:148–55.
- Michelsen B, Sexton J, Smolen JS, *et al*. Can disease activity in patients with psoriatic arthritis be adequately assessed by a modified Disease Activity Index for Psoriatic Arthritis (DAPSA) based on 28 joints? *Ann Rheum Dis* 2018;77:1736–41.
- Nell-Duxneuner VP, Stamm TA, Machold KP, *et al*. Evaluation of the appropriateness of composite disease activity measures for assessment of psoriatic arthritis. *Ann Rheum Dis* 2010;69:546–9.
- Tillett W, Helliwell P, Fitzgerald O, *et al*. AB0839 Reliability of composite measures for the assessment of psoriatic arthritis. *Ann Rheum Dis* 2020;79:1725.
- van Mens LJJ, van de Sande MGH, van Kuijk AWR, *et al*. Ideal target for psoriatic arthritis? Comparison of remission and low disease activity states in a real-life cohort. *Ann Rheum Dis* 2018;77:251–7.
- Singh JA, Guyatt G, Ogdie A, *et al*. Special Article: 2018 American College of Rheumatology/National Psoriasis Foundation Guideline for the Treatment of Psoriatic Arthritis. *Arthritis Care Res (Hoboken)* 2019;71:2–29.
- Gossec L, Smolen JS, Ramiro S, *et al*. European League Against Rheumatism (EULAR) recommendations for the management of psoriatic arthritis with pharmacological therapies: 2015 update. *Ann Rheum Dis* 2016;75:499–510.
- Sánchez BN, Budtz-Jørgensen E, Ryan LM, *et al*. Structural equation models: a review with applications to environmental epidemiology. *J Am Stat Assoc* 2005;100:1443–55.
- Kelava A. A Review of Confirmatory Factor Analysis for Applied Research. *J Educ Behav Stat* 2016;41:443–7.
- Linde L, Ørnberg LM, Rasmussen SH, *et al*. Commonalities and differences in set-up and data collection across European spondyloarthritis registries - results from the EuroSpA collaboration. *Arthritis Res Ther* 2023;25:205.
- Eurostat. Purchasing power adjusted GDP per capita. 2024. Available: https://ec.europa.eu/eurostat/databrowser/view/sdg_10_10/default/table?lang=en [Accessed 5 Dec 2024].
- Beaton D, Maxwell L, Grosskleg S, *et al*. OMERACT outcome measures in rheumatology. 2021. Available: https://omeract.org/wp-content/uploads/2021/12/OMERACT-Handbook-Chapter-5_Final_June-2-2021_a.pdf [Accessed 4 Sep 2023].
- American Educational Research Association. Standards for educational and psychological testing. Washington, DC American Educational Research Association; 2014.
- Akoglu H. User's guide to correlation coefficients. *Turk J Emerg Med* 2018;18:91–3.
- Hu LT, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Struct Equ Modeling* 1999;6:1–55.
- Rigdon EE. CFI versus RMSEA: A comparison of two fit indexes for structural equation modeling. *Struct Equ Modeling* 1996;3:369–79.
- Cheung GW, Rensvold RB. Evaluating Goodness-of-Fit Indexes for Testing Measurement Invariance. *Struct Equ Modeling* 2002;9:233–55.
- Tavakol M, Dennick R. Making sense of Cronbach's alpha. *Int J Med Educ* 2011;2:53–5.
- Nunnally J. C. Psychometric Theory. New York, NY: McGraw-Hill, 1967.
- Proft F, Schally J, Brandt HC, *et al*. Evaluation of the Disease Activity Index for Psoriatic Arthritis (DAPSA) with a quick quantitative C reactive protein assay (Q-DAPSA) in patients with psoriatic arthritis: a prospective multicentre cross-sectional study. *RMD Open* 2022;8:e002626.
- Theander E, Husmark T, Alenius G-M, *et al*. Early psoriatic arthritis: short symptom duration, male gender and preserved physical functioning at presentation predict favourable outcome at 5-year follow-up. Results from the Swedish Early Psoriatic Arthritis Register (SwePsA). *Ann Rheum Dis* 2014;73:407–13.
- Gorlier C, Orbai A-M, Puyraimond-Zemmour D, *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;78:201–8.
- 27 Ferreira RJO, Ndosi M, de Wit M, *et al*. Dual target strategy: a proposal to mitigate the risk of overtreatment and enhance patient satisfaction in rheumatoid arthritis. *Ann Rheum Dis* 2019;78:e109–10.
- 28 Landewé RBM. Response to: “Dual target strategy: a proposal to mitigate the risk of overtreatment and enhance patient satisfaction in rheumatoid arthritis” by Ferreira *et al*. *Ann Rheum Dis* 2019;78:e110:1.
- 29 Coates LC, van der Horst-Bruinsma IE, Lubrano E, *et al*. Sex-Specific Differences in Patients With Psoriatic Arthritis: A Systematic Review. *J Rheumatol* 2023;50:488–96.
- 30 Barnabe C, Bessette L, Flanagan C, *et al*. Sex differences in pain scores and localization in inflammatory arthritis: a systematic review and metaanalysis. *J Rheumatol* 2012;39:1221–30.
- 31 Lubrano E, Scriffignano S, Fatica M, *et al*. Psoriatic Arthritis in Males and Females: Differences and Similarities. *Rheumatol Ther* 2023;10:589–99.
- 32 Peterson RA. A meta-analysis of cronbach’s coefficient alpha. 1994. Available: <https://www.jstor.org/stable/2489828>
- 33 Brauner J, Christensen KB, Michelsen B, *et al*. OP0181 DAPSA and DAPSA28 internal construct validity assessed by confirmatory factor analysis and other psychometric tests based on European data. *Ann Rheum Dis* 2025;84:150–2.