Comparing patient-perceived and physician-perceived remission and low disease activity in psoriatic arthritis: an analysis of 410 patients from 14 countries

Clémence Gorlier, Ana-Maria Orbai, Déborah Puyraimond-Zemmour, Laura C Coates, Uta Kiltz, Ying-Ying Leung, Penelope Palominos, Juan D Cañete, Rossana Scrivo, Andra Balanescu, Emmanuelle Dernis, Sandra Talli, Adeline Ruyssen-Witrand, Martin Soubrier, Sibel Zehra Aydin, Lihi Eder, Inna Gaydukova, Ennio Lubrano, Umut Kalyoncu, Pascal Richette, M. Elaine Husni, Maarten de Wit, Josef S. Smolen, Laure Gossec

Affliations

Clémence Gorlier, Déborah Puyraimond-Zemmour, Laure Gossec: Sorbonne Université, Paris France; Pitié Salpêtrière hospital, APHP, Rheumatology department, Paris, France.

Ana-Maria Orbai, MD MHS, Johns Hopkins University School of Medicine, Division of Rheumatology, Baltimore, MD, USA

Laura C Coates, Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, University of Oxford, Oxford, UK

Uta Kiltz, Rheumazentrum Ruhrgebiet, Herne and Ruhr-Universität Bochum, , Germany

Ying-Ying Leung, Singapore General Hospital, Singapore

Penelope Palominos, Hospital de Clinicas de Porto Alegre, Porto Alegre, Brazil

Juan D Cañete, Hospital Clínic and IDIBAPS, Barcelona, Spain

Rossana Scrivo, Rheumatology Unit, Department of Internal Medicine and Medical Specialties, Sapienza Università di Roma, Rome, Italy

Andra Balanescu, Sf Maria Hospital, University of Medicine and Pharmacy Carol Davila, Bucharest, Romania

Emmanuelle Dernis, Le Mans Central Hospital, Le Mans, France

Sandra Tälli, Tallinn Central Hospital, Tallinn, Estonia

Adeline Ruyssen-Witrand, Rheumatology Unit, Toulouse university Hospital, UMR 1027, Inserm, Université Paul Sabatier Toulouse III, Toulouse, France

Martin Soubrier, Gabriel Montpied Hospital, Clermont Ferrand, France

Sibel Zehra Aydin, University of Ottawa, the Ottawa Hospital Research Institute, Ottawa, Canada

Lihi Eder, Women's College Hospital, University of Toronto, Toronto, ON, Canada

Inna Gaydukova, North-western State medical university, St.Petersburg, Russia, Russia

Ennio Lubrano, Academic Rheumatology Unit, Dipartimento di Medicina e Scienze della Salute "Vincenzo Tiberio", University of Molise,Campobasso, Italy.

Umut Kalyoncu, Hacettepe University Faculty of Medicine, Department of Internal Medicine, Division of Rheumatology, Ankara, Turkey

Pascal Richette, Hopital Lariboisiere Centre Viggo Petersen, service de rhumatologie, Paris, France ; Universite Paris Diderot UFR de Medecine, Inserm UMR1132 Bioscar, , Paris France

M. Elaine Husni, Cleveland Clinic, Department of rheumatic and Immunologic Diseases, Cleveland, USA

Maarten de Wit, Amsterdam University Medical Centre, Dept. Medical Humanities, Amsterdam Public Health (APH), Amsterdam, Netherlands

Josef Smolen, Division of Rheumatology, Department of Medicine 3, Medical University of Vienna, Vienna, Austria

Key Indexing Terms: remission, disease activity, composite score, psoriatic arthritis

Corresponding author:

Pr Laure GOSSEC, Hôpital Pitié- Salpêtrière, Service de Rhumatologie, 47-83 bd de l'hôpital, 75013 PARIS FRANCE

Email : laure.gossec@aphp.fr Tel=+33 142178421

Word count: 3602 words, 49 references, 4 tables, 2 figures, 3 online figures and 1 online table

Disclosures: None relevant to this paper

Funding: financial support from Pfizer through an unrestricted research grant. The fellow (CG) was additionally supported by a Master's bursary from Societe Francaise de Rhumatologie.

Laura Coates is funded by a National Institute for Health Research Clinician Scientist award. Her 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.

Ana-Maria Orbai 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 (Core B), a Rheumatology Research Foundation Scientist Development award, and a Staurulakis Family Discovery award. All statements in this report including its conclusions are the opinions of the authors and do not necessarily reflect those of NIH or NIAMS of the Foundation.

Key messages:

- Investigating an unselected, standard of care population of 410 psoriatic arthritis (PsA) patients, both remission and low disease activity were frequently attained: from 12.4% to 36.1% for remission and from 25.4% to 46.8% for LDA.
- Patient-perceived remission/low disease activity was frequent (65.4%) indicating patients often reported themselves in a low level of disease activity.
- Patient-perceived remission was as frequent as remission based on composite scores (VLDA/MDA or DAPSA); both were less frequent than physician-reported remission using a single question.
- Agreement between patient perceived remission/low disease activity and composite scores was fair to moderate. VLDA/MDA showed a lower sensitivity than DAPSA versus the patient perspective (52% vs 73%) but had a higher specificity (88 vs 77%). DAPSA based status had both sensitivity and specificity around 75% indicating that this score appears to better reflect patient perceived low disease activity.

Abstract (245 words):

Background

The objective was to compare different definitions of remission and low disease activity (LDA) in patients with psoriatic arthritis (PsA), based on both patients' and physicians' perspectives.

Methods

In ReFlap (NCT03119805), adults with physician-confirmed PsA and >2 years disease duration in 14 countries were included. Remission was defined as Very Low Disease Activity (VLDA), Disease Activity in PSoriatic Arthritis (DAPSA) ≤4, physician-and patient-perceived remission (specific question yes/no), and LDA as Minimal Disease Activity (MDA), DAPSA≤14, physician- and patient-perceived LDA. Frequencies of these definitions, their agreement (prevalence adjusted kappa), and sensitivity and specificity versus patient-defined status were assessed cross-sectionally.

Results

Of 410 patients, mean age (standard deviation) 53.9 (12.5) years, 50.7% male, disease duration 11.2 (8.2) years, 56.8% on biologics, remission/LDA was frequently attained: respectively for remission, from 12.4% (VLDA) to 36.1% (physician-perceived remission) and for LDA, from 25.4% (MDA) to 43.9% (patient-perceived LDA). Thus, patient-perceived remission/LDA was frequent (65.4%). Agreement between patient-perceived remission/LDA and composite scores was moderate to good (kappa range, 0.12-0.65). When using as reference, patient-perceived remission or LDA status, DAPSA-defined remission/LDA and VLDA/MDA had a sensitivity of 73.1% and 51.5% respectively, and a specificity of 76.8% and 88.0%, respectively. Physician-perceived remission/LDA using a single question was frequent (67.6%) but performed poorly against other definitions.

Conclusion

In this unselected population, remission/LDA was frequently attained. VLDA/MDA was a more stringent definition than DAPSA-based REM/LDA. DAPSA-based remission/LDA performed better than VLDA/MDA to detect patient-defined remission or remission/LDA. Further studies of long-term outcomes are needed.

Introduction

Psoriatic arthritis (PsA) is a complex inflammatory disease that spans a wide spectrum to include peripheral joints, skin, entheses, spine, and other adjacent tissues.

Recent management recommendations state that remission (REM), or, in some cases, low disease activity (LDA) is the treatment goal in PsA.(1–4) Several composite disease activity measures have been developed, and the current discussed treatment target definitions for REM/LDA are: VLDA (Very Low Disease Activity)/MDA (Minimal Disease Activity) (5-7) and DAPSA (Disease Activity index for Psoriatic Arthritis) cut-offs of $\leq$4/$\leq$14 (or clinical DAPSA, cDAPSA). (8-10) These definitions each have strengths and weaknesses which hamper achieving consensus on one definition.(11,12) To briefly summarise some of the issues: on the one hand, VLDA/MDA include a measure of function (Health Assessment Questionnaire, HAQ) which can be influenced by factors other than disease activity- this may be a methodological issue. On the other hand, DAPSA only assesses joints and not directly any other domain of PsA, such as entheses or skin, MDA does not assess dactylitis, and both do not assess all patient-important domains. While the Outcome Measures in Rheumatology Core Set states that all domains mentioned are of importance (13), the various multi-dimensional composite measures have major differences in their components and none uses all components. The question of unidimensional versus monodimensional scores has been widely addressed however there is currently no consensus in this respect. To this end we have used a unidimensional (DAPSA) and a multidimensional (MDA) instrument. Three recent studies have compared the VLDA/MDA outcomes to DAPSA outcomes in terms of frequency but did not assess the patient's perspective in parallel.(14-16)

REM/LDA from the patient's perspective has not been defined. The above composite measures factor in patient-reported outcomes including pain and patient global assessment.(5-10) However, they were developed with little patient involvement and cut-offs for REM/LDA were not patient-driven.(17, 18 This may be important since disagreements in the assessment of disease activity have a potential impact on treatment decisions and shared decision-making.(19-21) The only data available regarding the patient's assessment of REM/LDA are issued from studies on aspects of disease

impact.(22,23) However patient-perceived LDA or REM can be approached by specific designated questions, by the 'patient acceptable symptom state', or using low values of patient global assessment (PGA).(24-26) REM/LDA can also be defined, from the physician's perspective, as achieving a REM/LDA based on a global assessment of the physician (yes/no). Such single questions may have clinical relevance though they have not yet been assessed formally.

Since alignment between patients and health professionals in terms of treatment targets is thought to be a key component for shared decision-making,(27,28) it is of great interest to compare physician-perceived REM/LDA and composite scores with patient-perceived REM/LDA in the assessment of PsA.

The objectives of the present study were to assess the frequency of REM/LDA using different definitions according to the patient's and physician's perspective, and to assess agreement between these definitions.

Methods

Study population and study design

The ReFlaP (Remission/Flare in PsA) study was a prospective, multicenter international, longitudinal, observational study which took place in 21 centers in 14 countries (including 7 countries across Europe, the United Kingdom, Russia, Canada, the United States of America, Brazil, Turkey and Singapore) between June 2017 and August 2018 (NCT03119805). The objective of the study was to assess REM/LDA in PsA. Patients were seen twice; here, baseline data were used.

Adult patients with a diagnosis of PsA as defined by their rheumatologist and more than 2 years of disease duration were recruited. Investigators were advised to consider the CASPAR criteria for classification of PsA. Patients with no definite PsA or less than 2

years of disease duration, patients who didn't speak or read the local language or were not comfortable filling in a paper form in the local language were excluded. All patients granted informed consent, with ethical committee approval at each site. The inclusion of patients was performed consecutively.

Data collection

Medical data

The collected data included patient demographic variables (age, gender, work status, level of education) and the following disease characteristics: disease duration, predominant type of PsA (peripheral, axial or entheseal), current treatment (conventional disease-modifying drugs (csDMARDs) and/or biologic disease-modifying drugs (bDMARDs)). The Functional Comorbidity Index and the last available result (< 4 weeks) for C-Reactive Protein (CRP) were collected.(29) Physical examination included assessment with 66 swollen joint count (SJC), 68 tender joint count (TJC), tender entheseal points (by the Leeds Enthesitis Index), body surface area of psoriasis and physician global assessment (on a scale of 0-10).(30)

Patient-reported outcomes

PGA with a wording focused on disease activity was collected on a 0–10 numeric rating scale; as follows*: 'How active was your rheumatic disease on average during the last week?'* (from 'Not active' to 'Very active') and was used to calculate the composite scores.(31) This wording refers to the concept of disease activity and has been used in other rheumatic diseases.(31) As sensitivity analyses, this wording was replaced in the composite scores by wordings referring to global joint and global skin assessments.(32) Also collected were the HAQ Disability Index; and Patient Acceptable Symptom State, PASS (in the absence of a standardised PASS question, the following wording was used: "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?" yes/no).(33, 34) The PsA Impact of Disease (PsAID) assesses the impact of PsA on 12 aspects with a final result between 0 and 10 (higher results indicate a worse condition).(35)

The patient data collection form was translated by 2 persons into each local language according to usual procedures.

REM and LDA definitions

Composite scores

VLDA/MDA, DAPSA and cDAPSA were used to define REM and LDA (Table 1).

Physician perspective

Physicians were asked 2 separate single questions for REM/LDA, formulated by the steering committee as “At this time, is the psoriatic arthritis in remission, if this means: “the absence of clinical and laboratory evidence of significant inflammatory disease activity?”, and “At this time, is the psoriatic arthritis in low or minimal disease activity?”.

Of note the physicians answered these questions unblinded to other results (e.g. they could consult the patient questionnaires and CRP results if they wished as in their routine clinical practice). No instructions were given as to which aspects of disease should be considered when answering these questions, but the rheumatologists including patients into this study were all experienced in treating PsA and the question was related to PsA rather than to skin involvement which was addressed in a separate question.

Patient’s perspective

REM/LDA separate questions for patients were developed with input from 4 patient research partners with PsA and was based on previous work in the field of rheumatoid arthritis. (36, 37) The phrasing was the following: “At this time, is your psoriatic arthritis in remission, if this means: you feel your disease is as good as gone?’ (for 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 LDA).

From the patients’ perspective, two potential definitions for REM were used: patient-perceived remission (single question as above) and PGA$\leq$1. Also, two definitions for LDA were used: patient-perceived LDA (single question) and PGA$\leq$3. The PGA cut-offs were informed, for REM, by the rheumatoid arthritis international REM criteria since no cut-off

has been defined in PsA.(38) For LDA, the cutoff of PGA≤3 was selected by the steering committee. Such cutoffs are arbitrary and given issues around circularity between PGA and the composite scores, the PGA external criterion should be considered as indicative only.

As a comparison outcome, the Patient Acceptable Symptom State was compared to a state of LDA.

Statistical analysis

All patients with items available to calculate REM/LDA with all definitions were analyzed. Demographic, clinical and biologic variables were expressed as means +/- standard deviations (SDs) for continuous variables and as frequencies (percentages) for categorical variables. No imputation of missing data was performed; data were analysed on complete cases. To obtain an overview of the meaning of patient-defined disease states, patient characteristics in each self-defined disease state were described. Proportions achieving each REM/LDA criterion were calculated and for the composite score definitions REM and LDA groups were analysed separately, and then also combined. . Venn diagrams were used to represent the number of patients meeting different REM/LDA criteria. To assess performances of the composite scores, their sensitivity and specificity was calculated versus the reference definition which was here patient-perceived status (i.e., REM or REM/LDA). Thus, sensitivity was the % of patients in self-reported good status who was found in good status using the composite score, and specificity was the % of patients in self-reported lack of good status, who were found in lack of good status using the score.

The agreement between the tested definitions was established using 2×2 tables and calculation of Cohen's kappa and prevalence-adjusted bias-adjusted kappa (PABAK) where necessary, using Bennett's method.(39,40) In cases of discrepancy between Cohen's kappa and PABAK, the paradox of the kappa may apply and PABAK should be analysed preferentially. Usual cutoffs to interpret kappas were used, namely, 0.00-0.20 slight agreement, 0.21-0.40 fair, 0.41-0.60 moderate and 0.61-0.80 good agreement. R software, version 3.4.3, was used for all statistical analyses.

Results

Demographic and clinical characteristics

A total of 466 patients were included: 56 were ineligible (no confirmation of diagnosis, N=11, age below 18, N=1) or had missing data (mainly CRP, N=27 entheseal assessment, N=6 or HAQ, N=2; other criteria were missing in 9 patients). Thus, 410 with complete data were analyzed (Table 2). Of these, 50.7% were male and the mean disease duration 11.2 years. Disease activity was moderate and the majority were receiving conventional synthetic DMARDs (59.3%) and/or biologic DMARDs (56.8%).Disease activity was lower in patients in self-defined REM or LDA, supporting validity of the questions applied in the present study (Table 2).

Prevalence of REM/LDA according to the different definitions (Figure 1)

REM: the most frequent REM status was obtained using physician single question: 148 (36.1%) patients. cDAPSA (25.6% REM) and both of the patient-defined REM (single question, 21.5% or PGA<=1, 24.4%) were of similar frequency. DAPSA (19.0% REM) and especially VLDA (12.4%) were more stringent.

LDA: this status was frequent, in particular when using the patient single question (43.9%, figure 1). The definition leading least frequently to this status was MDA (25.4%)

*REM+LDA:*VLDA/MDA was difficult to reach with only 37.8% in REM/LDA; DAPSA was less limiting with 58.5% of patients. Patient-perceived REM/LDA and physician-perceived REM/LDA were also less limiting than VLDA/MDA and had similar frequencies (65.4% and 67.6% respectively).

Of note, 269 (65.6%) patients were in PASS.

Agreement between REM/LDA definitions

Agreements between definitions are shown in Table 3.

REM: there was a very high agreement between DAPSA and cDAPSA REM reflecting the similarity of the two definitions.(12,13) The agreement between DAPSA/cDAPSA and VLDA and between PGA≤1 and VLDA, cDAPSA and DAPSA was high, however the latter may reflect some circularity since PGA is a component of these measures. (4-10) The agreement between VLDA/cDAPSA/DAPSA and patient-perceived REM was moderate to good and comparable, (Table 3).

LDA: excluding expected high agreement between DAPSA and cDAPSA LDA, agreements were lower for LDA than for REM (Table 3).

Agreement between PASS and composite scores was moderate (kappa 0.56 and 0.59 and PABAK 0.33 and 0.58 for [VLDA or MDA] and [DAPSA REM or LDA], respectively, data not shown).

Sensitivity/specificity of different REM/LDA definitions versus the patient's assessment of status

Performances of different definitions are shown in Table 4 with detailed Venn diagrams in online supplementary Figures S1, S2 and S3.

REM:When using as reference, patient-perceived REM, sensitivity of DAPSA-defined REM and VLDA were respectively 47.7% and 38.6%, and specificities were respectively 88.8% and 94.7% (Table 4). Physician-perceived REM was less stringent thus leading to higher sensitivity but with lower specificity (Table 4).

LDA: there were 180 patients in patient-perceived LDA. Of these, 62 (sensitivity, 34.4%) met MDA criteria, 101 (56.1%) were in DAPSA-LDA, and 60 (33.3) were not in LDA according to any composite score (Table 4).

When analyzing as outcome, either patient-perceived REM or LDA (i.e., the sum of patients in these outcomes), sensitivity of DAPSA-defined REM/LDA and VLDA/MDA

versus patient-perceived status was respectively 73.1% and 51.5% (Figure 2). Conversely, specificity was respectively for DAPSA-defined REM/LDA and VLDA/MDA, 76.8% and 88.0%.

Main results were calculated when replacing in the composite scores, the PGA phrasing by phrasings referring to global assessment of joints and of skin psoriasis.(32). Results were very similar (online supplementary Table 1).

Discussion

This unique cohort of unselected patients with PsA brings important information on REM/LDA concepts and adds a dimension related to the patient's perspective. Defining a specific target for REM/LDA is of importance because a treat-to-target approach with either REM or LDA as the target is now recommended in standard care by guidelines for patients with PsA.(1,8) We were able to explore patient and physician-perceived REM/LDA using novel questions. We found that patient-perceived REM/LDA was frequent (65.4%) thus patient-perceived REM/LDA was similar in terms of prevalence to physician-perceived REM/LDA (67.6%) and to DAPSA-based REM/LDA (58.5%) compared with a lower frequency of MDA/VLDA (37.8%). When comparing patient-perceived status and composite scores, we found neither DAPSA-REM nor VLDA could detect all patients in self-reported REM though DAPSA performed better (sensitivity 47.7% and 38.6% respectively). When analyzing the status of REM/LDA pooled, agreement with composite scores was moderate to good; sensitivity was low for VLDA/MDA (51.5%) and higher for DAPSA-based cutoffs (73.1%) whereas specificity was high for both scores, though higher using VLDA/MDA (88.0% and 76.8%, respectively). Physician-perceived status appeared

too lenient when using a single question, with low agreements with other definitions of REM. Finally, agreements between definitions were moderate for LDA (when analysed alone), indicating the concept of LDA may need further exploration.

This study had strengths and weaknesses. Recruitment occurred in tertiary care centers as reflected by a high percentage of patients under biologics, which may limit external validity. Nevertheless, it is generalizable due to the international large-scale recruitment strategy of consecutive patients with PsA. Furthermore, frequencies of REM/LDA were similar to other studies which supports the validity of the present findings.(14-16) Another difficulty was to choose among many possible definitions of REM/LDA since no consensus exists. The instruments investigated in this study, DAPSA and MDA, are the ones recommended by an international Task Force to be applied when measuring disease activity in PsA.(3) This study brings new information on these instruments. Other possible definitions of REM/LDA provided by other measures (41,42) were not assessed, since they did not obtain a majority vote in the Treat To Target recommendations which were developed by a large international Task Force.(3) However, further research may explore such other instruments.

The scores were calculated using a wording for PGA, referring to disease activity and referring more to joints than skin; however, results were overall similar when performing the analyses with patient global questions referring to either joints, or skin. It is noteworthy that missing data was low (<15%) even though no queries were sent to the investigators, which supports the feasibility of these scores in clinical practice. A potential weakness is the use of non-validated single questions to explore patient- and physician-perceived REM/LDA. It was not possible to use consensual questions since none exist. Thus, questions were developed for the purpose of this study. Of note, great attention was paid to their elaboration process by involving patient research partners to ensure face and content validity, while physician-perceived REM/LDA questions were developed by the steering committee. Thus, these questions were developed with relevant input and support the REM/LDA concepts. However, they reflect more PsA concepts than skin psoriasis concepts – this ought to be taken into account when interpreting the study. . It should also be recognized that the present population had limited skin involvement, as is

often the case in PsA patients seen in rheumatology clinics (43). Results may differ in patients with more severe skin disease, e.g. PsA patients seen predominantly in dermatology offices, or in patients with less well-controlled disease.

This study focused on patient-perceived REM/LDA. Patients defined themselves as in REM/LDA in around 65% of cases (Figure 1). This is encouraging in terms of the overall disease burden of PsA.(44) and should be interpreted in the context that many of our study patients were receiving biologics. These results are in line with recent efforts to identify patients' priorities.(13,45) Interestingly, similar frequencies of low activity were found using REM/LDA questions and the Patient Acceptable Symptom State single question; this does not mean we suggest a Patient Acceptable Symptom State should be used as treatment target though; this criterion was used as grounding element only. Patient-perceived status refers to the disease process but also to patient expectations.(23) Considering recruitment occurred in 14 countries for the present study, it is interesting to note that patient status was self-reported as satisfactory so often, since recent data have indicated high patient expectations in countries with higher gross domestic product.(46) Such notions should be further explored.

When considering REM as the treatment target, we found composite scores to be only moderately in agreement with the patient perspective. In particular, 48.8% of patients in self-reported REM were not in VLDA or DAPSA-based REM, and 33.3% of those in self-reported LDA were not considered in LDA by composite measures. These figures lead to low sensitivities of composite scores to detect patient-defined REM, though DAPSA performed better than VLDA in this respect. Concordance was higher when pooling REM and LDA concepts. This may indicate limits of the composite scores to perfectly distinguish REM from LDA, and/or difficulties for patients to distinguish these states. LDA may be a personal concept and is more likely to carry different meanings for different people depending on their disease phenotype. Another explanation is that patients’ and physicians’ opinions on REM/LDA may differ and that composite measures may not entirely consider patient’s priorities.(13,47) Patients probably do not only refer to disease activity when considering the concept of REM; thus some discordance is expected. It

would be interesting to further investigate the connection between achieving different disease activity states and long-term prognosis.

In the present study, physician-perceived REM/LDA was explored using designated specific questions. We found that physicians defined patients as in REM much more often than composite scores or patients themselves. This indicates physicians' expectations for REM may be low, as has been previously suggested.(19-21,23,47)

Cross-tabulation of patient-perceived and physician-perceived REM/LDA is a novel contribution of our work. Agreement between patient and physician-perceived REM was not high and as stated, physicians were more lenient to define REM. However, the tendency was reversed for LDA: frequency of patient-perceived LDA was 43.9% versus 31.5% for physician-perceived LDA. Perhaps the concept of LDA needs to be further defined with both patients and physicians. Considerably higher agreement and concordance of patient-perceived REM/LDA with composite REM/LDA definitions versus physician perceived REM/LDA confirms that physicians should not base medical decisions or their global assessment/gestalt (as this may underestimate disease activity) but use validated scores instead.(48)

In the present study, we confirmed that frequency of REM and LDA was very variable according to the definition used and in particular, REM and LDA were more difficult to reach using VLDA/MDA than DAPSA-based cutoffs, as has been previously reported. (14-16) This may be because of the inclusion of diverse domains of PsA (and in particular skin involvement), or because of low cutoffs for each measure. The psychometric properties of VLDA/MDA with Boolean features also make them more strict.(38,49) Concerning agreements between these scores, kappas were also similar to the literature, with moderate agreement for REM but fair for LDA whatever the definition used.(14,15)

An original feature of our study was to cross-tabulate these composite measures with the patient's perspective as an external anchor. To provide data on using one measure over another is of great importance since there is no consensus on what measure should be used in PsA. Kappa agreements were moderate to good for both of the scores, and did not allow us to conclude. However, the comparison of these scores against patient-defined

status, performed here for the first time, was very informative. We found that more patients in patient-perceived good status were also in DAPSA-based good status, both for REM, LDA, and the combination. Of note, we advocate that REM should be the treatment goal, in accordance with recommendations; however the exploration of REM/LDA was also valuable.(1-4) In our study, patient-perceived REM/LDA occurred slightly more frequently as DAPSA-based definitions, with VLDA/MDA being rarer. DAPSA-based REM or REM/LDA had much higher sensitivity than VLDA/MDA against the reference of the patient-defined status, with only a slight loss of specificity. This means that DAPSA-based definitions correctly 'detected' much more patients in patient-defined REM or REM/LDA than VLDA/MDA. However, there were slightly more patients in DAPSA-based good status who did not report themselves in good status, than among patients in VLDA/MDA (as illustrated for REM/LDA in Figure 2). Thus each of these scores has different strengths depending on if the objective is sensitivity (i.e., to detect patient-defined good status: here DAPSA performed better) or specificity (i.e., to avoid over-detecting patients who did not self-report as doing well: here, VLDA/MDA performed better). However overall DAPSA-based cutoffs seemed to align better with the patient's perspective. These results suggest that DAPSA-based status is closer to patients' expectations than VLDA/MDA.

In conclusion, this international study of PsA disease activity highlights several important concepts regarding REM and LDA, including the aspect of truthfulness of the measures evaluated. Further studies of patients' expectations and studies demonstrating the prognostic value of different disease states/definitions for long-term outcomes, are needed to inform treatment targets.

Acknowledgements

We wish to thank all the patients who participated in the study and medical staff of all participating centers (in particular K. Fedorov, Germany).

We gratefully acknowledge the help of our patient research partners: Heidi Bertheussen (Norway), Laurence Carton (France) and Jim Walker (Scotland).

References

1. Gossec L, Smolen JS, Ramiro S, Wit M de, Cutolo M, Dougados M, et al. European League Against Rheumatism (EULAR) recommendations for the management of psoriatic arthritis with pharmacological therapies: 2015 update. Ann Rheum Dis. 2016;75(3):499–510.

2. Coates LC, Kavanaugh A, Mease PJ, Soriano ER, Laura Acosta-Felquer M, Armstrong AW, et al. Group for Research and Assessment of Psoriasis and Psoriatic Arthritis 2015 Treatment Recommendations for Psoriatic Arthritis. Arthritis Rheumatol 2016;68(5):1060–71.

3. Smolen JS, Schöls 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.

4. Gossec L, Coates LC, de Wit M, Kavanaugh A, Ramiro S, Mease PJ,et al. Management of psoriatic arthritis in 2016: a comparison of EULAR and GRAPPA recommendations. Nat Rev Rheumatol. 2016;12(12):743-750.

5. 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.

6. Coates LC, Helliwell PS. Validation of minimal disease activity criteria for psoriatic arthritis using interventional trial data. Arthritis Care Res. 2010;62(7):965–9.

7 Coates LC, Mease PJ, Gossec L, Kirkham B, Sherif B, Gaillez C, et al. Minimal Disease Activity among Active Psoriatic Arthritis Patients Treated with Secukinumab: 2-year Results from the FUTURE 2 Study. Arthritis Care Res (Hoboken). 2018 Feb 6 (epub)

8. 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.

9. Smolen JS, Schoels M, Aletaha D. Disease activity and response assessment in psoriatic arthritis using the Disease Activity index for PSoriatic Arthritis (DAPSA). A brief review. Clin Exp Rheumatol. 2015;33(5 Suppl 93):S48-50.

10 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.

11. Kerschbaumer A, Baker D, Smolen JS, Aletaha D. The effects of structural damage on functional disability in psoriatic arthritis. Ann Rheum Dis. 2017;76(12):2038–45.

12. Gossec L, McGonagle D, Korotaeva T, Lubrano E, de Miguel E, Østergaard M, et al. Minimal Disease Activity as a Treatment Target in Psoriatic Arthritis: A Review of the Literature. J Rheumatol. 2018;45(1):6–13.

13. 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.

14. Wervers K, Vis M, Tchetverikov I, Gerards AH, Kok MR, Appels CWY, et al. Burden of Psoriatic Arthritis in different definitions of disease activity: comparing Minimal Disease Activity and Disease Activity index for Psoriatic Arthritis. Arthritis Care Res. 2018 Apr 2; (epub)

15. van Mens LJJ, van de Sande MGH, van Kuijk AWR, Baeten D, Coates LC. Ideal target for psoriatic arthritis? Comparison of remission and low disease activity states in a real-life cohort. Ann Rheum Dis. 2018;77(2):251–7.

16. Smolen J, Aletaha D, Gladman D, Zhang Y, Ganz F. Outcomes associatded with achievement of various treatment targets in patients with psoriatic arthritis receiving adalimumab. Ann Rheum Dis 2017; 76 (suppl. 2):677.

17. Tillett W, Adebajo A, Brooke M, Campbell W, Coates LC, FitzGerald O, et al. Patient involvement in outcome measures for psoriatic arthritis. Curr Rheumatol Rep. 2014;16(5):418.

18. Helliwell PS, FitzGerald O, Fransen J. Composite disease activity and responder indices for psoriatic arthritis: a report from the GRAPPA 2013 meeting on development of cutoffs for both disease activity states and response. J Rheumatol. 2014;41(6):1212-7.

19. Desthieux C, Granger B, Balanescu AR, Balint P, Braun J, Canete JD, et al. Determinants of Patient-Physician Discordance in Global Assessment in Psoriatic Arthritis: A Multicenter European Study.Arthritis Care Res (Hoboken). 2017;69(10):1606-1611.

20. Michelsen B, Kristianslund EK, Hammer HB, Fagerli KM, Lie E, Wierød A, et al. Discordance between tender and swollen joint count as well as patient's and evaluator's global assessment may reduce likelihood of remission in patients with rheumatoid arthritis and psoriatic arthritis: data from the prospective multicentre NOR-DMARD study. Ann Rheum Dis. 2017;76(4):708-711

21. van Mens LJJ, Turina MC, van de Sande MGH, Nurmohamed MT, van Kuijk AWR, Baeten DLP. Residual disease activity in psoriatic arthritis: discordance between the rheumatologist's opinion and minimal disease activity measurement. Rheumatology (Oxford). 2018;57(2):283-290.

22. Dures E, Hewlett S, Lord J, Bowen C, McHugh N, PROMPT Study Group, et al. Important Treatment Outcomes for Patients with Psoriatic Arthritis: A Multisite Qualitative Study. The Patient. 2017 Aug;10(4):455–62.

23. Tillett W, Dures E, Hewlett S, Helliwell PS, FitzGerald O, Brooke M, et al. A Multicenter Nominal Group Study to Rank Outcomes Important to Patients, and Their Representation in Existing Composite Outcome Measures for Psoriatic Arthritis. J Rheumatol. 2017;44(10):1445–52.

24. Tubach F, Ravaud P, Martin-Mola E, Awada H, Bellamy N, Bombardier C, et al. Minimum clinically important improvement and patient acceptable symptom state in pain and function in rheumatoid arthritis, ankylosing spondylitis, chronic back pain, hand osteoarthritis, and hip and knee osteoarthritis: Results from a prospective multinational study. Arthritis Care Res. 2012 Nov;64(11):1699–707.

25. Heiberg T, Kvien TK, Mowinckel P, Aletaha D, Smolen JS, Hagen KB. Identification of disease activity and health status cut-off points for the symptom state acceptable to patients with rheumatoid arthritis. Ann Rheum Dis. 2008 Jul;67(7):967–71.

26. Lubrano E, Perrotta FM, Parsons WJ, Marchesoni A. Patient's Global Assessment as an Outcome Measure for Psoriatic Arthritis in Clinical Practice: A Surrogate for Measuring Low Disease Activity? J Rheumatol. 2015;42(12):2332-8.

27. Gossec L, Cantagrel A, Soubrier M, Berthelot J-M, Joubert J-M, Combe B, et al. An e-health interactive self-assessment website (Sanoia®) in rheumatoid arthritis. A 12-month randomized controlled trial in 320 patients. Jt Bone Spine Rev Rhum. 2017 Dec 12; (epub)

28. Toupin-April K, Barton J, Fraenkel L, Li L, Grandpierre V, Guillemin F, et al. Development of a Draft Core Set of Domains for Measuring Shared Decision Making in Osteoarthritis: An OMERACT Working Group on Shared Decision Making. J Rheumatol. 2015;42(12):2442–7.

29. Groll DL, To T, Bombardier C, Wright JG. The development of a comorbidity index with physical function as the outcome. J Clin Epidemiol. 2005 Jun;58(6):595–602.

30. Ramiro S, Smolen JS, Landewé R, Heijde DV, Gossec L. How are enthesitis, dactylitis and nail involvement measured and reported in recent clinical trials of psoriatic arthritis? A systematic literature review. Ann Rheum Dis. 2018;77(5):782-783.

31. Tälli S, Etcheto A, Fautrel B, Balanescu A, Braun J, Cañete JD, et al. Patient global assessment in psoriatic arthritis - what does it mean? An analysis of 223 patients from the Psoriatic arthritis impact of disease (PsAID) study. Jt Bone Spine Rev Rhum. 2016;83(3):335–40.

32. Cauli A, Gladman DD, Mathieu A, Olivieri I, Porru G, Tak PP,et al. Patient global assessment in psoriatic arthritis: a multicenter GRAPPA and OMERACT study. J Rheumatol. 2011;38(5):898-903.

331. 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.

34. Arnold MB, Khanna D, Denton CP, van Laar JM, Frech TM, Anderson M et al. Patient acceptable symptom state in scleroderma: results from the tocilizumab compared with placebo trial in active diffuse cutaneous systemic sclerosis. Rheumatology (Oxford). 2018;57:152-157.

35. 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.

36. Rasch LA, Boers M, Hill CL, Voshaar M, Hoogland W, de Wit M, et al. Validating Rheumatoid Arthritis Remission Using the Patients' Perspective: Results from a Special Interest Group at OMERACT 2016. J Rheumatol. 2017;44(12):1889–93.

37. Cheung PP, de Wit M, Bingham CO, Kirwan JR, Leong A, March LM, et al. Recommendations for the Involvement of Patient Research Partners (PRP) in OMERACT Working Groups. A Report from the OMERACT 2014 Working Group on PRP. J Rheumatol. 2016;43(1):187–93.

38. Felson DT, Smolen JS, Wells G, Zhang B, van Tuyl LHD, Funovits J, et al. American College of Rheumatology/European League Against Rheumatism provisional definition of remission in rheumatoid arthritis for clinical trials. Ann Rheum Dis. 2011;70(3):404-13.

39. Cohen J. A Coefficient of Agreement for Nominal Scales. Educ Psychol Meas. 19601;20(1):37–46.

40. Byrt T, Bishop J, Carlin JB. Bias, prevalence and kappa. J Clin Epidemiol. 1993;46(5):423–9.

41. Helliwell PS, Waxman R. Modification of the Psoriatic Arthritis Disease Activity Score (PASDAS). Ann Rheum Dis. 2018;77(3):467–8.

42. Helliwell PS, FitzGerald O, Fransen J, Gladman DD, Kreuger GG, Callis-Duffin K, et al. The development of candidate composite disease activity and responder indices for psoriatic arthritis (GRACE project). Ann Rheum Dis. 2013 Jun;72(6):986–91.

43.de Vlam K2, Merola JF, Birt J, Sandoval DM, Lobosco S, Moon R, et al. Skin Involvement in Psoriatic Arthritis Worsens Overall Disease Activity, Patient-Reported Outcomes, and Increases Healthcare Resource Utilization: An Observational, Cross-Sectional Study. Rheumatol Ther. 2018 Jul 6. doi: 10.1007/s40744-018-0120-8. [Epub ahead of print]

44. Gudu T, Gossec L. Quality of life in psoriatic arthritis. Expert Rev Clin Immunol. 2018;14(5):405–17.

45. 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.

46. Hifinger M, Putrik P, Ramiro S, Keszei AP, Hmamouchi I, Dougados M, et al. In rheumatoid arthritis, country of residence has an important influence on fatigue: results from the multinational COMORA study. Rheumatol Oxf Engl. 2016;55(4):735–44.

47. Lindstrom Egholm C, Krogh NS, Pincus T, Dreyer L, Ellingsen T, Glintborg B, et al. Discordance of Global Assessments by Patient and Physician Is Higher in Female than in Male Patients Regardless of the Physician's Sex: Data on Patients with Rheumatoid Arthritis, Axial Spondyloarthritis, and Psoriatic Arthritis from the DANBIO Registry. J Rheumatol. 2015;42(10):1781–5.

48. Coates LC, FitzGerald O, Merola JF, Smolen J, van Mens LJJ, Bertheussen H, et al. Group for Research and Assessment of Psoriasis and Psoriatic Arthritis/Outcome Measures in Rheumatology Consensus-Based Recommendations and Research Agenda for Use of Composite Measures and Treatment Targets in Psoriatic Arthritis. Arthritis Rheumatol. 2018;70(3):345–55.

49. Einarsson JT, Willim M, Ernestam S, Saxne T, Geborek P, Kapetanovic MC. Prevalence of sustained remission in rheumatoid arthritis: impact of criteria sets and disease duration, a Nationwide Study in Sweden. Rheumatol Oxf Engl. 2018 Mar 12 (epub);

Table 1 Composite indices used to define REM and LDA in PsA

Index	Components	Cut-off for REM	Cut-off for LDA
VLDA/MDA	Tender joints (≤1) Swollen joints (≤1) Skin psoriasis (PASI ≤1 or BSA ≤3%) Entheses (≤1) Pain (≤15) Patient global for joints and skin (≤20) HAQ (≤0.5)	VLDA: 7/7 of the criteria	MDA: 5/7 of the criteria
DAPSA	Tender joints Swollen joints Pain Patient global assessment CRP	DAPSA remission ≤4	DAPSA LDA: 5 to ≤14
c-DAPSA	Tender joints Swollen joints Pain Patient global assessment	c- DAPSA remission ≤4	c-DAPSA LDA: 5 to ≤13

REM: Remission; LDA: Low Disease Activity; MDA: Minimal Disease Activity; VLDA: Very Low Disease Activity; tender joint count on 68 joints; swollen joint count on 66 joints; PASI: Psoriasis Activity And Severity Index; BSA: body surface area; ; HAQ: Health Assessment Questionnaire; DAPSA: Disease Activity Index for Psoriatic Arthritis; cDAPSA: clinical DAPSA; CRP: C-reactive protein.

Table 2 Characteristics of 410 patients with PsA

	All (n=410)	Patients in self-defined REM (N=88)	Patients in self-defined LDA (N=180)	Patients in other disease states (N=142)
Male, n (%)	208 (50.7)	58 (65.9)	95(52.8)	55 (38.7)
Mean age, years (SD)	53.9 (12.5)	53.7 (13.5)	54.3 (12.3)	53.4 (12.1)
Mean PsA duration, years (SD)	11.2 (8.2)	11.9 (8.7)	11.3 (8.3)	10.8 (7.9)
Mean level of schooling, years (SD)	12.9 (3.4)	13.4 (3.4)	12.8 (3.5)	12.6 (3.5)
Paid work, n (%)	233(56.8)	53 (60.2)	106 (58.9)	74 (52.1)
Current smoking, n (%)	68 (16.6)	9 (10.2)	25 (13.9)	34 (23.9)
Elevated acute phase reactants (CRP >5mg/L), n (%)	156 (38.0)	23 (26.1)	60 (33.3)	73 (51.4)
Radiographic lesions according to CASPAR criteria, n (%)	124 (30.2)	26 (29.5)	51 (28.3)	47 (33.1)
Conventional synthetic DMARD intake, n (%)	243 (59.3)	54 (61.4)	112 (62.2)	77 (54.2)
Biologic DMARD intake, n (%)	233 (56.8)	53 (60.2)	108 (60.0)	72 (50.7)
Oral glucocorticoids, n (%)	67 (16.3)	10 (11.4)	26 (14.4)	31 (21.8)
Number of comorbidities, mean (SD)	1.3 (1.0)	1.4 (1.1)	1.2 (1.0)	1.3 (1.0)
No current psoriasis skin lesions, n (%)	142 (34.6)	45 (51.1)	63 (35.0)	34 (23.9)
Body surface area of psoriasis $\geq$5%, n (%)	38 (9.3)	3 (3.4)	14 (7.7)	19 (13.3)
Tender entheseal points, LEI mean (SD)	0.6 (1.4)	0.4 (1.3)	0.3 (0.9)	1.1 (1.8)
Tender joint count (0-68), mean (SD)	4.9 (9.8)	3.4 (10.6)	2.9 (6.8)	8.4 (11.5)
Swollen joint count (0-66), mean (SD)	2.4 (7.3)	0.9 (3.6)	1.6 (5.6)	4.3 (10.0)
Physician's global assessment of PsA, mean (SD)	3.1 (2.5)	1.7 (2.0)	2.6 (2.1)	4.7 (2.4)
Patient's assessment of pain (0-10), mean (SD)	4.1 (2.8)	2.2 (2.4)	3.5 (2.3)	6.2 (2.2)

Patient's global assessment of PsA (0-10), mean (SD)	4.2 (2.8)	2.4 (2.5)	3.5 (2.2)	6.2 (2.3)
DAPSA, mean (SD)	17.0 (17.7)	9.4 (15.5))	12.8 (15.0)	27.0 (17.8)
DAPSA ≤28, n (%)	344 (83.9)	84 (95.5)	167 (92.8)	49 (65.5)
HAQ (0-3), mean (SD)	0.68 (0.68)	0.36 (0.53)	0.54 (0.58)	1.06 (0.70)
PsAID12, mean (SD)	3.4 (2.5)	1.8 (1.9)	2.8 (2.1)	5.2 (2.1)

PsA: psoriatic arthritis; CRP: C-reactive protein; BSA: Body Surface Area; LEI: Leeds Enthesitis Index; DAPSA: Disease Activity Index for Psoriatic Arthritis ; HAQ: Health Assessment Questionnaire; PsAID: PsA Impact of Disease

Table 3 Agreement between different definitions of REM/LDA in 410 patients with PsA

REM					
	cDAPSA REM	VLDA	Physician-perceived REM	Patient - perceived REM	PGA ≤1
DAPSA REM	0.81 (0.87)	0.64 (0.81)	0.39 (0.49)	0.38 (0.60)	0.64 (0.76)
cDAPSA REM		0.58 (0.74)	0.44 (0.52)	0.40 (0.57)	0.73 (0.80)
VLDA			0.32 (0.46)	0.39 (0.65)	0.61 (0.76)
Physician-perceived REM				0.30 (0.41)	0.32 (0.41)
Patient - perceived REM					0.43 (0.60)

LDA					
	cDAPSA LDA	MDA	Physician-perceived LDA	Patient-perceived LDA	PGA >1 to 3
DAPSA LDA	0.77 (0.79)	0.31 (0.81)	0.24 (0.30)	0.30 (0.32)	0.28 (0.36)
cDAPSA LDA		0.23 (0.36)	0.24 (0.34)	0.25 (0.28)	0.33 (0.46)
MDA			0.12 (0.28)	0.17 (0.22)	0.14 (0.38)
Physician-perceived LDA				0.17 (0.20)	0.06 (0.25)

Results are presented as Cohen's kappa (prevalence-adjusted and bias-adjusted kappa). In cases of discrepancy the prevalence adjusted bias adjusted measures should be interpreted.

REM: remission; LDA: Low Disease Activity; PsA: psoriatic arthritis; DAPSA: Disease Activity Index for Psoriatic Arthritis; cDAPSA: clinical DAPSA; VLDA: Very Low Disease Activity; PGA: patient global assessment; MDA: minimal disease activity Patient perceived and physician perceived statuses are based on the single questions for each status.

Table 4. Assessment of sensitivity and specificity of different definitions of REM/LDA against the anchor of patient-perceived REM/LDA

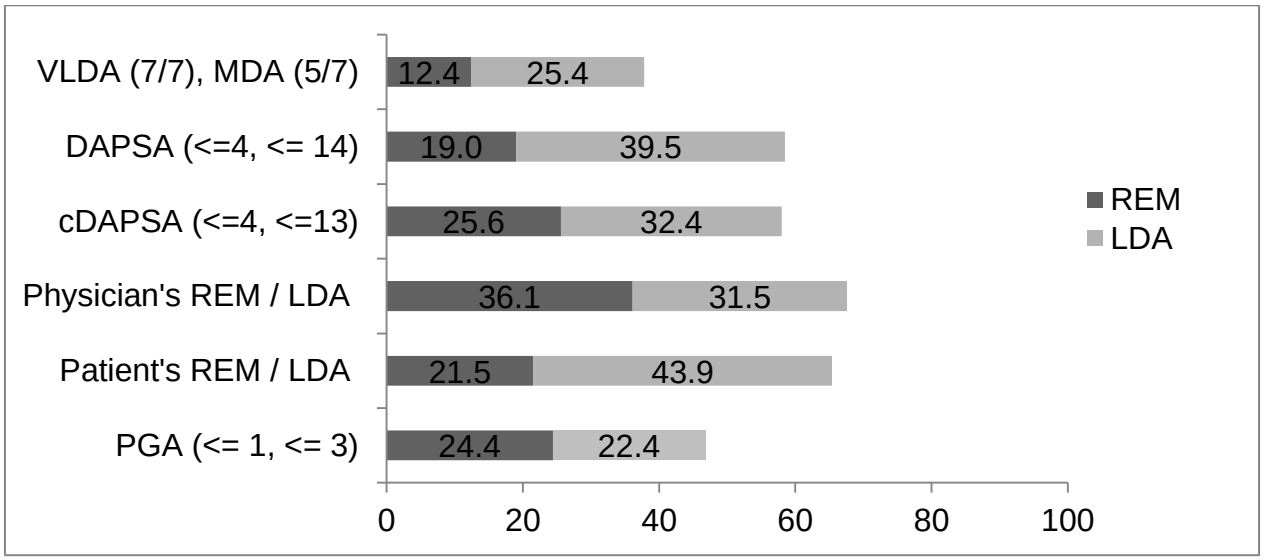
Definition tested	Property	Anchor: Patient-perceived REM	Anchor: Patient-perceived LDA	Anchor: Patient-perceived REM or LDA
VLDA/MDA	Sensitivity (N/N)	38.6% (34/88)	34.4% (62/180)	51.5% (138/268)
	Specificity (N/N)	94.7% (305/322)	81.7% (188/230)	88.0% (125/142)
DAPSA REM/LDA	Sensitivity (N/N)	47.7% (42/88)	56.1% (101/180)	73.1% (196/268)
	Specificity (N/N)	88.8% (286/322)	73.5% (169/230)	76.8% (109/142)
Physician-perceived REM/LDA	Sensitivity (N/N)	65.9% (58/88)	40.6% (73/180)	81.0% (217/268)
	Specificity (N/N)	72.0% (232/322)	75.7% (174/230)	57.7% (82/142)

Sensitivity N/N: N patients perceived as in the status by the score/ N patients in the status according to the patient-defined anchor status.

Specificity N/N: N patients perceived as NOT in the status by the score/ N patients NOT in the status according to the patients-defined anchor status.

REM: Remission; LDA: Low Disease Activity; MDA: minimal disease activity; DAPSA: Disease Activity Index for Psoriatic Arthritis.

Figure 1 Prevalence of REM/LDA according to different definitions in 410 patients with PsA



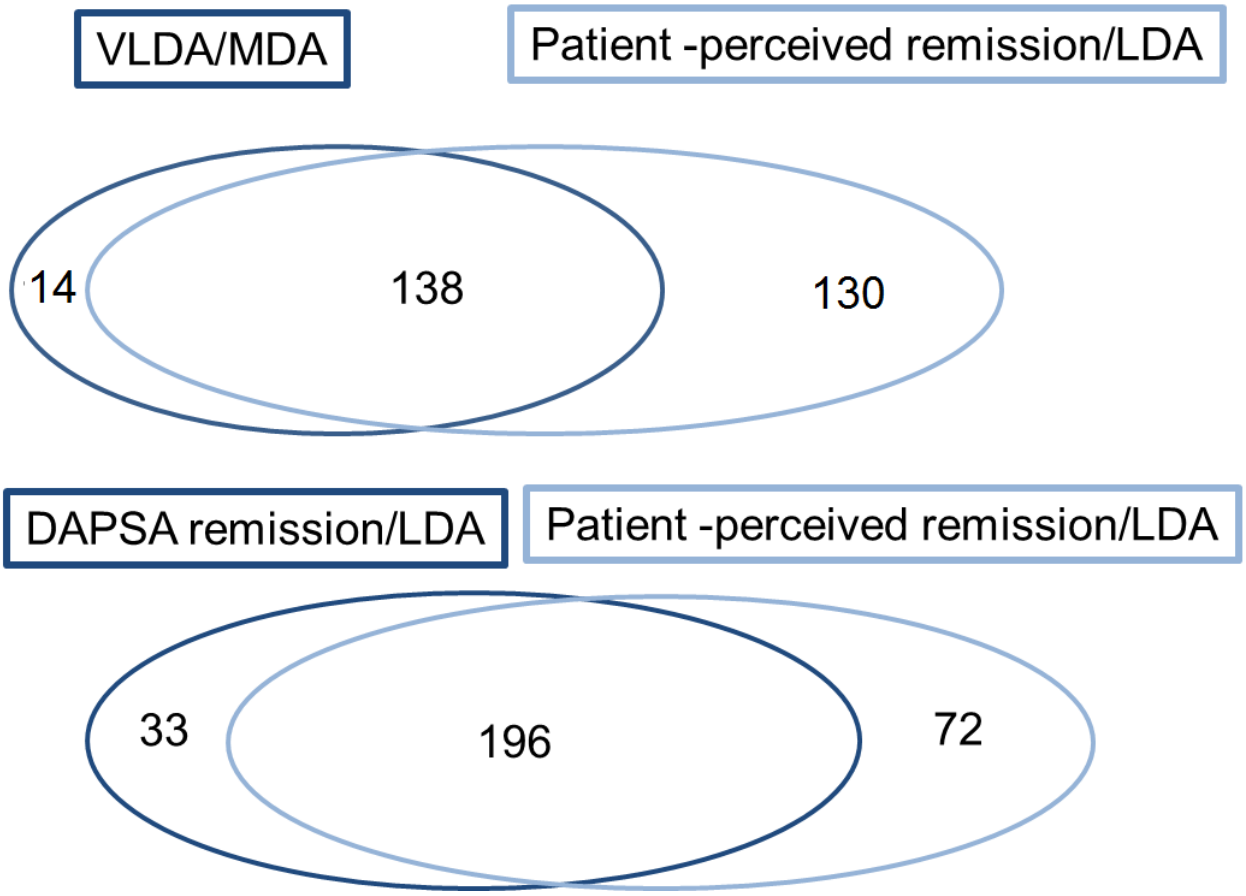
Results are presented for REM and LDA separately (without overlap of definitions).

REM: Remission; LDA: Low Disease Activity; VLDA: Very Low Disease Activity; MDA: minimal disease activity; DAPSA: Disease Activity Index for Psoriatic Arthritis; cDAPSA: clinical DAPSA; Physician's REM/LDA: physician's single question for REM/LDA; Patient's REM/LDA: patient's single question for REM/LDA; PGA: patient global assessment;.

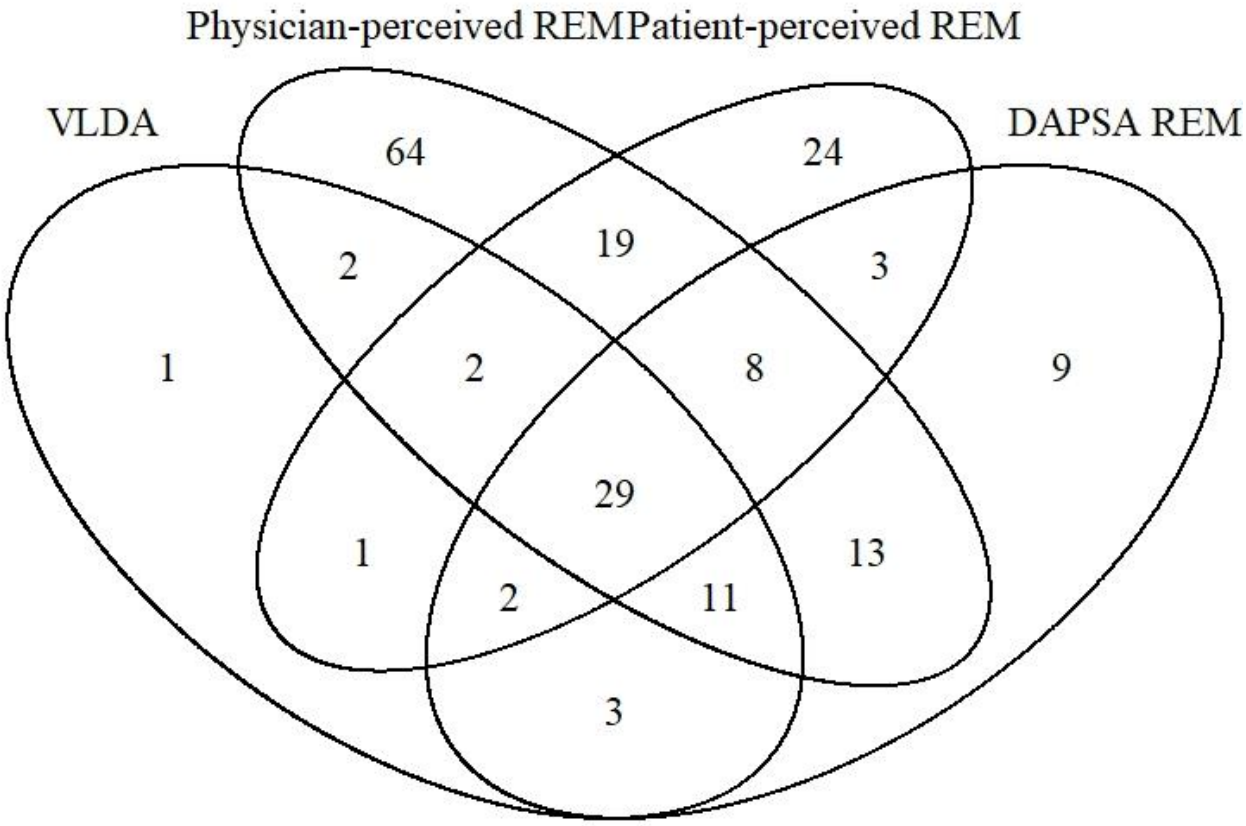
Figure 2. Venn diagram representing the number of patients meeting REM/LDA when comparing patient-perceived status and composite scores, among 410 PsA patients (of whom, 268 in patient-defined REM/LDA)

(a): VLDA/MDA versus patient perspective (sensitivity, 51.5%, specificity: 88.0%)

(b) DAPSA versus patient perspective (sensitivity, 73.1%, specificity: 76.8%)



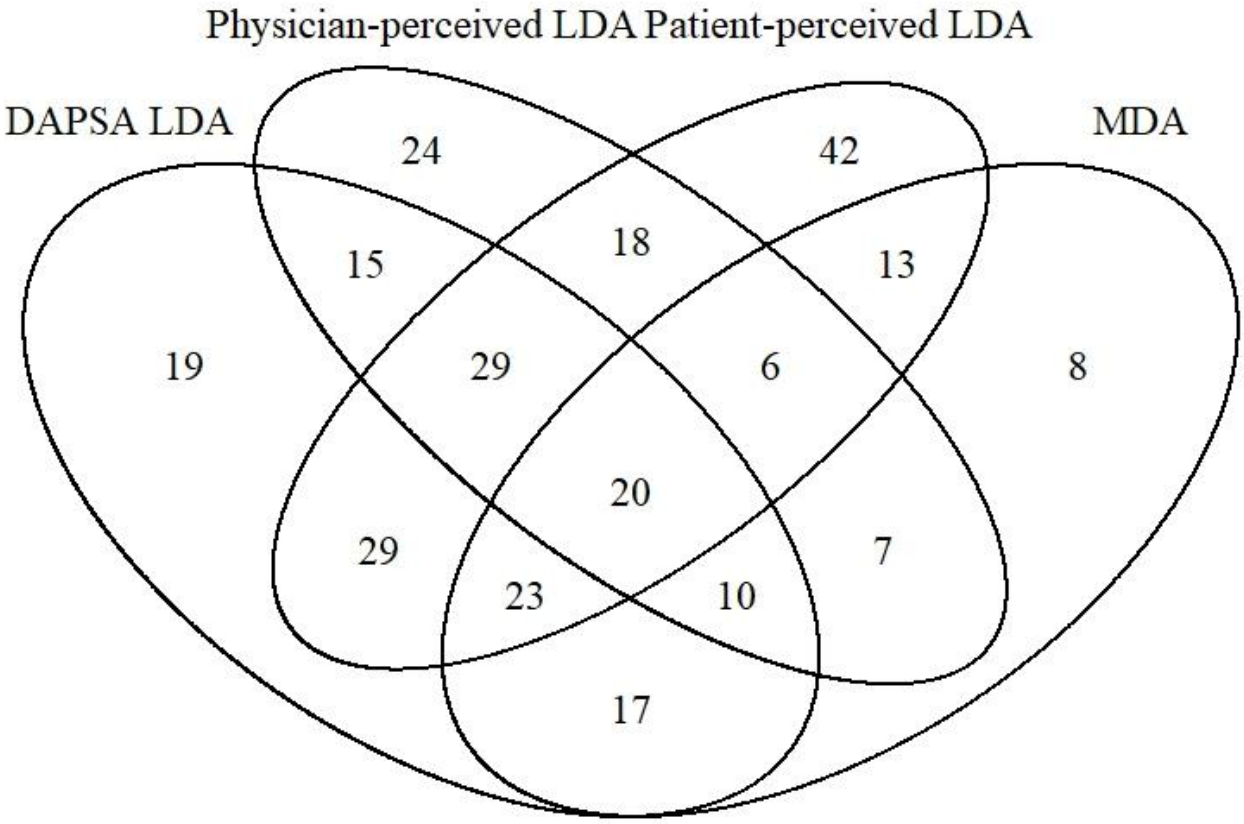
Online supplementary Figure 1 Venn diagram representing the number of patients meeting different REM criteria among 410 PsA patients



REM: Remission; VLDA: Very Low Disease Activity; DAPSA: Disease Activity Index for Psoriatic Arthritis.

The concordance between composite scores and patient-perceived REM was only moderate. Of 88 patients in patient-perceived REM, 34 (38.6%) met VLDA criteria, 42 (47.7%) were in DAPSA-REM, 43 (48.9%) were not in REM according to any composite score and 58 (65.9%) were in physician-perceived REM. Physician-perceived REM was the least stringent definition with 64/148 (43.2%) patients in physician-perceived REM, who were not in REM according to any other definition. Of 51 patients meeting VLDA criteria, 34 (66.7%) were in patient-perceived REM and of 78 patients in DAPSA REM, 42 (53.8%) were in patient-perceived REM. Furthermore, only 58/178 (32.6%) patients in patient perceived REM, were also in physician-perceived REM (and of these, 37, 63.8% were in DAPSA-REM and 31, 53.4% were in VLDA).

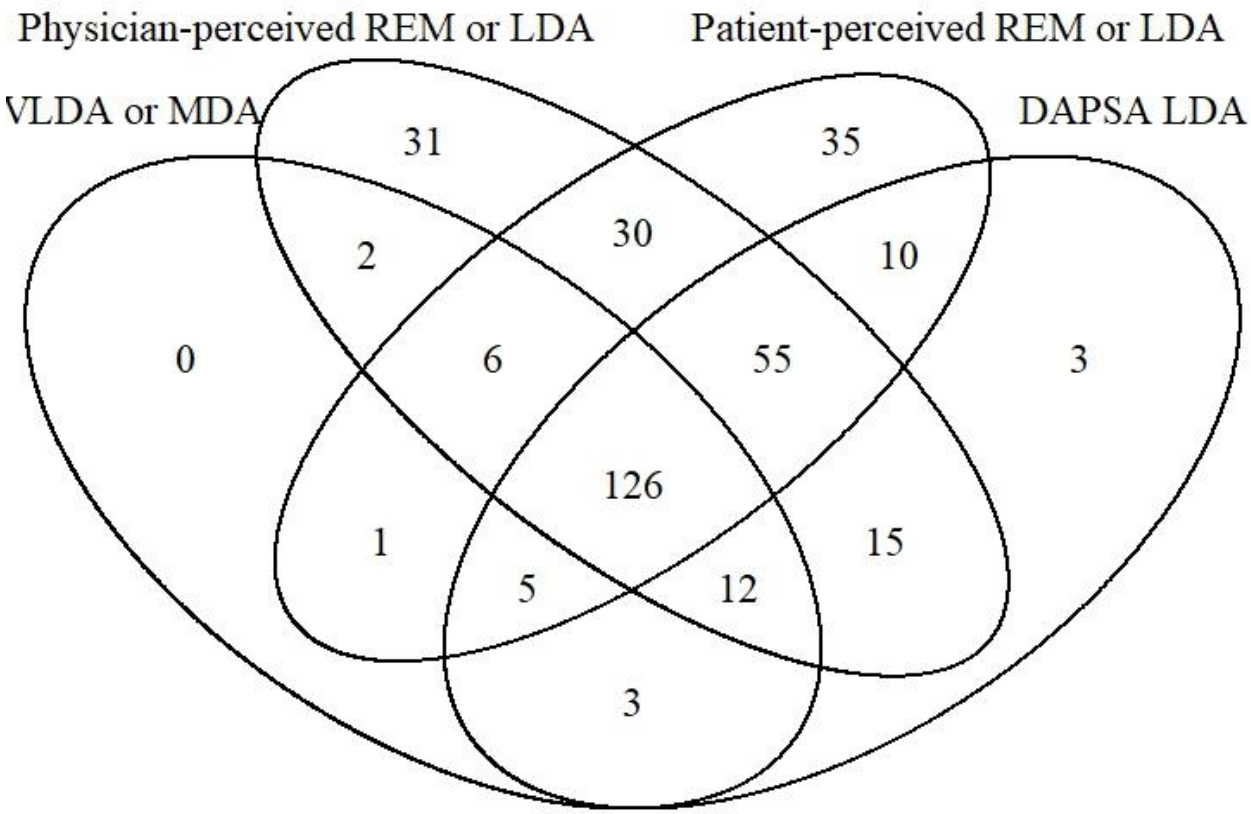
Online supplementary Figure 2 Venn diagram representing the number of patients meeting different LDA criteria among 410 PsA patients



LDA: Low Disease Activity; MDA: minimal disease activity; DAPSA: Disease Activity Index for Psoriatic Arthritis.

The concordance between composite scores and patient-perceived LDA was lower than for remission. There were 180 patients in patient-perceived LDA. Of these, 62 (sensitivity, 34.4%) met MDA criteria, 101 (56.1%) were in DAPSA-LDA, and 60 (33.3%) were not in LDA according to any composite score; 8/104 (7.7%) patients in MDA and 19/162 (11.7%) patients in DAPSA LDA were not included in another definition of low disease. Concordance was very low between physician and patient-perceived LDA. Of 104 patients in MDA, 62 (59.6%) were in patient-perceived LDA and of 162 patients in DAPSA LDA, 101 (62.3%) were in patient-perceived LDA.

Online supplementary Figure 3 Venn diagram representing the number of patients meeting different REM or LDA criteria (pooled analyses of REM/LDA) among 410 PsA patients



REM: Remission; LDA: Low Disease Activity; VLDA: Very Low Disease Activity; MDA: minimal disease activity; DAPSA: Disease Activity Index for Psoriatic Arthritis.

When analyzing as outcome, either patient-perceived REM or LDA (i.e., the sum of patients in these outcomes), DAPSA was a more inclusive definition than VLDA/MDA: among the 268 patients in patient-perceived REM/LDA, only 138 (51.5%) were in VLDA/MDA whereas 196 (73.1%) were in DAPSA REM/LDA; 65 (24.3%) were not in REM/LDA according to any composite score. Physician-defined good status (single questions) although very inclusive, did not cover well patient-defined good status.

Online supplementary Table 1. Prevalence and performance of composite scores calculated using alternative wordings of PGA

Score calculated using alternative PGA	Remission: prevalence (Se/Sp against patient question)	Low disease activity: prevalence (Se/Sp against patient question)	Remission OR Low disease activity: prevalence (Se/Sp against patient questions)
DAPSA – joints PGA	15.9% (44.3/91.9)	42.7% (61.6/71.7)	56.6% (73.1/74.6)
VLDA/MDA – joints PGA	12.4% (36.4/94.1)	24.7% (33.9/82.2)	37.1% (50.3/ 88.0)
DAPSA – skin PGA	18.3% (38.6/86.0)	44.1% (56.1/63.9)	59.3% (75.7/70.4)
VLDA/MDA - skin PGA	10.0% (29.5/94.1)	30.6% (41.7/74.8)	41.7% (56.3/84.5)

The joints PGA and skin PGA were formulated as follows: Considering all the ways your joints have affected you during the last week, circle the number that best describes how you have been doing; and Considering all the ways psoriasis (skin disease) has affected you during the last week, circle the number that best describes how you have been doing.

Sensitivity (Se) is defined as N patients in the status by the score/ N patients in the status according to the patient-defined anchor status. Specificity (Sp) is defined as N patients NOT in the status by the score/ N patients NOT in the status according to the patients-defined anchor status.