The Journal of Rheumatology 2024;51 Suppl 2:80–3
doi:10.3899/jrheum.2024-0515
First Release July 15 2024

The Journal of Rheumatology

Composite Outcome Measures for Psoriatic Arthritis: OMERACT and 3 and 4 Visual Analog Scale Progress in 2023

Ying Ying Leung[1], Dafna D. Gladman[2], Ana-Maria Orbai[3], and William Tillett[4]

ABSTRACT. The Group for Research and Assessment of Psoriasis and Psoriatic Arthritis (GRAPPA)–Outcome Measures in Rheumatology (OMERACT) psoriatic arthritis (PsA) working group provided updates at the GRAPPA 2023 annual meeting on its work to evaluate composite outcome measures for PsA. An ongoing systematic literature review is in progress to evaluate psychometric measurement properties using the OMERACT filter 2.2 for a list of candidate composite outcome measures, which include minimal disease activity (MDA), Disease Activity for Psoriatic Arthritis (DAPSA), American College of Rheumatology (ACR) response criteria, Psoriatic Arthritis Disease Activity Score (PASDAS), Composite Psoriatic Disease Activity Index (CPDAI), 3 visual analog scale (3VAS), and 4VAS. The performance of the 3VAS and 4VAS in clinical practice and a synthesis of new data were presented, including estimates for minimal clinically important differences and thresholds of meaning, discrimination and construct validity, and longitudinal construct validity. Numeric rating scale (NRS) versions of the VAS have also been tested. Performance characteristics and psychometric properties are similar to the ASSESS study, a UK multicenter study, indicating that the VAS scales may be feasible tools for routine clinical care with a preference for the 4VAS because of superior face validity and clinical utility.

Key Indexing Terms: composite outcome measures, GRAPPA, psoriasis, psoriatic arthritis

As part of the supplement series GRAPPA 2023, this report was reviewed internally and approved by the Guest Editors for integrity, accuracy, and consistency with scientific and ethical standards.

YYL is funded by the Clinician Scientist award of the National Medical Research Council, Singapore (NMRC/CSA-INV/0022/2017). The views expressed are those of the author(s) and not necessarily those of the NMRC.

[1]Y.Y. Leung, MBChB, MD, Duke-NUS Medical School, Department of Rheumatology and Immunology, Singapore General Hospital, Singapore; [2]D.D. Gladman, MD, University of Toronto, Schroeder Arthritis Institute, Krembil Research Institute, Psoriatic Arthritis Program, University Health Network, Toronto Western Hospital, Toronto, Ontario, Canada; [3]A.M. Orbai, MD, MHS, Psoriatic Arthritis Program, Division of Rheumatology, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA; [4]W. Tillett, BSc, MB ChB, PhD, Royal National Hospital for Rheumatic Diseases, University of Bath, Bath, UK.

YYL has received speaker fee from AbbVie, DKSH, Janssen, Novartis, and Pfizer. DDG has received research grants from AbbVie, Amgen, Eli Lilly, Janssen, Novartis, Pfizer, and UCB, as well as consulting fees from AbbVie, Amgen, BMS, Gilead, Eli Lilly, Janssen, Novartis, Pfizer, and UCB. AMO has received research grants to Johns Hopkins University from AbbVie, Amgen, Celgene, Gilead, Janssen, Lilly, and Novartis, and consulting fees from BMS, Janssen, and UCB. WT has received research grants from AbbVie, Amgen, Eli Lilly, Janssen, Pfizer, and UCB and consulting or speaker fees from AbbVie, Amgen, BMS, Celgene, Eli Lilly, GSK, Janssen, MSD, Novartis, Ono, Pfizer, and UCB.

This paper does not require institutional review board approval.

Address correspondence to Dr. Y.Y. Leung, Associate Professor, Department of Rheumatology and Immunology, Singapore General Hospital, 20 College Road, the Academia, Singapore S169856. Email: katyccc@hotmail.com.

Accepted for publication May 23, 2024.

Outcome Measures in Rheumatology composite measure working group updates

The Group for Research and Assessment of Psoriasis and Psoriatic Arthritis (GRAPPA)–Outcome Measures in Rheumatology (OMERACT) outcome measure working group has the objective to standardize outcome measures for psoriatic arthritis (PsA) for use in randomized controlled trials (RCTs) and longitudinal studies.[1] A working group for composite outcome measures was set up and started the appraisal of PsA composite outcome measures since 2022. The purpose of use for each composite outcome measure has been defined previously.[2] Different composite outcome measures may be required in different trial settings and scenarios. After a consensus procedure, the composite measures prioritized for evaluation are the American College of Rheumatology (ACR) response criteria, minimal disease activity (MDA), Disease Activity in Psoriatic Arthritis (DAPSA) and clinical DAPSA (cDAPSA; without C reactive protein [CRP]), Psoriatic Arthritis Disease Activity Score (PASDAS), Composite Psoriatic Disease Activity Index (CPDAI), and 3 visual analog scale (3VAS) and 4VAS.[2]

The next step in the workstream is deciding what domains should be included in the composite measures. None of the existing composites cover all the domains in the core domain set.[3] Composite outcome measures in PsA have been described previously and are summarized for reference (Supplementary Table S1, available with the online version of this article).[2,4-12] Following breakout group discussions at the GRAPPA 2019 annual meeting, preliminary voting by GRAPPA stakeholders endorsed MDA as a treatment target both for RCTs and clin-

ical practice and PASDAS as a composite endpoint in RCTs.[10] However, there is a lack of international consensus on which composite outcome measures should be used in different scenarios. Consensus across the community exists on the following aspects: (1) PsA-specific composites are needed rather than utilizing outcome measures developed for other forms of arthritis, (2) patient participation is important in the development and evaluation of composites, (3) comprehensive inclusion of PsA domains is essential, and (4) disease activity and disease impact should be measured separately.[12] Uncertainties remain regarding how to balance the combination of domains that run different clinical courses and may not track together. Considering cases where a composite outcome measure is used as an outcome in clinical trials, combining domains that do not always track together may reduce the sensitivity to change or responsiveness of the measure.[13] In contrast, when a composite outcome measure is used as a tool to evaluate the total burden of disease, combining a broader representation of domains becomes more important. A systematic literature review is in progress to evaluate the psychometric properties using the OMERACT filter 2.2[14] for candidate composite measures for use in clinical trials and longitudinal studies.

Testing of the 3VAS and 4VAS

Concomitantly, the 3VAS and 4VAS instruments have undergone the stages of development and are currently undergoing validation. Treat-to-target has been shown to improve clinical response, patient-reported outcomes, and quality of life in PsA,[15] yet a widely accepted target for use in routine clinical care remains elusive.[9] It was on this backdrop that the benefits, limitations, and barriers to a wider uptake of composite measures of PsA were the subject of a workshop at the GRAPPA 2019 annual meeting.[10] The majority (89%) of GRAPPA members agreed there was a need for a PsA-specific composite measure for routine practice; however, most were either using no measure at all in their practices or using the Disease Activity Score in 28 joints (DAS28). Breakout groups identified feasibility as the most significant barrier to wider adoption. In particular, the time taken to conduct the CPDAI or PASDAS was not feasible within the time constraints of routine care and CRP was identified as a barrier to using the DAPSA. GRAPPA members voted to test shortened versions of the CPDAI and GRAPPA Composite Exercise (GRACE) for use in routine care, and the 3VAS and 4VAS were developed and evaluated in a UK multicenter observational study.[16]

The 3VAS comprises the physician global, patient global, and patient skin assessments, and the 4VAS comprises the physician global, patient pain, patient joints, and patient skin assessments. Each score is divided by the denominator to give a 0-10 scale. The 3VAS and 4VAS were found to have superior performance characteristics to the simplified (sCPDAI), DAS28, DAPSA, and Routine Assessment of Patient Index Data 3 (RAPID3) in the ASSESS study, including responsiveness and correlation with treatment change.[16] The results of the ASSESS study were presented and discussed at the GRAPPA 2020 annual meeting and further testing in external datasets was advised. A summary of progress and new data on the 3VAS and 4VAS since the 2020 workshop are provided below.

Thresholds of meaning

Thresholds of meaning and estimates for minimal clinically important difference (MCID) and minimal detectable change (MDC) were estimated in the ASSESS dataset from UK routine care.[17] In the ASSESS study, 141 patients from 6 centers across the UK were recruited from routine care and evaluated using a panel of clinical measures and patient-reported outcome measures (PROMs) at baseline, 3 months, and 6 months.[16] Thresholds of meaning and MCID were derived using the health anchor method and 2 distribution methods. The mean of the 3 methods gave an MCID of 1.0 for both 3VAS and 4VAS. MDC was 3.1 for 3VAS and 2.5 for 4VAS.[17] Thresholds of meaning were triangulated from established cut-off values for the patient global VAS, PASDAS, and DAPSA (Figure).

External estimates of MCID, MDC, and disease activity thresholds were evaluated in posthoc pooled analyses of 3 guselkumab (GUS) phase III studies.[18] Measurement error, which is study-specific, should be considered when applying MCID estimates.[19] The analyses were generally comparable to those previously reported in the ASSESS study.[16] Further, achieving VAS states of low disease activity and remission were associated with reduced rates of structural damage progression in the DISCOVER-2 trial dataset among biologic-naïve patients treated with GUS for active PsA.[20]

Construct validity and discrimination

Construct validity, longitudinal construct validity, and discrimination were also tested in posthoc pooled analyses of 3 phase III studies of GUS in PsA. The 3VAS and 4VAS were able to discriminate between placebo and GUS treatment, and there was strong correlation with GRACE and PASDAS from baseline to follow-up at 48 weeks.[17]

3VAS and 4VAS testing in early PsA

The 3VAS and 4VAS have also been tested in the Dutch Early Psoriatic Arthritis (DEPAR) cohort as part of a GRAPPA 2021 pilot research award to Dr. Fazira Kasiem.[21] Kasiem et al evaluated 410 people with PsA treated during routine care in the DEPAR dataset. The 3VAS and 4VAS had similar responsiveness to other shorter, more feasible composite measures (DAS28 and DAPSA) and was superior to the RAPID3. The 3VAS and 4VAS had superior correlation with the PASDAS. The construct validity and responsiveness are similar between the 3VAS and 4VAS, similar to analyses in other datasets. Thresholds of meaning were evaluated using mean PROMs for physical function (Health Assessment Questionnaire [HAQ]), fatigue (Bristol Rheumatoid Arthritis Fatigue [BRAF]), quality of life (EuroQol-5 Dimension questionnaire [EQ-5D]), and impact of disease (Psoriatic Arthritis Impact of Disease [PSAID]). Mean PROM scores correlated appropriately when stratified by low, moderate, and high disease activity, providing supportive evidence for the thresholds of meaning.[22]

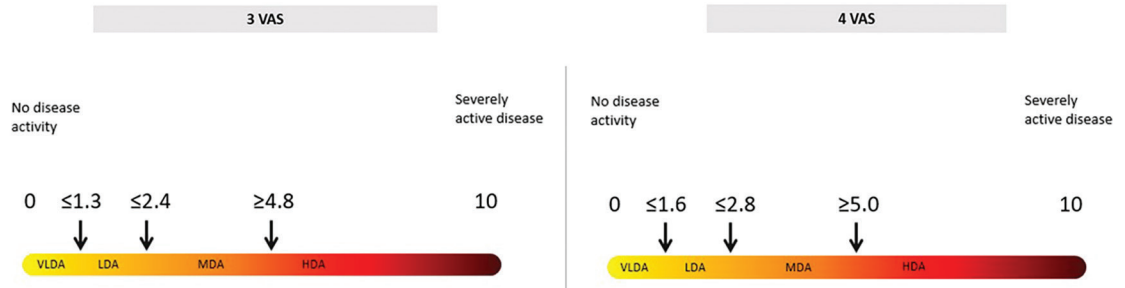

Figure. Thresholds of meaning adapted from Tillet et al.[17] 3VAS includes physician global, patient global, and patient skin assessments. 4VAS includes physician global, patient pain, patient joints, and patient skin assessments. HDA: high disease activity; LDA: low disease activity; MDA: moderate disease activity; VAS: visual analog scale; VLDA: very low disease activity.

Testing of numerical rating scale versions: 3NRS and 4NRS

Testing of numerical rating scale (NRS) versions was recommended due to increased utility and superior psychometric properties of NRS scales. NRS versions of the 3VAS and 4VAS were evaluated in the upadacitinib SELECT-PsA trial dataset. The 3NRS and 4NRS were able to discriminate between placebo and treatment groups in both biologic-naïve patients treated with upadacitinib (SELECT-PsA 1) and biologic-resistant patient populations (SELECT-PsA 2).[23] The 3NRS and 4NRS correlated well with other clinical measures and PROMs, including those focused on joints (DAPSA) or multiple manifestations (PASDAS).[23]

NRS versions have also been tested in a UK observational dataset. Data were collected prospectively across 3 UK hospital trusts from 2018 to 2019 as part of a study assessing the use of NRS in PROMs in PsA. Data from 209 patients were analyzed. There was good agreement between VAS and NRS for the patient-reported components of 3VAS and 4VAS, supporting that VAS scores are reproducible as NRS scores. Both NRS and VAS versions of the 3VAS and 4VAS scales correlate with disease activity and life impact.[24]

Physician VAS

An important component of the 3VAS and 4VAS scores is the physician global VAS (PhVAS). Work by GRAPPA has shown the PhVAS assessment to be a reliable tool to assess musculoskeletal (MSK) and dermatological disease activity in an international study of 8 countries, 16 centers, and 319 patients.[25] However, some uncertainties remain, including limited instructional material to conduct the PhVAS, little understanding of use of the PhVAS by the wider multidisciplinary team, and knowing which assessments (history/examination) should be conducted to inform a PhVAS. To address these questions, a study is underway to develop instructional material and test the PhVAS among the wider clinical team providing routine care (eg, trainees, nurses, and fellows).

Discussion

Discussion followed on the need for a feasible measure for routine care and the conflict between pushing for more thorough assessment tools routinely vs using abbreviated tools such as the 3VAS and 4VAS. Specifically, concerns were voiced that the reduction of numerous clinical assessments to the PhVAS scales may result in clinicians not examining patients thoroughly enough. The counterpoint is that when the GRAPPA membership, which comprises clinicians interested in best practices for PsA, was surveyed in 2019, members were either not using a composite measure or using the DAS28, which is inadequate in the assessment of articular disease in PsA and does not capture other domains of disease.[10]

Which instrument should be used going forward? The 4VAS had superior ability to detect treatment change (*t* score) and magnitude of response, but the 3VAS demonstrated better responsiveness using the standardized response mean in the ASSESS study. Further testing confirmed that the performance characteristics remain very similar in both clinical trial and observational datasets. Discussion leaned toward favoring the 4VAS over the 3VAS on the basis of superior face validity and clinical utility through explicit measurement of skin and joint disease as well as pain, which is a high priority for patients.

In summary, the GRAPPA-OMERACT composites working group is evaluating the psychometric properties of composite measures for clinical trials and will shortly report the findings of the systematic literature review. Testing of the 3VAS and 4VAS in prospective observational and clinical trial datasets is planned, with a view to determine which should be the preferred instrument.

ACKNOWLEDGMENT

We thank DerMEDit (www.dermedit.com) for editing services in preparation of this manuscript.

ONLINE SUPPLEMENT

Supplementary material accompanies the online version of this article.

REFERENCES

1. Tillett W, Orbai AM, Ogdie A, et al. GRAPPA-OMERACT initiative to standardise outcomes in psoriatic arthritis clinical trials and longitudinal observational studies. Ann Rheum Dis 2018;77:e23.
2. Leung YY, Tillett W, de Wit M, et al. Initiating evaluation of composite outcome measures for psoriatic arthritis: 2022 updates from the GRAPPA-OMERACT working group. J Rheumatol 2023;50:53-7.
3. Orbai AM, de Wit M, Mease P, et al. International patient and physician consensus on a psoriatic arthritis core outcome set for clinical trials. Ann Rheum Dis 2017;76:673-80.
4. Gladman DD, Mease PJ, Strand V, et al. Consensus on a core set of

domains for psoriatic arthritis. J Rheumatol 2007;34:1167-70.
5. 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;41:1212-17.
6. Coates LC, FitzGerald O, Mease PJ, et al. Development of a disease activity and responder index for psoriatic arthritis—report of the psoriatic arthritis module at OMERACT 11. J Rheumatol 2014;41:782-91.
7. Smolen JS, Braun J, Dougados M, et al. Treating spondyloarthritis, including ankylosing spondylitis and psoriatic arthritis, to target: recommendations of an international task force. Ann Rheum Dis 2014;73:6-16.
8. Smolen JS, Schöls M, Braun J, 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;77:3-17.
9. Coates LC, FitzGerald O, Merola JF, 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;70:345-55.
10. Tillett W, McHugh N, Orbai A-M, et al. Outcomes of the 2019 GRAPPA Workshop on continuous composite indices for the assessment of psoriatic arthritis and membership-recommended next steps. J Rheumatol 2020;96:11-8.
11. Gladman DD, Landewé R, McHugh NJ, et al. Composite measures in psoriatic arthritis: GRAPPA 2008. J Rheumatol 2010;37:453-61.
12. Helliwell PS, Fitzgerald O, Mease PJ. Development of composite measures for psoriatic arthritis: a report from the GRAPPA 2010 annual meeting. J Rheumatol 2012;39:398-403.
13. Wells GA, Tugwell P, Tomasson G, et al. Composite outcomes at OMERACT: multi-outcome domains and composite outcome domains. Semin Arthritis Rheum 2021;51:1370-7.
14. Maxwell LJ, Beaton DE, Boers M, et al. The evolution of instrument selection for inclusion in core outcome sets at OMERACT: Filter 2.2. Semin Arthritis Rheum 2021;51:1320-30.
15. Coates LC, Moverley AR, McParland L, et al. Effect of tight control of inflammation in early psoriatic arthritis (TICOPA): a UK multicentre, open-label, randomised controlled trial. Lancet 2015;386:2489-98.
16. Tillett W, FitzGerald O, Coates LC, et al. Composite measures for routine clinical practice in psoriatic arthritis: testing of shortened versions in a UK multicenter study. J Rheumatol Suppl 2021; 97:45-9.
17. Tillett W, Day J, Mchugh N, Fitzgerald O, Coates LC, Helliwell P. OPO224 Continuous composite measures for routine care in psoriatic arthritis: thresholds of meaning and clinically important difference estimates for the 3 and 4 VAS scales from a UK multicentre study [abstract]. Ann Rheum Dis 2021;80:135.
18. Tillett W, Coates L, Vis M, et al. 1017: Minimal important difference (MID), minimal detectable change (MDC), and disease activity thresholds for two novel composite instruments (3VAS, 4VAS) in patients with psoriatic arthritis: pooled analysis of three phase 3 studies [abstract]. Arthritis Rheumatol 2022;74.
19. Cella D, Nolla K, Peipert JD. The challenge of using patient reported outcome measures in clinical practice: how do we get there? J Patient Rep Outcomes 2024;8:35.
20. Tillett W, Coates L, Vis M, et al. Early improvement in 3 visual analogue scale (3VAS)/4VAS predicts reduced rates of radiographic change in bio-naive active psoriatic arthritis patients receiving guselkumab treatment [abstract]. Ann Rheum Dis 2023;75.
21. Kasiem FR, Yamada D, Marin J, et al. GRAPPA 2021 pilot grant award reports. J Rheumatol 2023;50:36-7.
22. Kasiem FR, Kok MR, Luime JJ, et al. Construct validity and responsiveness of feasible composite disease activity measures for use in daily clinical practice in patients with psoriatic arthritis. RMD Open 2023;9:e002972.
23. Tillett W, Coates L, Kishimoto M, et al. AB0904 Evaluating numeric rating scale versions of the 3 and 4 visual analog scale (3/4-VAS) composite measures in patients with active psoriatic arthritis from the SELECT-PsA program [abstract]. Arthritis Rheumatol 2022;81:1580-1.
24. Day J, Ye W, Tillett W, Coates LC. POS1084 Comparison of numerical rating scale (NRS) and visual analogue scale (VAS) in the patient reported outcome measures of the 3VAS and 4VAS in psoriatic arthritis [abstract]. Ann Rheum Dis 2021;80:821.
25. Cauli A, Gladman DD, Mathieu A, et al. Physician's global assessment in psoriatic arthritis: a multicenter GRAPPA study. J Rheumatol 2018;45:1256-62.