The Journal of Rheumatology 2025;52 Suppl 3:87–90
doi:10.3899/jrheum.2025-0268
First Release August 21 2025

The Journal of Rheumatology

Composite Outcome Measures for Psoriatic Arthritis: Project Updates 2024

Andre L. Ribeiro[1], Tommy Kok Annfeldt[2], Dafna D. Gladman[3], William Tillett[4], Ana-Maria Orbai[5], and Ying Ying Leung[6]

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 2024 annual meeting on their assessment of composite outcome measures for PsA. The group presented the progress of a systematic literature review on the psychometric properties of the following candidate composite outcome measures using the OMERACT filter 2.2: (1) minimal disease activity (MDA), (2) Disease Activity for Psoriatic Arthritis (DAPSA), (3) American College of Rheumatology (ACR) response criteria, (4) Psoriatic Arthritis Disease Activity Score (PASDAS), (5) Composite Psoriatic Disease Activity Index (CPDAI), (6) 3-item visual analog scale (3VAS), and (7) 4VAS. A Delphi exercise for patient research partners (PRPs) on domain match and feasibility is ongoing. Following analysis and endorsement of domain match and feasibility by PRPs, the working group will seek endorsement from the GRAPPA community. In addition, the group illustrated a new research proposal for using network metaanalysis to quantitatively compare the responsiveness of these various composite outcome measures.

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

OMERACT composite measure working group updates

The Group for Research and Assessment of Psoriasis and Psoriatic Arthritis (GRAPPA)–Outcome Measures in Rheumatology (OMERACT) working group aims to standardize outcome measures for PsA for use in randomized controlled trials (RCT) and longitudinal observational studies (LOS) to enhance quality and consistency of PsA research.[1] Composite outcome measures are essential tools in clinical trials that enhance the rigor and efficiency of these studies.[2] The GRAPPA-OMERACT working group initiated the evaluation of composite outcome measures for PsA in 2022, and the updates are reported herein.

Following a consensus process in 2023, the prioritized 7 composite outcome measures for evaluation include the American College of Rheumatology (ACR) response criteria, minimal disease activity (MDA), Disease Activity in Psoriatic Arthritis (DAPSA), Psoriatic Arthritis Disease Activity Score (PASDAS), Composite Psoriatic Disease Activity Index (CPDAI), the 3-item visual analog scale (3VAS), and 4VAS measures.[3] The specific purpose of use for each of these measures and the patient population for which each should be applied has been clearly defined in earlier work.[4] For instance, MDA is a composite for patients with PsA to be used in RCT or LOS as a responder index to assess low disease activity.

Systematic literature review on 5 psychometric properties of the composite outcome measures. The working group completed a systematic literature review to assess the measurement properties of the composite outcome measures. The initial literature search identified 1450 articles, of which 65 were included in the final analysis. Three researchers (ALR, R. Holland, and YYL) worked in pairs, independently screened titles, abstracts, and full text (if appropriate) for eligibility. Disputes were resolved by consensus of the 3 researchers via webinar. The OMERACT filter 2.2 was applied to assess the following 5 measurement properties: construct validity, test-retest reliability, longitudinal construct, RCT discrimination, and threshold of meaning.[5] The evidence to support each measurement property was synthesized based on the number of quality studies, performance of the measurement property, and the consistency of data across all studies, and was presented in a summary of measurement properties (SOMP) table for each composite outcome measure.[6] The SOMP tables illustrated the number of studies and detailed supporting evidence for the 5 measurement properties.[6] The final ratings for each measurement property were rated as GREEN (good evidence supporting this property; passes this element of the OMERACT filter 2.2); AMBER (some caution, but good enough); RED (stop; evidence against this property or only poor-quality evidence); or WHITE (no or inadequate data).

[1]A.L. Ribeiro, MD, MSc, Women's College Hospital, University of Toronto, Toronto, Ontario, Canada; [2]T.K. Annfeldt, MSc Pharm, Section for Biostatistics and Evidence-Based Research, The Parker Institute, Bispebjerg and Frederiksberg Hospital, Copenhagen, Denmark; [3]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; [4]W. Tillet, MB ChB, PhD, Royal National Hospital for Rheumatic Diseases, University of Bath, Bath, UK; [5]A.M. Orbai, MD, MHS, Psoriatic Arthritis Program, Division of Rheumatology, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA; [6]Y.Y. Leung, MB ChB, MD, Duke-NUS Medical School, Department of Rheumatology and Immunology, Singapore General Hospital, Singapore.

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

Accepted for publication March 9, 2025.

In summary, PASDAS had GREEN ratings for all 5 measurement properties. There were no data for test-retest reliability for MDA. The working group evaluated test-retest reliability and/or interobserver reliability for each component of MDA and assigned an overall GREEN rating. DAPSA and MDA had GREEN ratings for 4 out of 5 measurement properties and AMBER for the threshold of meaning. There was a general lack of data on measurement properties for ACR responder criteria, 3VAS, and 4VAS (Figure 1).

Domain match and feasibility of the composite outcome measures. The working group held virtual discussions to appraise domain match and feasibility of composite outcome measures and achieved consensus through 2 rounds of Delphi exercises. The working group members were given the SOMP tables, with the 5 measurement properties for each composite outcome measure, and voted on the following items:

1. Given your evaluation of the data, can you indicate the domain match for each "composite instrument" to be used for its "intended purpose"?
2. Given your evaluation of the data, can you assess the feasibility of using each composite outcome measure in clinical trials?

These questions could be answered as yes, "uncertain," or no, with consensus defined as the option chosen by at least 70% of participants in the Delphi exercise. If there was consensus on yes for domain match (item 1 above) and/or feasibility (item 2 above), the respective measurement property received a GREEN designation.

In the first Delphi round, MDA and DAPSA received a GREEN rating for both domain match (81.8% and 77.3%, respectively) and feasibility (100% and 86.4%, respectively) based on working group consensus on yes. PASDAS received a GREEN rating for domain match. However, there was no consensus on domain match for ACR20/50/70, 3VAS, and 4VAS. The CPDAI received a yes vote for domain match from 59.1% of the working group, but feasibility was challenged for both CPDAI and PASDAS (Figure 2).

In the second Delphi round, we explored if the lack of consensus was related to differences between RCT vs LOS settings. Working group members were asked to rate CPDAI, PASDAS, and 3VAS or 4VAS separately in RCT vs LOS settings. Consensus (85.7%) for yes and therefore (GREEN) designation in the feasibility domain was achieved for PASDAS in the RCT setting (Figure 2).

Further steps toward project completion. In the near future, the working group will seek the perspectives of patient research partners (PRPs) on domain match and feasibility. After obtaining endorsement from PRPs, the working group will engage the GRAPPA community to endorse the final composite outcome measures for PsA. The working group is focusing on using the current data to guide optimizing the presentation of DAPSA, MDA, and PASDAS for endorsement.

Comparative responsiveness of composite outcome measures. The OMERACT filter evaluates "discrimination of instruments" qualitatively. However, the quantitative comparison of composite outcome measures in RCTs remains unclear. To address this, a new metaepidemiological study is being conducted to determine which of the 7 selected composite outcome measures has the greatest odds of responding to experimental intervention in PsA RCTs. Although network metaanalysis is commonly used to compare the efficacy of 3 or more interventions across a network of studies, it is novel in evaluating the comparative responsiveness of outcome measures. The working group has previously employed a similar methodology to evaluate the comparative responsiveness of physical function outcome measures in PsA, with a manuscript reporting these findings currently under finalization for publication.

A systematic literature review will be conducted to identify articles that have reported change scores of composite outcome measures in RCTs that evaluate biologic or targeted synthetic disease-modifying antirheumatic drugs against a placebo comparator in patients with PsA. The standardized mean difference (SMD) will be used to analyze and compare the various composite outcome measures across trials through network metaanalysis. The SMD between intervention and comparator groups, along with the corresponding standard error (SE [SMD]),[7] represents a relative measure of the net benefit relative to the variability in the outcome measure among participants within the same study (Figure 3).

This study faces the added complexity of dealing with both dichotomous (MDA, ACR20) and continuous composite outcome measures (DAPSA, PASDAS, CPDAI, 3VAS, 4VAS). For dichotomous composite outcome measures, the odds ratio (OR) and the corresponding SE [ie, log(OR) and SE(log(OR))] will be used as summary measures. For continuous composite outcome measures, the SMDs will be converted to the corresponding OR to facilitate comparison across dichotomous and continuous scales.[8] This conversion allows for a unified analysis of the comparative effectiveness of these diverse outcome measures.

DISCUSSION

A discussion followed the presentations of these projects and findings at the GRAPPA 2024 annual meeting. Concerns were raised that skin domains were not included in some of the composite outcome measures. The working group acknowledged that determining which domains should be included in a composite outcome measure is a complex issue that has been widely discussed.[3] When developing PASDAS, for example, the skin domain did not appear to influence clinical decision making with regard to treatment change and was therefore excluded from the measure in a data-driven process.[9] Further, post hoc analysis of data from a large registry and a cohort study suggested that skin involvement may be indirectly represented in PASDAS through the patient and physician global VAS scores.[10] The discussion highlighted the importance of clearly defining the population of interest and the type of study being conducted, as these factors influence the choice of composite outcome measures and specific instruments. It was also noted that composite outcome measures complement individual measures rather than replace them, as all items present in the core domain set are still required to be measured in all RCTs.[11]

	Results from WG Delphi (Await PRP)		Results from systematic literature review					
Composites for PsA	**Truth**	**Feasibility**	**Truth**	**Discrimination**				
	Domain match	Feasibility	Construct validity	Test-retest reliability	Long'l construct validity	Clinical trial discriminat ion	Thresholds of meaning	**(Tentative) Overall**
ACR20/50/70	No consensus	GREEN	AMBER	NA	NA	NA	NA	**NA**
MDA	GREEN	GREEN	GREEN	GREEN	GREEN	GREEN	AMBER	**AMBER**
CPDAI	No consensus (or AMBER)	No consensus (or AMBER for RCT)	AMBER	AMBER	GREEN	GREEN	AMBER	**NA**
DAPSA	GREEN	GREEN	GREEN	AMBER	GREEN	GREEN	AMBER	**AMBER**
PASDAS	GREEN	GREEN (RCT)	GREEN	GREEN	GREEN	GREEN	GREEN	**GREEN** (RCT)
3VAS, 4VAS	No consensus (or AMBER)	GREEN	AMBER	NA	AMBER	NA	NA	**NA**

Figure 1. Summary of overall measurement properties and results of domain match and feasibility from Delphi exercises. Properties were rated GREEN (good evidence supporting this property; passes this element of the OMERACT filter 2.2), AMBER (some caution, but good enough), or WHITE (no/inadequate data). 3VAS: 3-item visual analog scale; 4VAS: 4-item visual analog scale; ACR: American College of Rheumatology response criteria; CPDAI: Composite Psoriatic Disease Activity Index; DAPSA: Disease Activity in Psoriatic Arthritis; Long'l: longitudinal; MDA: minimal disease activity; NA: not enough data to answer; PASDAS: Psoriatic Arthritis Disease Activity Score; PRP: patient research partner; PsA: psoriatic arthritis; RCT: randomized clinical trial; WG: working group.

	Delphi 1 Response rate 95.7%		Delphi 2 Response rate 91.3%				**Overall**	
			RCT		LOS			
Composites for PsA	**Truth**	**Feasibility**	**Truth**	**Feasibility**	**Truth**	**Feasibility**	**Truth**	**Feasibility**
	Domain match	Feasibility	Domain match	Feasibility	Domain match	Feasibility	Domain match	Feasibility
ACR20/50/70	No consensus Yes: 40.9% Uncertain: 13.6% No: 45.5%	GREEN (yes: 81.8%)	-	-	-	-	No consensus	GREEN
MDA	GREEN (yes: 81.8%)	GREEN (Yes: 100%)	-	-	-	-	GREEN	GREEN
CPDAI	No consensus Yes: 59.1% Uncertain: 22.7% No: 18.2%	No consensus Yes: 31.8% Uncertain: 36.4% No: 31.8%	AMBER Yes: 59.1% Uncertain: 28.6% No: 9.5%	AMBER Yes: 57.1% Uncertain: 28.6% No: 14.3%	No consensus Yes=47.6% Uncertain=38.1% No=14.3%	No consensus Yes: 38.1% Uncertain: 23.8% No: 38.1%	No consensus (or AMBER)	No consensus (or AMBER for RCT)
DAPSA	GREEN (Yes: 77.3%)	GREEN (Yes: 86.4%)	-	-	-	-	GREEN	GREEN
PASDAS	GREEN (Yes: 77.3%)	No consensus Yes: 54.6% Uncertain: 31.8% No: 13.6%	-	GREEN (yes=85.7%)	-	Yes: 52.4% Uncertain: 28.6% No: 19.1%	GREEN	GREEN (RCT)
3VAS, 4VAS	No consensus Yes: 31.8% Uncertain: 40.9% No: 27.3%	GREEN (Yes: 77.3%)	AMBER *Yes: 52.4% *Uncertain 33.3% *No: 14.3%	-	No consensus Yes: 33.3% Uncertain:38.1% No: 28.6%	-	No consensus (or AMBER)	GREEN

* Clinical practice

Figure 2. Results of working group Delphi for domain match and feasibility for the composite outcome measures. Properties were rated GREEN (good evidence supporting this property; passes this element of the OMERACT filter 2.2), AMBER (some caution, but good enough), or WHITE (no/inadequate data). 3VAS: 3-item visual analog scale; 4VAS: 4-item visual analog scale; ACR: American College of Rheumatology response criteria; CPDAI: Composite Psoriatic Disease Activity Index; DAPSA: Disease Activity in Psoriatic Arthritis; LOS: longitudinal observational studies; MDA: minimal disease activity; PASDAS: Psoriatic Arthritis Disease Activity Score; PsA: psoriatic arthritis; RCT: randomized clinical trial.

$$SMD = \frac{\bar{x}_1 - \bar{x}_2}{SDp}$$

Figure 3. The formula for SMD. $\bar{x}_1$ represents the mean difference in the intervention group; $\bar{x}_2$ represents the mean difference in the comparator group; SD_p represents the pooled SD. SMD: standardized mean difference.

In summary, the GRAPPA-OMERACT working group presented their progress toward recommending composite outcome measures for PsA. After seeking PRPs' perspectives on domain match and feasibility, DAPSA, MDA, and PASDAS are likely candidates for endorsement by the GRAPPA community. A new network metaanalysis is underway to address the comparative responsiveness of these composite outcome measures.

ACKNOWLEDGMENT

We thank DerMEDit (www.dermedit.com) for editing services in preparation of this manuscript. We thank the researchers who participated in the composite working group and the systematic literature review for measurement properties for the 7 composite outcome measures: Ying-Ying Leung, Tommy Kok Annfeldt, Richard Holland, André Lucas Ribeiro, Shounak Ghosh, Lourdes Maria Perez Chada, Maria Sole Chimenti, Latika Gupta, Philip S. Helliwell, Dafna D. Gladman, William Tillett, Ana-Maria Orbai, Laura C. Coates, Maarten de Wit, Christine Lindsay, Niti Goel, Philip Mease, Alexis Ogdie, Laure Gossec, Oliver FitzGerald, Joseph Merola, Vibeke Strand, Robin Christensen. Their detail engagement will be reported in the final systematic literature review article.

FUNDING

TKA is supported by a core grant from the Oak Foundation (OFIL-24-074) to the Section for Biostatistics and Evidence-Based Research, the Parker Institute, Bispebjerg and Frederiksberg Hospital. YYL is funded by the Clinician Scientist award of the National Medical Research Council, Singapore (MOH-CSAINV24jul-0007). The views expressed are those of the author(s) and not necessarily those of the NMRC.

COMPETING INTERESTS

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. 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-Pharma, 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. YYL has received speaker fees from AbbVie, DKSH, Janssen, Novartis, and Pfizer. The remaining authors declare no conflicts of interest relevant to this article.

ETHICS AND PATIENT CONSENT

Institutional review board approval and patient consent were not required.

PEER REVIEW

As part of the supplement series GRAPPA 2024, this report was reviewed internally and approved by the Guest Editors for integrity, accuracy, and consistency with scientific and ethical standards.

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. 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.
3. Leung YY, Gladman DD, Orbai A-M, Tillett W. Composite outcome measures for psoriatic arthritis: OMERACT and 3 and 4 visual analog scale progress in 2023. J Rheumatol 2024;51 Suppl 2:80-3.
4. Leung Y-Y, 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 Suppl 2:53-7.
5. Beaton DE, Maxwell LJ, Shea BJ, et al. Instrument selection using the OMERACT Filter 2.1: the OMERACT methodology. J Rheumatol 2019;46:1028-35.
6. 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.
7. Wells G, Li T, Maxwell L, Maclean R, Tugwell P. Responsiveness of patient reported outcomes including fatigue, sleep quality, activity limitation, and quality of life following treatment with abatacept for rheumatoid arthritis. Ann Rheum Dis 2008;67:260-5.
8. Chinn S. A simple method for converting an odds ratio to effect size for use in meta-analysis. Stat Med 2000;19:3127-31.
9. Helliwell PS, FitzGerald O, Fransen J, et al. The development of candidate composite disease activity and responder indices for psoriatic arthritis (GRACE project). Ann Rheum Dis 2013; 72:986-91.
10. Mulder MLM, Jones GT, Rotariu O, Helliwell PS. POS0980 Exploring the relationship between skin disease activity and the PASDAS in psoriatic arthritis [abstract]. Ann Rheum Dis 2024;83 Suppl 1:699.
11. 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.