Development of Composite Measures for Psoriatic Arthritis: A Report from the GRAPPA 2010 Annual Meeting

PHILIP S. HELLIWELL, OLIVER FITZGERALD, and PHILIP J. MEASE

ABSTRACT. Composite disease outcome measures have been used in rheumatology for some time, but a disease-specific composite measure for psoriatic arthritis (PsA) has not yet been validated. Currently, instruments developed for use in rheumatoid arthritis are employed in PsA and include the American College of Rheumatology response criteria (ACR20, 50, and 70) and the Disease Activity Score for 28 and 44 joints (DAS28 and DAS44); however, these instruments do not cover the full spectrum of psoriatic disease. A composite measure is one way of incorporating an assessment of all relevant clinical outcomes into one single measure. By definition, it incorporates several dimensions of disease status, often by combining these different domains into a single score, which in the case of PsA includes joints, skin, entheses, dactylitis, and axial disease. New indices that combine these diverse clinical manifestations of PsA are under development and, in some cases, in the validation phase. The Group for Research and Assessment of Psoriasis and Psoriatic Arthritis (GRAPPA) established the GRAPPA Composite Exercise (GRACE) project to compare existing and emerging composite measures and to develop a new index. At the GRAPPA 2010 meeting, initial results from this project were presented, and existing and new candidate measures were compared. (J Rheumatol 2012;39:398–403; doi:10.3899/jrheum.111233)

Key Indexing Terms:
PSORIATIC ARTHRITIS   PSORIASIS   OUTCOME MEASURES

Psoriatic arthritis (PsA) is a heterogeneous disorder affecting peripheral and axial joints as well as having other features such as dactylitis, enthesitis, and skin and nail disease. Although not all these clinical features may occur together at any one time, it is important to be able to assess them all in order to assess their influence on the patient and the response to treatment, which may not be consistent across features.

A composite measure is one way of assessing all relevant clinical outcomes in a single instrument. By definition, it incorporates several dimensions of disease status, often by combining these different domains into a single score. Such instruments are well established in rheumatoid arthritis (RA), and these have been adopted for use in clinical trials involving patients with PsA. Measures adopted from RA include the American College of Rheumatology (ACR) responder index[1] and the Disease Activity Score for 28 joints (DAS28)[2]. The ACR responder index measures improvement in tender and swollen joint counts plus improvement in at least 3 of the following 5 measures: acute-phase reactant, patient global assessment of disease activity by visual analog scale (VAS), physician global assessment of disease activity by VAS, pain by VAS, and physical function using the Health Assessment Questionnaire (HAQ). The ACR20, 50, and 70 scores refer to ≥ 20%/50%/70% improvements in these measures[1].

In PsA, the number of joints assessed optimally involves a 68-tender, 66-swollen joint count, which includes the distal interphalangeal joints of the fingers[3]. The DAS28 in RA includes 28-joint tender and swollen counts, patient global, and either erythrocyte sedimentation rate or C-reactive protein (CRP). Although the DAS28 has been shown to distinguish PsA patients treated with anti-tumor necrosis factor agents from those receiving placebo, it was noted that 25% of the patients would not have been included in this study because the primary joints involved were below the knees, which are not assessed as part of the DAS28[4].

A number of additional composite measures for assessing disease activity in PsA have been proposed. A composite measure for defining "minimal disease activity" (MDA) has been validated and includes assessments of joints, skin, entheses, and physical function[5]. The MDA criteria define a low disease state and can be used as a responder index in addition to a target for treatment interventions. Three other disease-specific measures have been suggested. First, an adaptation of the Disease Activity index for Reactive Arthritis (DAREA)[6] has been renamed the Disease Activity index for

From the University of Leeds, Leeds, United Kingdom; St. Vincent's University Hospital, Dublin, Ireland; and the Swedish Medical Center, University of Washington, Seattle, Washington, USA.

Supported by GRAPPA.

P.S. Helliwell, DM, PhD, FRCP, Senior Lecturer in Rheumatology, University of Leeds; O. FitzGerald, MD, FRCPI FRCP(UK), St. Vincent's University Hospital; P.J. Mease, MD, Director, Rheumatology Research, Swedish Medical Center and Clinical Professor, University of Washington.

Address correspondence to Dr. P.S. Helliwell, Section of Musculoskeletal Disease, University of Leeds, 2nd Floor, Chapel Allerton Hospital, Harehills Lane, Leeds, LS7 4SA, United Kingdom; E-mail: p.helliwell@leeds.ac.uk

PSoriatic Arthritis (DAPSA), developed from a clinical cohort[7] and validated using clinical trial data[8]. Second, a weighted articular responder index, the Psoriatic Arthritis Joint Activity Index (PsAJAI), has been developed from pooled data from random clinical trials of biologic agents in PsA[9,10]. A response, according to this measure, is defined as a 30% improvement in core measures, with weights of 2 given to tender and swollen joint counts, CRP, and physician global assessment of disease activity, and weights of 1 given to patient pain and global scores, and the Health Assessment Questionnaire. In both of these instruments, the analytic method used in their development led to factoring out skin disease, which was therefore recommended to be measured separately from the musculoskeletal components, as has been the case with the ACR and DAS scoring systems.

Third, a domain-based approach has been proposed with the development of a composite measure known as the Composite Psoriatic Disease Activity Index (CPDAI)[11]. In the CPDAI, disease involvement is assessed in up to 5 domains: peripheral joints, skin, entheses, dactylitis, and spinal manifestations. For each domain, instruments are used to assess both the extent of disease activity and the effect of involvement in that domain on patient function and health-related quality of life. Domains are scored 0–3, with empirical cutoffs for disease severity/activity proposed in each, largely based on the literature. Individual domain scores are summed to give an overall, composite score (range 0–15)[11]. This instrument has also been validated in a large clinical trial dataset [Psoriasis Randomized Etanercept Study in Subjects with Psoriatic Arthritis (PRESTA)][12]. In an open-label period of the PRESTA study where 2 dose regimes of etanercept were administered, the composite measure (CPDAI) was able to demonstrate the differential response to the 2 doses, whereas the DAPSA was unable to discern this difference.

The GRAPPA Exercise to Develop a Composite Measure for PsA
The process of developing a composite measure for psoriatic disease started at the 8th meeting of OMERACT (Outcome Measures in Rheumatology) and was further developed at the annual Group for Research and Assessment of Psoriasis and Psoriatic Arthritis (GRAPPA) conferences in Leeds 2008 and Stockholm 2009[13,14]. The methodological approach followed the one used in the development of the Ankylosing Spondylitis Disease Activity Score (ASDAS)[15] and that suggested by Fransen, et al[16]. Further details of the assessments and instruments used in the clinical record form are given in a previous publication[14]. In brief, we aimed for a sample size of 300 with baseline, 3, 6, and 12 month data. At each timepoint, the surrogate for disease activity was a change in disease-modifying medication. At the time of the GRAPPA meeting in December 2010, 471 baseline subject case report forms had been received from the GRAPPA Composite Exercise (GRACE) project, and data on 268 of these were available for

the 3-month visit. Table 1 gives the baseline characteristics of these subjects divided into 2 groups according to the disease activity construct.

Analysis of the baseline data using methodology similar to that used in the development of the ASDAS (transformation of variables, factor analysis, discriminant function analysis, and linear regression) revealed that just 3 variables explained over 90% of the variance in scores. This putative index is given as: PsA index = [0.539 × patient global (mm)] + [0.194 × patient skin (mm)] + [0.438 × physician global (mm)].

In the GRACE study, the patient global scores were identical to those proposed by Cauli, et al[17], which used the following questions:
• Global VAS: In all the ways in which your PSORIASIS and ARTHRITIS as a whole affects you, how would you rate the way you felt over the past week?
• Skin VAS: In all the ways your PSORIASIS affects you, how would you rate the way you felt over the past week?

No specific question was given for the physician global VAS; the respondent was simply asked to mark on a 100-mm line the global assessment from "not active at all" to "extremely active."

To establish the desirability function (DF) cutoffs for disease activity, disease states were derived from members of the GRAPPA group using an online survey technique. These cutoffs were used to derive a series of linear functions for each variable so that the variable was rescaled to give a value from 0 (completely unacceptable) to 1 (no better outcome possible). These individual functions were then combined into a single measure to give the arithmetic mean (AM), again with a score range of 0 to 1. Two separate composite scores were developed:
• AM_DF1: includes swollen joint count, tender joint count, patient skin VAS, patient joints VAS, patient global VAS, HAQ, and a quality of life measure, the PsAQoL[18]
• AM_DF2: the same as AM_DF1 with the addition of the psoriasis area and severity index (PASI).

The performance of these 3 new measures (PsA Index, AM_DF1, AM_DF2) was compared to that of existing composite measures and the results are given in Table 2. In this table, the new measures are compared to the DAS28, DAPSA, CPDAI, and a new version of the CPDAI in which the cutoffs for each domain have been revised to be consistent with those used in the derivation of the DF. Two sets of data are used to compare the measures: the z statistic of the Mann-Whitney U test and the area under the receiver operating curve. In both cases, the larger the figure the better the measure discriminates between those with "active" disease and those with stable disease. For both statistics, the largest figure is found for the PsA Index. Similar results were obtained for subgroup analysis when patients with oligoarthritis (< 5 swollen/tender joints) were selected. Selecting for severe skin involvement (PASI score > 10) gave statistics of smaller magnitude such that, in the case of the CPDAI new and DAPSA, the z statistic did not reach the usual level of significance (0.05).

Characteristic	Treatment Change, n = 158	No Treatment Change, n = 313
Female/male	71/86 (1 unknown)	130/181 (2 unknown)
Mean (SD) age, yrs	50.2 (13.2)	51.2 (13.2)
Mean (SD) duration psoriasis, yrs	17.5 (14.2)	18.6 (13.5)
Mean (SD) duration arthritis, yrs	8.3 (9.5)	10.3 (9.7)
Mean (SD) tender joint count	10.4 (12.1)	5.6 (8.6)
Mean (SD) swollen joint count	4.7 (6.9)	2.6 (5.9)
No. (%) with skin involvement	130 (82)	243 (78)
No. (%) with nail involvement	84 (53)	154 (49)
Mean (SD) PASI	5.3 (8.4)	3.6 (5.2)
Mean (SD) mNAPSI	7.5 (13.3)	6.5 (10.9)
No. (%) with dactylitis	35 (22)	60 (19)
No. (%) with enthesitis	90 (57)	140 (45)
No. (%) with axial involvement	32 (23)	71 (20)

mNAPSI: Modified Nail Psoriasis Severity Index; PASI: Psoriasis Area and Severity Index.

Table 2. Comparison of measures at baseline, based on decision to change treatment (GRACE study; patients who had treatment change at baseline).

	Baseline Scores, mean (SD)			
	Treatment Not Changed	Treatment Changed	Z (MWU)	Area (ROC)
PsA Index	35.5	58.6	8.95	0.796
AM1_DF	0.69	0.51	8.05	0.774
AM2_DF	0.70	0.54	8.00	0.776
CPDAI	4.64	6.76	7.08	0.740
CPDAI new	5.81	8.30	6.85	0.715
DAPSA	17.87	30.34	6.95	0.737
DAS28	2.58	3.50	6.68	0.729

AM1_DF: composite score based on desirability functions; AM2_DF: AM1_DF with addition of PASI (Psoriasis Area and Severity Index); Area (ROC): area under the receiver operating curve; CPDAI: Composite Psoriatic Disease Activity Index; CPDAI new: CPDAI with new cutoffs (see text); DAPSA: Disease Activity for Psoriatic Arthritis; DAS28: Disease Activity Score for 28 joints; Z (MWU): Mann-Whitney U test, and associated z statistic; PsA Index; putative new composite score for PsA.

Table 3 gives the GRACE data for those who had a treatment change at baseline and for whom 3-month followup data were available. The scores for each measure are given and compared using the t statistic. Two other statistics were calculated: Cohen's effect size[19] and the standardized response mean[20]. For both measures, the larger the score the better the measure is able to record a response to the intervention — they are measures of responsiveness. Once more the PsA Index performs well, as do the DF and the CPDAI.

Results of Breakout Group Discussions
Following presentation of the data, GRAPPA delegates split into a number of small groups charged with discussing the content and performance of the measures. Delegates were asked to consider the measures in terms of the OMERACT filter: truth, discrimination, and feasibility. At a plenary session, group representatives discussed these deliberations and the main issues are listed below:

• Cultural differences may be apparent with response measure. Should we do different analyses by continent?
• What would be the influence of fibromyalgia and other comorbidities on the VAS scores?
• With reference to the proposed PsA Index, it was felt that the measure lacked face validity. In addition, one group questioned the inter- and intrarater reliability of the physician VAS scores. Another group reported recent results indicating good intrarater reliability for patient VAS scores, with an ICC of 0.87 for global disease activity[17].
• An assessment of enthesitis should be included in the DF scores
• Will the composite measures truly measure the full spectrum of disease?
• In the development, the measures should be tested for their prognostic ability
• The concept of a unidimensional instrument for a multidimensional disease was questioned. In particular, the concep-

Mean Scores	Baseline	3 mo	Change, Baseline — 3mo	t	Effect Size	SRM
PsA Index	58.6	39.0	16.1	4.44	0.83	0.57
AM1_DF	0.51	0.64	–0.12	4.40	0.67	0.67
AM2_DF	0.54	0.66	–0.11	4.76	0.67	0.69
CPDAI	6.76	5.40	1.53	3.87	0.50	0.62
CPDAI new	8.30	7.00	1.03	3.98	0.38	0.50
DAPSA	30.3	21.2	11.24	3.54	0.36	0.65
DAS28	3.5	2.9	0.72	2.74	0.45	0.54

AM1_DF: composite score based on desirability functions; AM2_DF: AM1_DF with addition of PASI (Psoriasis Area and Severity Index); CPDAI: Composite Psoriatic Disease Activity Index; CPDAI new: CPDAI with new cutoffs (see text); DAPSA: Disease Activity for Psoriatic Arthritis; DAS28: Disease Activity Score for 28 joints; PsA Index; putative new composite score for PsA; SRM: standardized response mean; t: t statistic to compare means.

tual model of combining skin and joint measures into one measure was thought to lack validity. It was suggested that one solution to this issue might be to express the measures within a 2-dimensional framework, with skin on one axis and musculoskeletal (possibly a composite of joints, spine, dactylitis, and enthesitis) on the other

• The measure should be useful in the clinic and in the research setting

• The skin component needs to include all the different ways and sites that can be affected by psoriasis, such as the scalp and genital areas

• Much discussion was devoted to the physician global VAS. Many people felt that this was an unreliable estimate of the true state of the patient's disease as assessed by the physician. Others countered that to make the judgment about global VAS, the physician must have examined the joints and the skin, so that it was equivalent to performing the objective measures

• In order to measure the possible differential response of the different domains to different treatments, the composite measure must have representation of individual domains nested within it. Otherwise, valuable information would potentially be lost

• To measure low disease activity states and remission, a composite measure must include all relevant domains of disease, and scores using this measure must reflect this, such that a low score cannot be obtained if any one domain still demonstrates significant disease activity

• Some discussants suggested that a comorbidity component be added to the composite measure, e.g., a measure of risk factors for cardiovascular disease.

Some discussants argued that it is conceptually flawed to attempt to combine assessments of skin and joints into one measure. Following the suggestion of one of the breakout groups, a new concept for recording disease activity in 2 dimensions was developed (Figure 1). Disease activity in skin is recorded along one axis and disease activity in the musculoskeletal system on the other. Whatever measure of disease
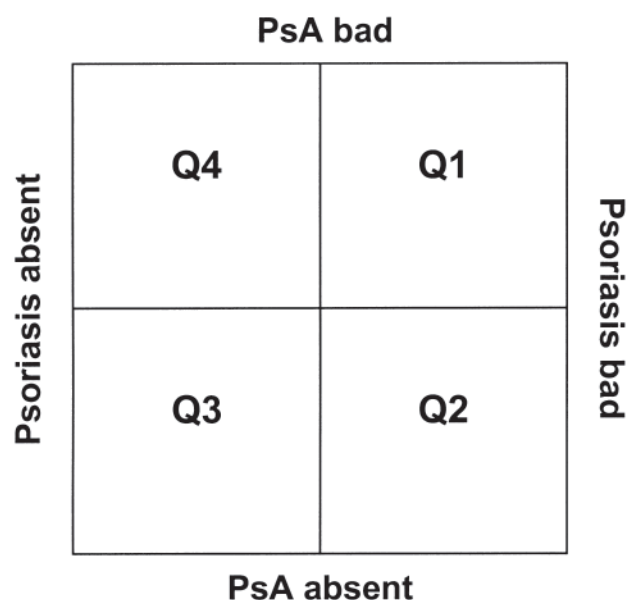

Figure 1. A proposed method for combining disease activity data for skin and musculoskeletal systems. For a particular patient, disease activity is represented by a vector. Changes in activity from visit to visit can be tracked by vector mathematics, but the observer can see at a glance the activity in each component of the disease. PsA: psoriatic arthritis; PSO: psoriasis; Q: quadrant.

activity is used, a division is made such that high and not-high disease activity can be defined. The resulting 4-quadrant graph allows the observer to identify the disease activity of the patient at a glance, and this can be quantified using vector mathematics.

Voting Procedures
At the end of the plenary sessions, delegates were asked to vote on each of the measures tested; the questions and results are provided in Table 4. The CPDAI ranked highest in the domains of truth and discrimination, and the proposed PsA Index in the feasibility domain. The attendees voted the fol-

Table 4. Voting responses from the attendees. The figures denote the percentage of positive ("yes") responses to each question.

| Measure | Truth | Does this measure fulfill the following requirements of the OMERACT filter? | | Should this measure be considered for further study? |
		Discrimination	Feasibility	
PsA Index	47	64	86	73
AM_DF1	67	40	71	25
AM_DF2	74	82	62	70
CPDAI	94	88	69	88
CPDAI new	89	78	65	53
DAPSA	41	36	66	19
DAS28	37	27	73	28

AM1_DF: composite score based on desirability functions; AM2_DF: AM1_DF with addition of PASI (Psoriasis Area and Severity Index); CPDAI: Composite Psoriatic Disease Activity Index; CPDAI new: CPDAI with new cutoffs (see text); DAPSA: Disease Activity for Psoriatic Arthritis; DAS28: Disease Activity Score for 28 joints; OMERACT: Outcome Measures in Rheumatology; PsA Index; putative new composite score for PsA.

lowing measures to consider for further study: PsA Index, AM_DF2, and CPDAI.

Additional Remarks

The GRAPPA initiative to develop a new composite measure for PsA has been under way for 3 years and is reaching the point where candidate measures are being adopted for further study in new interventional studies. In addition, if the data are appropriate, these measures can be tested in existing databases, as described[14]. At this time, GRAPPA is not committed to any single measure. Development of a single composite measure that has linear properties will enable the determination of cutoffs for low, moderate, and high disease activity, as well as the magnitude of any change. A single index also permits an assessment of disease activity at a glance. However — and this point was highlighted in the plenary session — it is important to retain the information on the different domains of disease, if only to provide reassurance to the observer who has traditionally relied on these individual assessments to guide their therapeutic choices.

Conclusions

The search for a better composite measure for PsA is further advanced but requires more developmental and validation work. Incorporating clinical assessment into a single composite measure presents several challenges. Existing composite measures mostly assess the articular component of the disease, but to combine all manifestations of this heterogeneous disease may be conceptually incorrect. Once candidate measures are available, the performance of these measures against existing measures will be necessary before final acceptance by the rheumatology community.

REFERENCES

1. Felson DT, Anderson JJ, Boers M, Bombardier C, Furst D, Goldsmith C, et al. American College of Rheumatology. Preliminary definition of improvement in rheumatoid arthritis. Arthritis Rheum 1995;38:727-35.
2. Prevoo ML, van Gestel AM, van 't Hof MA, van Rijswijk MH, van de Putte LB, van Riel PL. Remission in a prospective study of patients with rheumatoid arthritis. American Rheumatism Association preliminary remission criteria in relation to the disease activity score. Br J Rheumatol 1996;35:1101-5.
3. Mease PJ, Kivitz AJ, Burch FX, Siegel EL, Cohen SB, Ory P, et al. Etanercept treatment of psoriatic arthritis: Safety, efficacy, and effect on disease progression. Arthritis Rheum 2004;50:2264-72.
4. Fransen J, Antoni C, Mease PJ, Uter W, Kavanaugh A, Kalden JR, et al. Performance of response criteria for assessing peripheral arthritis in patients with psoriatic arthritis: Analysis of data from randomised controlled trials of two tumour necrosis factor inhibitors. Ann Rheum Dis 2006;65:1373-8.
5. Coates LC, Fransen J, Helliwell PS. Defining minimal disease activity in psoriatic arthritis: A proposed objective target for treatment. Ann Rheum Dis 2010;69:48-53.
6. Eberl G, Studnicka-Benke A, Hitzelhammer H, Gschnait F, Smolen JS. Development of a disease activity index for the assessment of reactive arthritis (DAREA). Rheumatology 2000;39:148-55.
7. Nell-Duxneuner VP, Stamm TA, Machold KP, Pflugbeil S, Aletaha D, Smolen JS. Evaluation of the appropriateness of composite disease activity measures for assessment of psoriatic arthritis. Ann Rheum Dis 2010;69:546-9.
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;69:1441-7.
9. Gladman DD, Tom BD, Mease PJ, Farewell VT. Informing response criteria for psoriatic arthritis. I: Discrimination models based on data from 3 anti-tumor necrosis factor randomized studies. J Rheumatol 2010;37:1892-7.
10. Gladman DD, Tom BD, Mease PJ, Farewell VT. Informing response criteria for psoriatic arthritis (PsA). II: Further considerations and a proposal — the PsA joint activity index. J Rheumatol 2010;37:2559-65.
11. Mumtaz A, Gallagher P, Kirby B, Waxman R, Coates LC, Veale JD, et al. Development of a preliminary composite disease activity index in psoriatic arthritis. Ann Rheum Dis 2011;70:272-7.
12. FitzGerald O, Helliwell P, Mumtaz A, Coates L, Pedersen P, Molta CT. Application of composite disease activity scores in psoriatic arthritis to the PRESTA dataset [abstract]. Arthritis Rheum 2010;62 Suppl:s214.
13. Gladman DD, Landewe R, McHugh NJ, Fitzgerald O, Thaci D,

Coates L, et al. Composite measures in psoriatic arthritis: GRAPPA 2008. J Rheumatol 2010;37:453-61.

14. Helliwell PS, Fitzgerald O, Strand CV, Mease PJ. Composite measures in psoriatic arthritis: A report from the GRAPPA 2009 annual meeting. J Rheumatol 2011;38:540-5.

15. van der Heijde D, Lie E, Kvien TK, Sieper J, Van den Bosch F, Listing J, et al. ASDAS, a highly discriminatory ASAS-endorsed disease activity score in patients with ankylosing spondylitis. Ann Rheum Dis 2009;68:1811-8.

16. Fransen J, Kavanaugh A, Borm G. Desirability scores for assessing multiple outcomes in systemic rheumatic diseases. Communications in Statistics – Theory and Methods 2009;38:3461-71.

17. 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;38:898-903.

18. McKenna SP, Doward LC, Whalley D, Tennant A, Emery P, Veale DJ. Development of the PsAQoL: A quality of life instrument specific to psoriatic arthritis. Ann Rheum Dis 2004;63:162-9.

19. Kazis LE, Anderson JJ, Meenan RF. Effect sizes for interpreting changes in health status. Med Care 1989;27 Suppl:S178-89.

20. Liang MH, Fossel AH, Larson MG. Comparisons of five health status instruments for orthopedic evaluation. Med Care 1990;28:632-42.