PsAID12 provisionally endorsed at OMERACT2018 as core outcome measure to assess Psoriatic Arthritis-specific Health-Related Quality of Life in clinical trials

Ana-Maria Orbai, Richard Holland, Ying Ying Leung, William Tillett, Niti Goel, Robin Christensen, Neil McHugh, Laure Gossec, Maarten de Wit, Pil Højgaard, Laura C. Coates, Philip Mease, Julie Birt, Lara Fallon, Oliver FitzGerald, Alexis Ogdie, Beverly Shea, Vibeke Strand, Kristina Callis-Duffin, Peter Tugwell, Dorcas Beaton, Dafna D. Gladman
University, Palo Alto, CA, USA, vibekestr@me.com; K Callis Duffin, MD, Department of Dermatology, University of Utah, Salt Lake City, Utah, USA, Kristina.duffin@hsc.utah.edu; P Tugwell, O.C., MD, FRCPC, Division of Rheumatology, Department of Medicine, and School of Epidemiology and Public Health, Faculty of Medicine, University of Ottawa, Clinical Epidemiology Program, Ottawa Hospital Research Institute, Ottawa, Canada, Tugwell.BB@uOttawa.ca; D Beaton, BScOT, PhD, Musculoskeletal Health and Outcomes Research, St. Michael's Hospital and Institute for Work and Health, and Department of Occupational Science and Occupational Therapy, Rehabilitation Sciences Institute, and the Institute for Health Policy Management and Evaluation, University of Toronto, Toronto, Ontario, Canada, BeatonD@smh.ca; DD Gladman, MD, FRCPC, Professor of Medicine, University of Toronto; Senior Scientist, Krembil Research Institute; Director, Psoriatic Arthritis Program, University Health Network; Toronto, Ontario, Canada, dafna.gladman@utoronto.ca;

Corresponding Author: Ana-Maria Orbai, MD MHS, Assistant Professor of Medicine, Rheumatology, Director Psoriatic Arthritis Program, Johns Hopkins University School of Medicine, Division of Rheumatology, 5200 Eastern Avenue MFL Center Tower Suite 4100, Baltimore, MD 21224, Email: aorbai1@jhmi.edu, Phone: +1 (410) 550-8089; Fax: +1 (410) 550-2072
Abstract

Objective: The GRAPPA-OMERACT psoriatic arthritis (PsA) working group is developing a Core Outcome Measurement Set for PsA clinical trials (randomized controlled trials and longitudinal observational studies) using the OMERACT Filter 2.1 instrument selection algorithm. Our objective was to assess the Psoriatic Arthritis Impact of Disease questionnaire (PsAID12) for the measurement of the core domain PsA-specific Health Related Quality of Life (HRQOL).

Methods: PsAID12 measurement property evidence gathered in a systematic literature review, and additional analyses conducted in longitudinal observational studies (LOS) were used to inform a stakeholder consensus process. Analyses that had not been published were independently reviewed by the OMERACT technical advisory group. Data and process were presented, discussed in breakout groups, and voted on, at the OMERACT conference (Terrigal, Australia, May 2018).

Results: PsAID12 fulfilled with green (good to go) OMERACT standards for domain match, feasibility, reliability, and construct/longitudinal construct validity. Discrimination and thresholds of meaning were amber (caution but good enough to
go forward). The overall working group recommendation was amber/provisional endorsement of PsAID12 for measuring PsA specific-HRQOL in randomized controlled trials (RCTs) and LOS. Of 96 participants who voted at the PsA OMERACT workshop 87.5% (84) voted “yes” endorsing this recommendation; 14 were patient research partners and 93% (13) voted “yes”; 82 participants were not PRPs and 87% (71) voted “yes”.

**Conclusion:** At OMERACT 2018, PsAID12 was the first patient reported outcome measure provisionally endorsed as core outcome measure for disease-specific HRQOL in PsA clinical trials. PsAID12 discrimination and improvement thresholds will be studied in future RCTs.
Introduction

Group for Research and Assessment of Psoriasis and Psoriatic Arthritis (GRAPPA) and Outcome Measures in Rheumatology (OMERACT) are collaborating to develop a Core Outcome Measurement Set for Psoriatic Arthritis (PsA). Aiming to standardize outcome measures for PsA randomized controlled trials (RCTs) and longitudinal observational studies (LOS) the project is based on the GRAPPA-OMERACT updated PsA core domain set (1,2) and new OMERACT guidance on instrument selection (3,4).

The OMERACT Filter 2.1 instrument selection process (3,4) requires a solid definition of the target domain as documented in the Core Domain Set. It consists of two major steps: 1) confirmation of content validity (domain match) and feasibility for the domain and population of interest; and 2) appraisal of construct validity and discrimination (reliability, responsiveness, thresholds of meaning). The evidence is appraised at each step and scored: Green (good to go), Amber (caution but good enough to go forward), and Red (stop, do not continue) (3). For any candidate outcome measure, the OMERACT selection process can be completed only if content validity and feasibility have initially been demonstrated (3,4).

HRQOL is a PsA core domain (1) but there are few outcome measures with dedicated content validity studies to support PsA-specific HRQOL domain match (6). This paper reports the first patient reported outcome (PRO), the Psoriatic Arthritis Impact of Disease (PsAID12) (7,8), appraised through OMERACT filter 2.1 - from core domain match with PsA-specific HRQOL, to provisional endorsement as core outcome measure for PsA clinical trials (randomized controlled trials and longitudinal observational studies).
Methods

The GRAPPA-OMERACT project is known by the acronym COMPACT (Core Outcome Measures for Psoriatic Arthritis Clinical Trials) (5). Figure 1 illustrates the COMPACT study design. It was applied to PsAID12, the first PRO evaluated through the OMERACT Filter 2.1. Through steps of evidence review and multi-stakeholder consensus guided by Filter 2.1, a conclusion can be achieved on suitability of a candidate as core outcome measure for the desired core domain. To arrive at a complete PsA core outcome measurement set, each PsA core domain has to have at least one corresponding outcome measure which successfully met all requirements of the OMERACT Filter 2.1 Instrument Selection criteria.

<Insert Figure 1>

The working group (6) completed a systematic literature review (SLR) of the evidence for measurement properties of all PRO used in PsA. PsAID12 had evidence for adequate validity and reliability for the core domain PsA-specific HRQOL. At the time of the SLR, PsAID9, the 9 item version of PsAID12, did not have enough evidence to be considered on its own. More importantly, PRPs had indicated that PsAID12 met the first step of truth/domain match better than PsAID9. The SLR was updated with a focus on PsAID12 to support the ensuing multi-stakeholder consensus process (9).

**The PsAID12 outcome measure:** The PsAID12 questionnaire (7, 8) was developed with PRPs included in domain generation, scoring and item formulation (6). PsAID12 is available free of charge for any application in several languages.
PsAID12 is composed of 12 numeric rating scales (NRS, range 0-11): 1) pain, 2) fatigue, 3) skin problems, 4) work and/or leisure activities, 5) functional capacity, 6) discomfort, 7) sleep disturbance, 8) coping, 9) anxiety, fear and uncertainty, 10) embarrassment and/or shame, 11) social participation, 12) depression. The total score is the sum of weighted NRS scores. The NRS weights represent patient rankings of importance for each of the 12 items. The weighted raw score is divided by 20 to yield a final PsAID12 score (range 0-10). Higher scores represent worse life impact. The individual NRS’ of the PsAID12 have been evaluated in the Bath, UK cohort (10).

**Domain match and feasibility:** Stakeholders were involved in three stages. First GRAPPA PRPs participated in a face-to-face pre-meeting (GRAPPA meeting July 2017). The PsA-specific HRQOL domain description was presented followed by completion by the PRPs of OMERACT truth and feasibility questionnaires for PsAID12 (3). Second, during the GRAPPA 2017 annual meeting COMPACT workshop, one breakout group was dedicated to PsAID12. Facilitators provided the HRQOL domain description, the PsAID12 questionnaire and scoring method; participants then anonymously completed each item of the OMERACT domain match and feasibility questionnaires (11). Third, the working group voted on domain match and feasibility.

**Construct validity and discrimination:** SLR PsAID12 articles were screened and selected (3) and the evidence for measurement properties was extracted (9) and placed in the Summary of Measurement Properties (SOMP) table (See Figure 2) to record location, quality and results of these studies. Two reviewers checked if the
study had used good enough methods to avoid a risk of bias in its results. The COSMIN-OMERACT Good-Methods-Checklist was used and consensus sought between two reviewers (AMO, PH, RH, WT). Evidence from studies using good methods was extracted and compared to measurement standards and results tracked on SOMP. Evidence from new studies was used to fill gaps and its quality was independently appraised by OMERACT’s technical advisory group (10,12). Final rating of synthesis of evidence for each measurement property was done according to OMERACT guidance looking for consistent, good performance from studies identified as having good methods for a green rating. Amber was assigned when a non-critical limitation in the evidence was found. Amber ratings could go forward if accompanied by a research plan. These ratings were recorded on the SOMP and an overall rating for the instrument given.

Results including the SOMP were sent out to the OMERACT community in pre-reading materials. At OMERACT 2018 the evidence was presented in a plenary session, and discussed in eight breakout groups. After breakout group results were reported, a plenary vote was held with all OMERACT attendees.

Results

SLR results and additional analyses are reported separately (9). Briefly seven studies met inclusion criteria and were used as evidence: Domain match (n=3), Feasibility (n=2), Construct validity (n=5), Test retest reliability (n=2), Longitudinal construct validity (n=2), Thresholds of meaning (n=2).

**Domain Match and Feasibility:** Twelve GRAPPA PRPs participated in the face-to-face meeting: nine voted green for PsAID12 domain match with the core domain
PsA- specific HRQOL, and three voted amber. There was unanimous green vote for PsAID12 feasibility. At the GRAPPA 2017 annual meeting 24 stakeholders voted majority amber for PsAID12 domain match, and majority green for feasibility (11). Twenty-two working group members voted: 17 (82%) green for domain match and 20 (91%) green for feasibility; the rest of the working group votes were amber.

**Construct validity and Discrimination:** Of five studies assessing construct validity two had good and three had amber methods. PsAID12 construct validity was adequate or better based on the studies’ hypotheses. We conducted our own analysis for construct validity in an international longitudinal dataset (12). Hypotheses were confirmed for correlations with similar/dissimilar constructs, and for known groups where we expected better HRQoL/PsAID12 scores in PsA subgroups with low disease activity versus high disease activity. This analysis constituted our 6th piece of evidence for construct validity with quality recorded as green.

There were two good quality studies for test-retest reliability with consistent adequate or better intra-class correlation coefficients in stable PsA patients (ICC 0.91-0.95). For longitudinal construct validity there were two good quality studies which replicated their hypotheses regarding longitudinal magnitude of change in patients who changed treatment compared with patients who were stable over time; the two studies had consistent results of good longitudinal construct validity.

For Discrimination in RCTs, because no RCTs had included PsAID12 at the time of this work, we conducted a separate analysis in a LOS, considered bronze level evidence based on OMERACT guidance (3). In the LOS a subset of patients had improved after medication change and could be compared to a subset of patients in the same study that had been stable with no change in therapy (10). We pre-
specified that PsAID12 effect size would be large in the improved group and small in
the stable group. The hypothesis was met and the level of evidence for
Discrimination in RCTs recorded as amber. For green level of evidence the analysis
should have taken place in an RCT which was not available at that time, and be
consistent across at least two studies.

**Summary of PsAID12 measurement properties:** The PsAID12 SOMP is
presented in Figure 2. The combined rating of all the evidence and measurement
properties was supportive of an amber recommendation, or provisional endorsement
as core instrument along with a roadmap/research agenda to complete final
endorsement. The working group is committed to derive evidence on discrimination
in RCTs once these datasets become available, and for thresholds of meaning in an
ongoing longitudinal study.

<Insert Figure 2>

**OMERACT workshop, endorsement, and research agenda:** OMERACT PsA
workshop participants, including PRPs (two per group, two groups only had one)
provided feedback in eight breakout groups (Supplement Table 1). Suggestions
were made as follows: to evaluate double barreled NRS; to evaluate coping, and
embarrassment/shame NRS for their sensitivity to change; to evaluate if all PsAID12
NRS change in the same direction with an intervention; and importantly, to define
clinically meaningful changes at the person level.

At the OMERACT PsA workshop the working group formulated a recommendation
for amber/provisional endorsement of PsAID12 to measure PsA specific-HRQOL in
RCTs and LOS. Of 96 participants who voted at the PsA OMERACT workshop 87.5% (84) voted “yes” endorsing this recommendation; 14 were patient research partners and 93% (13) voted “yes”; 82 participants were not PRPs and 87% (71) voted “yes”. The working group set a research agenda to fully endorse PsAID12, which includes implementation in RCTs to examine discrimination, and validation of PsAID12 thresholds for score interpretation.

Core outcome measurement set uptake and implementation strategies were discussed and the following were recommended for PsAID12: dissemination of PsAID12 provisional endorsement as core outcome measure for PsA RCTs and LOS; implementation in RCTs; and SLRs of clinical trial outcomes to assess if core outcome measures are being used.

Discussion

The PsAID12 is the first PRO provisionally endorsed as core outcome measure for PsA RCTs and LOS, based on OMERACT Filter 2.1 criteria. PsAID12 had a solid foundation of evidence for its measurement properties, derived from seven studies conducted with good methods. In addition, truth/domain match received strong endorsement from the beginning from both PRPs and the working group; feasibility was highly endorsed by all stakeholders. There was evidence for consistent good measurement properties across studies, and no evidence for poor measurement properties, allowing us to conclude that it performed well as a measure of PsA-specific HRQoL in cross-sectional as well as longitudinal studies.

Our results support the OMERACT approach to sequential core set development: a disease specific core domain set, followed by the core outcome measurement set.
The rigorous process of instrument appraisal helps identify high quality instruments to capture the core domains for PsA patients; and combines identifying existing evidence with guiding the working group to create new evidence to fill any gaps. The involvement of multiple stakeholders in core set development is key to dissemination and implementation. Standardization of PsA measurement will create the evidence necessary for core domain inclusion in treatment recommendations.

HRQoL has been routinely measured in PsA RCTs using the generic health survey Medical Outcomes Study Short Form-36 (SF-36) (13) although its content validity for PsA has not been studied (6). Content validity evidence would be necessary to pass the first critical step of the OMERACT filter (3,4). Psoriatic Arthritis Quality of Life (PsAQoL) is a PsA-specific measure but the GRAPPA-OMERACT literature review and appraisal of measurement properties (6) showed that PsAID12 content validity was superior to PsAQoL for PsA-specific HRQoL (6). Feasibility is also limiting for copyrighted measures (SF-36, PsAQoL), especially for longitudinal observational studies.

In summary, PsAID12 received overwhelming support at OMERACT 2018, and was provisionally endorsed as core instrument for PsA clinical trials. The provisional designation has been chosen until PsAID12 discrimination in RCTs and clinically meaningful thresholds for score interpretation can be appraised (PsAID12 research agenda). Additional candidate outcome measures are entering the OMERACT process of domain match and appraisal. The end result will be a complete PsA Core Outcome Measurement Set where each core domain will have at least one recommended core instrument. In the interim, clinical trial designers can use the current findings to inform their outcome measure choice for PsA-specific HRQoL in clinical trials.
Acknowledgement:

Study collaborators: C.A. Lindsay, PharmD, Patient Research Partner employed by Amgen Inc., Thousand Oaks, California, USA; D. Veale, MD, Adjunct Professor, School of Medicine, St. Vincent’s University Hospital, University College Dublin, Dublin, Ireland.
References:


12. de Wit M, Kvien TK, Gossec L. Patient participation as an integral part of patient-reported outcomes development ensures the representation of the patient voice: a case study from the field of rheumatology. RMD Open 2015;1:e000129.


TABLES AND FIGURE LEGENDS

Figure 1. The OMERACT Filter 2.1 Instrument selection algorithm is represented at the left (OMERACT handbook). The COMPACT workflow through OMERACT Filter 2.1 is represented at the right. The four signaling questions must be addressed in OMERACT Filter 2.1 instrument selection (items 1-4) in the left part of the figure. Following the OMERACT process stakeholders provide input initially into domain match and feasibility, and subsequently into the approach and results of evaluation of additional measurement properties, as well as the final voting for inclusion as core outcome measure(s). For each domain and candidate outcome measure the OFISA Filter 2.1 is applied until a core outcome measure has been selected. The process is repeated for each core domain. Once each core domain has at least one corresponding core outcome measurement instrument the PsA Core Outcome Measurement Set is complete. Abbreviations: OMERACT Outcome Measures in Rheumatology, GRAPPA Group for Research and Assessment of Psoriasis and Psoriatic Arthritis, COMPACT Core Outcome Measures in Psoriatic Arthritis, LOS longitudinal observational studies, PsA psoriatic arthritis, SLR systematic literature review, RCT randomized controlled trials, COSMIN Consensus based standards for the assessment of health measurement instruments, TAG OMERACT Technical Advisory Group.
Figure 2. OMERACT Summary of evidence for measurement properties of PsAID12. Color designates quality of evidence: dark grey stands for green=good methods used, use this evidence; light grey stands for amber=some cautions but we will use this evidence. In the rating row, color designates overall evidence-based instrument rating for the core instrument set: dark grey stands for green= at least 2 pieces of evidence with good methods and consistent findings of adequate or better performance; light grey stands for amber=in between green and red; red= inadequate performance in at least 1 study that used good methods, white=no evidence. Arithmetic signs designate the performance of the instrument according to that study (for each measurement property studied): “+”adequate or better performance, “+/-”equivocal performance, “-”less than adequate/poor performance. Abbreviations: TAG OMERACT Technical advisory group.