Endorsement of the 66/68 joint count for the measurement of musculoskeletal disease activity: OMERACT 2018 Psoriatic Arthritis workshop report

Alí Duarte-García*[1,2], Ying Ying Leung* [3], Laura C. Coates[4], Dorcas Beaton[5], Robin Christensen[6], Ethan T. Craig[7], Maarten de Wit[8], Lihi Eder,[9] Lara Fallon[10], Oliver FitzGerald[11], Dafna D. Gladman[12], Niti Goel[13,14], Richard Holland[15], Chris Lindsay[16], Lara Maxwell[17], Philip Mease[18], Ana Maria Orbai[19], Bev Shea[20], Vibeke Strand[21], Douglas J Veale[22], William Tillett**[23], Alexis Ogdie **[7]

*Co-first author

**Co-senior author

[1] Division of Rheumatology, Department of Medicine, Mayo Clinic, Rochester, MN, USA

[2] Robert D. and Patricia E. Kern Center for the Science of Health Care Delivery, Mayo Clinic, Rochester, MN, USA

[3] Department of Rheumatology and Immunology, Singapore General Hospital, Singapore, Singapore

[4] University of Leeds, Leeds, UK, and University of Oxford, Oxford, UK

[5] 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, ON, Canada.

[6] Musculoskeletal Statistics Unit: The Parker Institute, Bispebjerg and Frederiksberg Hospital & Department of Rheumatology, Odense University Hospital, Denmark

[7] Division of Rheumatology, Department of Medicine, University of Pennsylvania, Philadelphia, PA, USA

[8] Department of Medical Humanities, Patient Research Partner, Amsterdam University Medical Centre, , Amsterdam, The Netherlands.

[9] Department of Medicine, University of Toronto, Women's College Hospital, Toronto, ON, Canada

[10] Pfizer Inc., Montreal, QC, Canada

[11] Department of Rheumatology, St Vincent's University Hospital and Conway Institute for Biomolecular Research, University College Dublin, Ireland.

[12] Department of Medicine, University of Toronto, Toronto Western Hospital, Toronto, ON, Canada

[13] Patient Research Partner; Division of Rheumatology, Duke University School of Medicine, Durham, NC; Kezar Life Sciences, South San Francisco, CA, USA

[14] Division of Rheumatology, Duke University School of Medicine, Durham, NC, USA

[15] Royal Prince Alfred Hospital Medical Centre, Sydney, Australia

[16] Patient Research Partner, employed by Amgen Inc, Thousand Oaks, CA, USA

[17] Cochrane Musculoskeletal Group, Ottawa Hospital Research Institute, Centre for Practice-Changing Research, Ottawa, ON, Canada

[18] Swedish-Providence-St. John's Health Systems and University of Washington, Seattle, WA, USA

[19] Division of Rheumatology, Johns Hopkins University School of Medicine, Baltimore, MD, USA

[20] Ottawa Hospital Research Institute, and School of Epidemiology and Public Health, University of Ottawa, ON, Canada

[21] Division of Immunology/Rheumatology, Stanford University School of Medicine, Palo Alto, CA, USA

[22] St. Vincent's University Hospital and University College Dublin, Dublin, Ireland

[23] Royal National Hospital for Rheumatic Diseases and the University of Bath, Bath, UK

Abstract word count: 250

Manuscript word count: 2822

Figures/Tables: 6

Key words: OMERACT, Psoriatic Arthritis, Outcomes measures, core set

Corresponding author:
Alexis Ogdie, MD
University of Pennsylvania
White Building Room 5023, 3400 Spruce St, Philadelphia, Pennsylvania 19104
United States
E-mail: alexis.ogdie@uphs.upenn.edu

ABSTRACT

Objective: The psoriatic arthritis (PsA) core domain set for randomized controlled trials (RCTs) and longitudinal observational studies (LOS) has recently been updated. The joint counts are central to the measurement of the peripheral arthritis component of the musculoskeletal (MSK) disease activity domain. We report the Outcome Measures in Rheumatology (OMERACT) 2018 meeting approaches to seek endorsement of the 66/68-swollen and tender joint count (SJC66/TJC68) for inclusion in the PsA Core Outcome Measurement Set.

Methods: Using the OMERACT Filter 2.1 Instrument Selection Process, the SJC66/TJC68 was assessed for (1) domain match, (2) feasibility, (3) numerical sense (construct validity), and (4) discrimination (test retest reliability, longitudinal construct validity, sensitivity in clinical trials and thresholds of meaning). A protocol was designed to assess the measurement properties of the SJC66/TJC68 joint count. The results were summarized in a "Summary of Measurement Properties Table" developed by OMERACT. OMERACT members discussed and voted on whether the strength of the evidence supported that the SJC66/TJC68 had passed the OMERACT Filter as an outcome measurement instrument for the PsA Core Outcome Measurement Set.

Results: OMERACT delegates endorsed the use of the SJC66/TJC68 for the measurement of the peripheral arthritis component of the MSK disease activity domain: Among patient research partners, 100% voted for a "green" endorsement, whereas among the group of "other stakeholders", 85% voted for a "green" endorsement.

Conclusion: The SJC66/TJC68 is the first fully endorsed outcome measurement instrument using the OMERACT Filter 2.1 and the first instrument fully endorsed within the PsA Core Outcome Measurement Set.

INTRODUCTION

Psoriatic arthritis (PsA) is a chronic inflammatory musculoskeletal and skin disease that is clinically heterogeneous with distinct manifestations including peripheral arthritis, spondylitis, enthesitis, and dactylitis, as well as skin and nail features. Additionally, the disease affects many domains of patients' lives including fatigue, participation, and emotional wellbeing. The Group for Research and Assessment of Psoriasis and Psoriatic Arthritis (GRAPPA)-Outcome Measures in Rheumatology (OMERACT) working group developed a Core Domain Set (**Figure 1**) to specify which key domains should be measured in randomized controlled trials (RCTs) and longitudinal observational studies (LOS) for PsA. This was endorsed at the 2016 OMERACT meeting.(1, 2) Since that time, multiple work streams have been initiated as part of the Core Outcome Measures for Psoriatic Clinical Trials (COMPACT) study.(3, 4) The GRAPPA-OMERACT working group has been evaluating the measurement properties of multiple outcome measurement instruments to develop a PsA Core Outcome Measurement Set (COS) that would assist in standardizing what is measured in RCT's and how they are measured (Domains and Instruments).(5, 6)

Among the domains included in the COS, musculoskeletal (MSK) disease activity is considered one of the most important for both patients and clinicians.(1) The MSK disease activity domain includes peripheral joints, enthesitis, dactylitis, and spine symptoms. The tender and swollen joint counts are central to the measurement of the peripheral arthritis element of MSK disease activity. While several joint counts exist,(7) there are no existing recommendations about which joint count to use in RCTs or LOS measuring peripheral arthritis in PsA, and none have moved through the Instrument Selection Process described by OMERACT.

The goal of the PsA workshop at OMERACT was to seek endorsement of the 66/68-Swollen and Tender Joint Counts (SJC66/TJC68) (**Figure 2**) as one of the instruments for the PsA Core Outcome Measurement Set. In this paper, we describe the instrument selection process as recommended by OMERACT, summarize the plenary presentation, and present the voting results and discussion points from the PsA workshop and breakout groups at the OMERACT 2018 meeting.

METHODS AND RESULTS

Patient Engagement in the Working Group

One of the key tenets of the OMERACT process is involving patient research partners (PRPs) in the process of developing core outcome sets. In the work presented in this paper, PRPs have been involved in all aspects of the project: Three PRPs are part of the GRAPPA-OMERACT working group steering committee. They have reviewed and provided feedback on protocols, pre-reading materials, and presentations, and helped plan the workshop. Furthermore, PRPs from GRAPPA and OMERACT have participated in small groups and were involved in surveys and web-based seminars.

Instrument Selection Process

Using the OMERACT Filter 2.1 Instrument Selection Process (**Figure 3**), an instrument is first assessed for "Truth: domain match" and "Feasibility" and, if these two steps are met, the instrument may progress to the subsequent steps, "Truth 2: Numerical Sense" (i.e., construct validity) and "Discrimination" (measured by test retest reliability, longitudinal construct validity, ability to distinguish between treatment and placebo groups in clinical trials and thresholds of meaning).(8, 9) To seek endorsement of an instrument, the working group assembles the evidence for the instrument, appraises it and provides an overall assessment of the instrument using a "Summary of Measurement Properties" or SOMP table. In the absence of evidence in the available literature, new studies may be performed by the working

group to fill the evidence void. The working group makes a recommendation for endorsement and the attendees then vote if they agree with this recommendation. At OMERACT, the voting groups are split into patient research partners and other stakeholders to ensure the patient voice is adequately represented. At least 70% agreement among voting attendees at the session from both stakeholder groups suggests consensus with the working group recommendation.(6) For a more in depth review of the instrument selection process, see the OMERACT handbook.(8) The research protocol was reviewed and approved by the Institutional Review Board (IRB) of the University of Pennsylvania (IRB PROTOCOL#: 829776) for the patient research partner surveys and webinars, the rest of the project components were deemed exempt from IRB review.

Evaluation of Joint Counts using the OMERACT process

A systematic literature search (SLR) was first performed to identify instruments that had been used to measure MSK disease activity which includes peripheral joint activity, enthesitis, dactylitis, and spine symptoms in PsA and to assess their measurement properties.(10, 11) In this report, we focused on the evidence evaluating the SJC28/TJC28, SJC66/TJC68, and SJC76/TJC78 joint counts. We addressed domain match and feasibility at the GRAPPA meeting in 2017 (Amsterdam) as well as with the working group and patient research partners (described in more detail below). We assessed the measurement properties of the joint counts in the literature (and applied the OMERACT Good Methods Checklist to assess data quality) and analyzed measurement properties in clinical trial and longitudinal observational study datasets (obtained from companies and principal investigators). The working group requested data from phase III trials published between 2010-2017 and was included from 7 phase III RCTs, The Tight Control of Inflammation in Psoriatic Arthritis trial and one LOS, the Psoriatic Arthritis Research Consortium. A priori, a standardized protocol was designed to address content validity, construct validity, responsiveness and discrimination.

We used this data to complete the summary of measurement properties table and presented this to the working group for a final recommendation. The results were then presented at the OMERACT meeting in Terrigal.

The PsA OMERACT Core Set Workshop at the GRAPPA 2017 Meeting

Domain Match and Feasibility of the Joint Counts: Clinicians and other Stakeholders

Domain match and feasibility for the SJC66/TJC68 were addressed at GRAPPA 2017 in a breakout group discussion and, following GRAPPA, among working group members using a web-based survey. During the GRAPPA meeting, content validity and feasibility were addressed within a small group with clinicians, two patients, and a patient advocate; the voting sheets were completed by 22 people.(12) There was consensus (20/22, 91%) among the group that the SJC66/TJC68 was a match for the MSK disease activity/peripheral arthritis domain and that there was adequacy of content and no redundancies. With regards to feasibility, all the voters agreed that the SJC66/TJC68 was feasible.

Eighteen working group members completed a follow up online survey. This survey documented the reasons for selecting SJC66/TJC68 count over the comparators (28 and 76/78 joint counts). The 28 joint count is a core measure for rheumatoid arthritis (RA) and is frequently performed in clinical practice. The 76/78 joint count is performed in some trials. Other joint counts beyond the 28, 66/68, and 76/78 (i.e., 32, 44, Ritchie index) were not sufficiently used in RCTs or LOS to merit inclusion.(7)

Concerns have been raised about these joint counts in PsA: the 28 joint count does not include the joints of the feet, and these joints are frequently affected in PsA, this concern was raised by both PRPs and clinicians; the 76/78 joint count includes the carpometacarpal (CMC) joints, typically involved in osteoarthritis and thus tenderness in this joint is difficult to attribute to PsA, and it separately includes the toe proximal and distal interphalangeal joints, these joints are difficult to decipher individually on exam decreasing feasibility.

The 28 joint count did not meet domain match (does not cover key joints) and the 76/78 joint count had lower feasibility (difficult to distinguish between toe joints) and reduced domain match (CMC joint more often an osteoarthritis joints) compared to the SJC66/TJC68. Given the results of the above discussions and surveys with all stakeholders, the working group decided to only move forward the SJC66/TJC68 through the OMERACT filter (**Figure 4**).

Domain Match and Feasibility of the Joint Counts: Patient Research Partners

To assess domain match and feasibility from the PRPs perspective, a web-based survey was a designed with an embedded video of a clinician (AO) performing the SJC66/TJC68. Respondents were asked to note whether they agreed that the SJC66/TJC68 measured their perception of "peripheral arthritis disease activity" and whether it was feasible to complete within RCT or LOS visits. PRP representatives of GRAPPA and OMERACT were invited to participate in the survey, and fourteen responded. Among those that responded, 9 voted "green" and 3 voted "amber" and 1 voted "white" . For feasibility, 13 voted "green" and 1 voted "white." After completion of the survey, two web-based seminars were held with the participating PRPs to discuss the results by first reporting the survey results and then the discussion was opened for comments, questions, or concerns. We began the webinar by discussing the survey questions and results. Points of confusion with the domain were that several patients did not endorse for "domain match" because the SJC66/TJC68 did not include the entheses or the spine. AO reminded the group that enthesitis and spine symptoms are assessed using separate measures, and this explanation was satisfactory to those who voted "no" (although the group did not re-vote as the vote was mainly used to start the discussion). Some patients advocated for inclusion of the CMC joint as a common source of pain. PRPs also noted that the feet and ankles are essential for inclusion in assessing peripheral arthritis in PsA.

Regarding feasibility, all patients felt that the SJC66/TJC68 is feasible however, the only concern raised was that when patients are in a lot of pain, getting shoes on and off is uncomfortable and can decrease feasibility. Additionally, the patient needs sufficient time to respond during examination (i.e., if the SJC66/TJC68 is performed too quickly, there will not be sufficient opportunity to say "yes" to a tender joint). It was also noted by several PRPs that for the joint count to be a valid assessment of peripheral arthritis, particularly tenderness, there needs to be communication between the physician and patient. There was discussion about the fact that the joint examination may miss a joint that was active within the past week but is not active today. Finally, patients felt that there was no clear meaning for "tender" and that communication from the physician prior to the joint count is assessed is needed.

Numerical Sense (construct validity) and Discrimination

We addressed numerical sense and discrimination via an SLR and analysis of RCT datasets. In the SLR, 1921 unique references regarding the four components of the MSK disease activity domain were identified, 159 were eligible for full-text article assessment and 87 of these were excluded in this phase. 59 of the 72 remaining were excluded since they involved other components of the MSK disease domain (e.g. dactylitis) not pertinent for this report or due to lack of enough data regarding the SJC66 and TJC68. 13 SJC66/TJC68 unique references were included in the good methods analysis. The good methods checklist is applied at the level of the instrument and measurement property tested rather than the level of the study, in our case, no study had some red and other evidence that was amber or green. Three studies had all their evidence as red and therefore were excluded, leaving 10 studies for inclusion (**Figure 5**).

The list of articles and summary of findings was included in **Table 1**. The results suggest that SJC66 and TJC68 have construct validity. TJC68 has adequate interrater reliability while SJC66 does not have adequate interrater reliability (ICC<0.75).(13) Regarding responsiveness and discrimination, SJC66 and

TJC68 change over time in response to treatment (placebo did change as well but less) and the change in SJC66 and TJC68 can distinguish between patients receiving an effective therapy compared to placebo. We similarly addressed measurement properties including responsiveness and discrimination in RCT datasets (manuscript in progress, data presented at the OMERACT meeting). Standardized response means ranged from -0.9 to -0.5 for the SJC66 and -0.9 to -0.4 for the TJC68, thus mostly in the moderate effect range. Standardized mean difference (treatment compared to placebo) range from -0.7 to -0.2 for the SJC66 and -0.6 to -0.2 for the TJC68.

The working group concluded that the SJC66/TJC68 meets the OMERACT criteria for domain match, feasibility, truth and discrimination. The instruments shortfalls are relatively low inter-rater reliability for the SJC only and a lack of studies addressing intra-rater reliability of the TJC/SJC in PsA (**Table 1**).

OMERACT 2018 PsA Workshop: Plenary Presentation and Breakout Group Discussions

In the plenary presentation, we presented the evidence that addressed each of the 4 steps of the OMERACT Filter 2.1 Instrument Selection Process for SJC66/TJC68. Data from these studies were summarized in the SOMP Table (**Table 1**).

After the plenary presentation, eight breakout groups were asked to discuss the four measurement properties (content validity/domain match, feasibility, construct validity, and discrimination) and vote on agreement with the working group's assessment of green ("good to go"), amber ("some concerns raised"), or red ("not endorsed"). Breakout groups were facilitated by one OMERACT-trained facilitator and one reporter; reporters were part of the working group or experienced researchers. During breakout groups, the participants had the option to raise concerns regarding the working group assessment of "green". Overall, most participants agreed with our assignment of "green" for content

validity/domain match, feasibility, and construct validity. Feasibility concerns came up for some groups in that the SJC66/TJC68 takes longer than the reduced joint counts, but overall, the majority felt that the SJC66/TJC68 is feasible in the setting of a RCT or LOS. In some groups, concerns were raised about discrimination, mainly centered around the insufficient data for test-retest reliability and thresholds of meaning (both with only one unpublished study available in PsA). Additionally, the concern about the relatively low interrater reliability of the SJC was raised. This was countered by the argument that in most RCTs, the assessor is the same throughout the study and test-retest reliability, or intrarater reliability, in a single unpublished study was found to be quite high (ICC 0.8-0.9) (Tillett et al unpublished). Furthermore, clinicians are generally asked to undergo training prior to trial participation to increase interrater reliability.(14) Reasons for endorsement of the SJC66/TJC68 that were raised included sending a clear message that this is the preferred joint count based upon evidence in order to assist in standardizing joint counts among RCTs.

A broader discussion was raised during the small groups regarding the meaning of full endorsement of an instrument ("green") or provisional endorsement ("amber). Some wondered whether a 'green' instrument would then become mandatory, similar to the inner circle of a core domain set. However, in the PsA workshop, "green" was used to denote the sufficient measurement properties to confidently say this is a "good" instrument, and "amber" was used to indicate that although this is a good instrument which could be used, further research is still required on its measurement properties. It is possible that multiple instruments for the same domain will pass through the filter at a "green" level, thus requiring a subsequent consensus process to identify the best instrument. Additional discussion then turned to define what is good enough? We assigned an "amber" to test-retest reliability and thresholds of meaning because of only one unpublished study for each. The OMERACT handbook suggests that the instrument should then be amber. However, the working group felt that the

instrument (SJC66/TJC68) should be endorsed as 'green' given the data in all other domains collectively being excellent; the studies evaluating test-retest reliability and thresholds of meaning being sufficient; and further research on these domains, though supportive, are not critical to further inform the preferred use of the SJC66/TJC68 over other joint counts.

Vote for the 66/68 Joint Counts

Following report back from the groups and discussion, we held a vote for the endorsement of the instrument. Among PRPs, 100% (of 14 patient votes in total) voted for a "green" endorsement. Among all other stakeholders, 88% (84 of 96 votes in total) voted for a "green" endorsement.

CONCLUSIONS

Through the years, the lack of standardization of the instruments to measure peripheral arthritis in PsA has resulted in the use of different instruments to assess peripheral arthritis in RCTs and LOS.

After a careful assessment by PRPs, clinicians, methodologists, representatives of the pharmaceutical industry and other stakeholders, and in accordance with the OMERACT filter 2.1, the evidence supporting the measurement properties of the SJC66/TJC68 was assessed and resulted in full endorsement ("green") by OMERACT as an instrument to measure MSK disease activity/peripheral arthritis in PsA. The SJC66/TJC68 is the first "green" instrument to enter the PsA Core Outcome Measurement Set.

The MSK disease activity domain includes the heterogeneous disease manifestations of PsA: enthesitis, dactylitis, spondylitis/axial arthritis and peripheral arthritis. An ongoing program will assess and eventually seek endorsement of the optimal instruments that measure the other components of the MSK disease activity domain. While the joint count is the first to go through the filter for this domain,

others will be added in the future as the additional work streams work through additional systematic literature reviews and consensus processes. In the meantime, PRPs, regulatory agencies, investigators developing protocols for RCTs and LOS, and other stakeholders can be confident with the SJC66/TJC68 and its adoption of the SJC66/TJC68 will be monitored in published RCTs and LOS.

ACKNOWLEDGEMENTS AND DECLARATIONS

Acknowledgements: We thank Kathleen Bush and Christina Burgese for administrative support. We would like to thank Janssen Scientific Affairs LLC for their assistance in identifying access to trial data through the YODA (Yale Open Data Access) Project. We would like to thank UCB, Novartis, and Pfizer for their scientific partnership by analyzing their clinical data of the certoluzimab RAPID-PsA, secukinumab FUTURE I & II, and tofacitinib PsA OPAL studies respectively to support the OMERACT-GRAPPA working group.

Ethical Approval: The SLR and physician surveys were deemed exempt from IRB approval. Trial participants in the original trials completed informed consent prior to participation. Patients who did not give consent for their data to be used for other studies were excluded from the additional trial analyses. Finally, IRB approval was obtained from the University of Pennsylvania for the patient research partner surveys and webinars.

Funding: Several parts of this study were funded by the Rheumatology Research Foundation Innovative Research Award (PI Ogdie) and some by R01- AR072363 (PI Ogdie). The Parker Institute, Bispebjerg and Frederiksberg Hospital (RC) is supported by a core grant from the Oak Foundation (OCAY-13-309). 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. Ali Duarte-Garcia is supported by the Mayo Clinic Robert D. and Patricia E. Kern Center for the Science of Health Care Delivery, which receives no industry funding

TABLE AND FIGURE LEGENDS

Table 1: Summary of Measurement Properties Table for TJC68/SJC66

The color indicates the good methods checklist/recommendation; the +/- sign indicates the adequacy of the data in support of the instrument.

For test-retest reliability, we have provided primary data for test-retest reliability (1 study). After identifying the gap in test-retest reliability in PsA, a hand search for test-retest reliability in RA was conducted to provide additional evidence and context about this property in the assessment of peripheral arthritis (2 studies).

Some studies were excluded because the measurement properties tested did not meet the good methods checklist: Gladman et al.1990 (15) (test-retest reliability, necessary summary statistics not provided), Englebrecht et al. 2010 (16) (construct validity, only tested correlation among subsets of TJC 68 and SJC66), Schoels et al. 2010 (17) (construct validity and longitudinal construct validity, didn't test associations of interest between joint counts and other individual measures).

Figure 1 Updated 2016 PsA Core Domain Set.

Final set of core domains to be measured in randomized clinical trials and longitudinal observational studies, endorsed by OMERACT in 2016.

*MSK disease activity includes peripheral joints, enthesitis, dactylitis, and spine symptoms

Abbreviations: MSK = musculoskeletal

Figure 2 The SJC66/TJC68 Joint Count

66 swollen and 68 tender joints are assessed (the hips are not assessed for swelling). The joint count is scored as a sum of the tender joints and a sum of the swollen joints.

Figure 3: OMERACT instrument selection process

Selecting an instrument the OMERACT way means proceeding through the steps shown in the figure: first defining the domain to be measured, identifying candidate instruments, assessing whether the instruments match the domain and feasibility, narrowing the list (removing instruments that do not match the domain), gathering evidence about measurement properties in a systematic literature search and data analysis, identifying the winners, and taking to OMERACT for endorsement. This figure was designed by Dorcas Beaton, OMERACT Handbook.

Figure 4 The OMERACT Filter 2.1 Instrument Selection Process for the of the 66/68-Swollen and Tender Joint Counts (SJC66/TJC68).

*Discrimination includes longitudinal construct validity, clinical trial discrimination (green circles), test-retest reliability and thresholds of meaning (amber circles).

Abbreviations: JC 28=28/28 Swollen and Tender Joint Counts; SJC76/TJC78=76/78-Swollen and Tender Joint Counts

Figure 5 PRISMA diagram

We conducted a systematic literature review for all four components of "MSK Disease Activity" which includes peripheral arthritis, enthesitis, dactylitis, and spine symptoms. In the current report, we focused on the TJC68 and SJC66 and thus many manuscripts went through the good methods checklist but only 10 papers included measurement properties that met these criteria and applied to the joint counts. In our case, no study had some red and other evidence that was amber or green. Three studies had all their evidence as red and therefore were excluded. Abbreviations: SLR=Systematic Literature Review; SJC66 = 66 Swollen Joint Count; TJC78 = 68 Tender Joint Count.

References

1. Orbai A-M, 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;76:673-80.
2. Orbai A-M, de Wit M, Mease PJ, Callis Duffin K, Elmamoun M, Tillett W, et al. Updating the psoriatic arthritis (psa) core domain set: A report from the psa workshop at omeract 2016. J Rheumatol 2017;44:1522-8.
3. Ogdie A, de Wit M, Callis Duffin K, Campbell W, Chau J, Coates LC, et al. Defining outcome measures for psoriatic arthritis: A report from the grappa-omeract working group. J Rheumatol 2017;44:697-700.
4. Tillett W, Orbai AM, Ogdie A, Leung YY, Strand V, Gladman DD, et al. Grappa-omeract initiative to standardise outcomes in psoriatic arthritis clinical trials and longitudinal observational studies. Ann Rheum Dis 2018;77:e23.
5. Prinsen CA, Vohra S, Rose MR, Boers M, Tugwell P, Clarke M, et al. How to select outcome measurement instruments for outcomes included in a "core outcome set" - a practical guideline. Trials 2016;17:449.
6. Beaton DE ML, Shea BJ, Wells GA, Boers M, Grosskleg S, et al. Instrument selection using the omeract filter 2.1: The omeract methodology. J Rheumatol 2018. (submitted)
7. Sokka T, Pincus T. Quantitative joint assessment in rheumatoid arthritis. Clin exp rheumatol 2005;23:S58.
8. Boers M, Kirwan JR, Tugwell P, Beaton D, Bingham CO, Conaghan PG, et al. Omeract handbook. Available from: https://omeract.org/resources.
9. Boers M, Kirwan JR, Wells G, Beaton D, Gossec L, d'Agostino MA, et al. Developing core outcome measurement sets for clinical trials: Omeract filter 2.0. J clin epidemiol 2014;67:745-53.
10. Duarte-Garcia A, Eder L, Stephens-Shields A, Goel N, de Witt M, Gladman D, et al. Construct validity of the swollen and tender joint counts for the measurement of msk disease activity in psoriatic arthritis[abstract]. Arthritis Rheumatol 2018;70 (suppl 10).
11. Duarte-Garcia A, Eder L, Goel N, Christensen R, de Wit M, FitzGerald O, et al. The 66/68 joint count for the measurement of msk disease activity/peripheral joint activity in psa: A grappa-omeract working group initiative [abstract]. Arthritis Rheumatol 2018;70 (suppl 10).
12. Holland R, Tillett W, Ogdie A, Leung YY, Gladman DD, Callis Duffin K, et al. Content and face validity and feasibility of 5 candidate instruments for psoriatic arthritis randomized controlled trials: The psa omeract core set workshop at the grappa 2017 annual meeting. J Rheumatol Suppl 2018;94:17-25.
13. Streiner DL, Norman GR, Cairney J. Health measurement scales: A practical guide to their development and use. Oxford University Press, USA; 2015.
14. Salvarani C, Girolomoni G, Di Lernia V, Gisondi P, Tripepi G, Egan CG, et al. Impact of training on concordance among rheumatologists and dermatologists in the assessment of patients with psoriasis and psoriatic arthritis. Semin Arthritis Rheum 2016;46:305-11.

15. Gladman DD, Farewell V, Buskila D, Goodman R, Hamilton L, Langevitz P, et al. Reliability of measurements of active and damaged joints in psoriatic arthritis. J Rheum 1990;17:62-4.
16. Englbrecht M, Wang Y, Ronneberger M, Manger B, Vastesaeger N, Veale DJ, et al. Measuring joint involvement in polyarticular psoriatic arthritis: An introduction of alternatives. Arthritis Care Res (Hoboken) 2010;62:977-83.
17. 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.

Author/Year	Truth Domain Match*	Feasibility*	Truth	Discrimination				
			Construct Validity	Test-Retest Reliability	Interrater Reliability	Longitudinal Construct Validity	Clinical Trial Discrimination	Thresholds of Meaning
Aalbers,2015(13)						+	+/-	
Chandran, 2009(14)					+(TJC)/-(SJC)			
Fransen, 2006(15)						+	+	
Gladman, 2004(16)					+(TJC)/-(SJC)			
Gladman, 2007(17)					+(TJC)/-(SJC)			
Husic, 2014(18)			+					
Leung, 2012(19)			+					
Lubrano, 2015(20)			+					
Tillet 2012(21)					-			
Salvarani, 2016(12)		+			+(TJC)/-(SJC)			
Duarte-Garcia 2018 (data analysis; unpublished)			+/-			+	+	
Tillet 2018 (unpublished)				+				+
GRAPPA and Working Group Surveys	+	+						
Patient input	+	+						
Thompson, 1991(22)				+				
Deandrade 1965(23)				+				
Total available studies for each property	2	3	4	3	5	3	3	1
Total studies available for synthesis	2	3	4	3	5	3	3	1
Rating (RAGW) [put on Master Checklist]	Green	Green	Green	Amber	Green for TJC Amber for SJC	Green	Green	Amber
Overall Rating for Instrument across properties	Green: Working Groups is Recommending Endorsement							

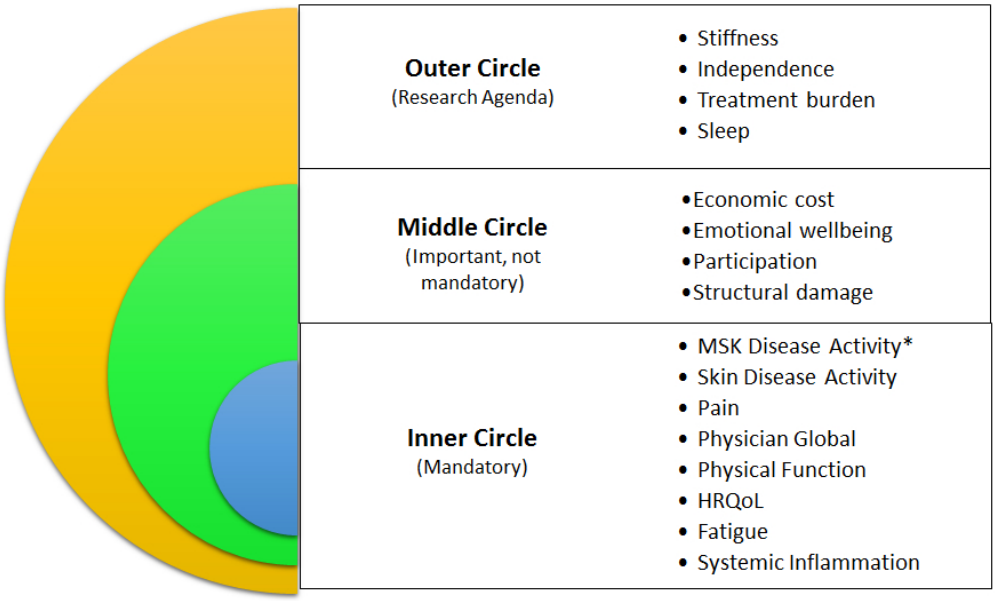

Figure 1 Updated 2016 PsA Core Domain Set. Final set of core domains to be measured in randomized clinical trials and longitudinal observational studies, endorsed by OMERACT in 2016.*MSK disease activity includes peripheral joints, enthesitis, dactylitis, and spine symptomsAbbreviations: MSK = musculoskeletal

308x185mm (72 x 72 DPI)

	Swollen joints	Tender joints
Temporomandibular joint	_____(0-2)	_____ (0-2)
Sternoclavicular joints	_____ (0-2)	_____ (0-2)
Acromioclavicular joints	_____ (0-2)	_____ (0-2)
Glenohumeral(s)	_____ (0-2)	_____ (0-2)
Elbow(s)	_____ (0-2)	_____ (0-2)
Wrist(s)	_____ (0-2)	_____ (0-2)
Metacarpal phalangeal joints	_____ (0-10)	_____ (0-10)
Finger Proximal interphalangeal joints	_____ (0-10)	_____ (0-10)
Finger Distal interphalangeal joints	_____ (0-8)	_____ (0-8)
Hip(s)	NA	_____ (0-2)
Knee(s)	_____ (0-2)	_____ (0-2)
Ankle(s)	_____ (0-2)	_____ (0-2)
Tarsus/Midfoot(feet)	_____ (0-2)	_____ (0-2)
Metatarsal phalangeal joints	_____ (0-10)	_____ (0-10)
Toe PIP(s)	_____ (0-10)	_____ (0-10)
Total joint counts	_____ (0-66)	_____ (0-68)

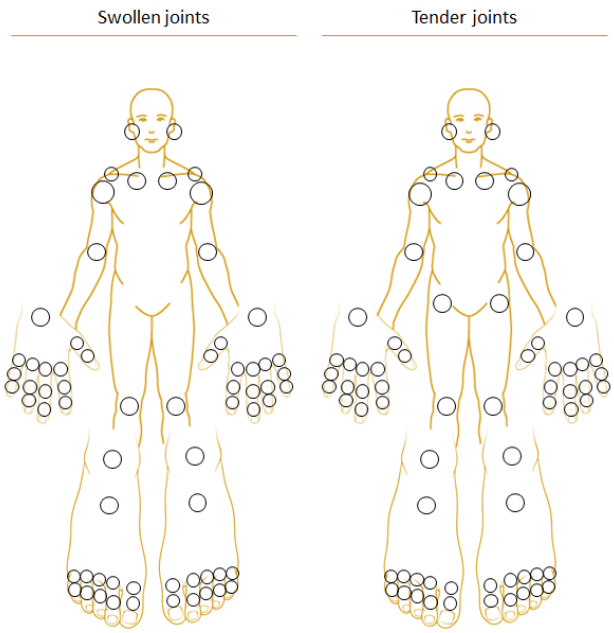

Figure 2 The SJC66/TJC68 Joint Count
66 swollen and 68 tender joints are assessed (the hips are not assessed for swelling). The joint count is scored as a sum of the tender joints and a sum of the swollen joints.

333x216mm (72 x 72 DPI)

How to choose an instrument the OMERACT way

Understand what you are trying to measure
Review domain definition and context of measurement

Find the candidate instruments
Then for each instrument...

Domain match?
Ask a lot of people, including patients; is this instrument capturing the chosen domain well?

Narrow the field
Set aside any red-flagged instruments. Short list the instruments.

Is it feasible to use?
Is it practical to use this instrument? Think about burden, time, cost, equipment.

Gather the evidence
Find, appraise and synthesize the measurement property evidence. Look for consistent findings of good performance from good quality studies.

Identify the winners
Those that have passed the Filter 2.1. Discuss work with Technical Advisory Group

bring it to a vote

brought to you by: Technical Advisory Group of
OMERACT Filter 2.1
Standards • Methods • Support

Figure 3. OMERACT instrument selection process: Selecting an instrument the OMERACT way means proceeding through the steps shown in the figure: first defining the domain to be measured, identifying candidate instruments, assessing whether the instruments match the domain and feasibility, narrowing the list (removing instruments that do not match the domain), gathering evidence about measurement properties in a systematic literature search and data analysis, identifying the winners, and taking to OMERACT for endorsement. This figure was designed by Dorcas Beaton, OMERACT Handbook.

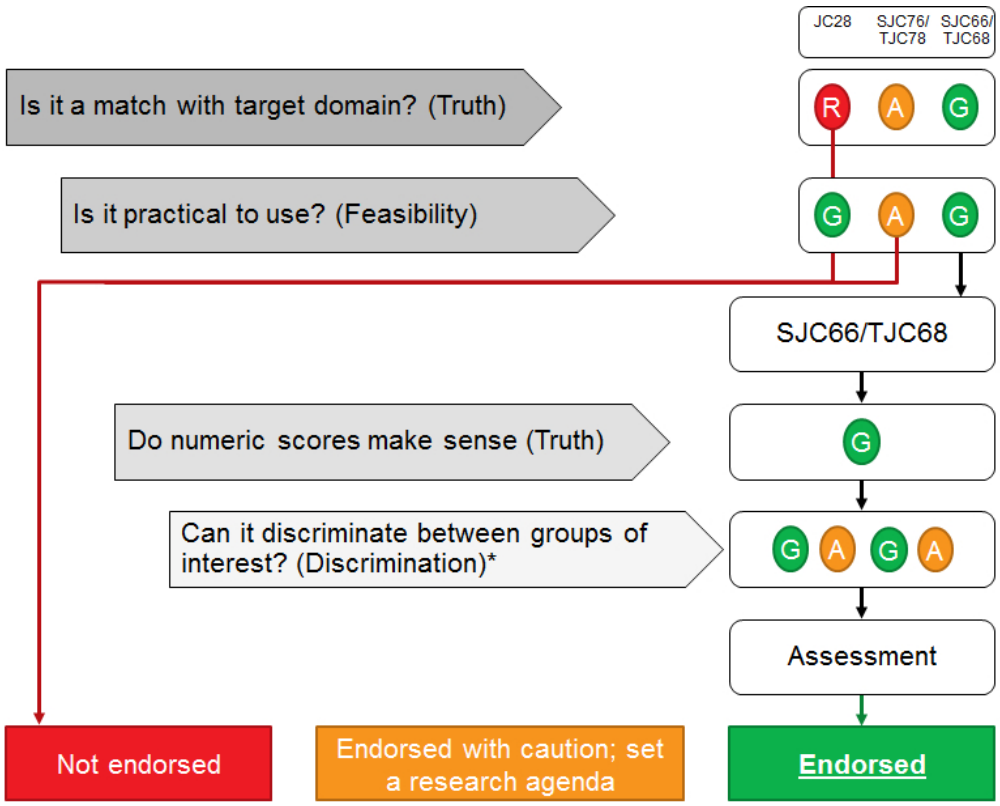

Figure 4 The OMERACT Filter 2.1 Instrument Selection Process for the of the 66/68-Swollen and Tender Joint Counts (SJC66/TJC68).
*Discrimination includes longitudinal construct validity, clinical trial discrimination (green circles), test-retest reliability and thresholds of meaning (amber circles).
Abbreviations: JC 28=28/28 Swollen and Tender Joint Counts; SJC76/TJC78=76/78-Swollen and Tender Joint Counts

270x217mm (72 x 72 DPI)

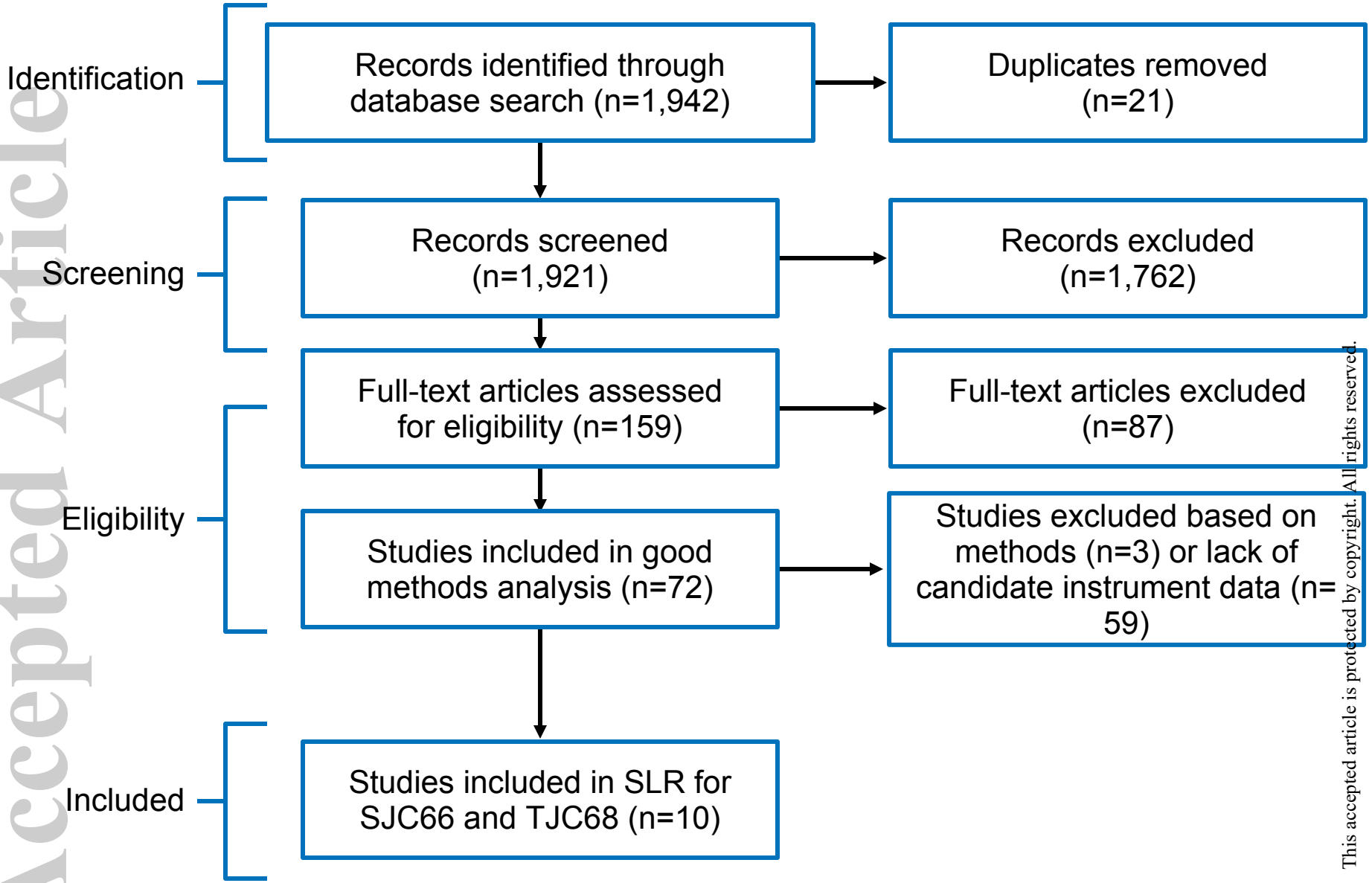

Figure 5 PRISMA diagram

We conducted a systematic literature review for all four components of "MSK Disease Activity" which includes peripheral arthritis, enthesitis, dactylitis, and spine symptoms. In the current report, we focused on the TJC68 and SJC66 and thus many manuscripts went through the good methods checklist but only 10 met criteria and applied to the joint counts. Abbreviations: SLR=Systematic Literature Review; SJC66 = 66 Swollen Joint Count; TJC78 = 68 Tender Joint Count.