Dialogue on Developing Consensus on Measurement and Presentation of Patient-important Outcomes, Using Pain Outcomes as an Exemplar, in Systematic Reviews: A Preconference Meeting at OMERACT 12

Peter S. Tugwell, Lara J. Maxwell, Dorcas E. Beaton, Jason W. Busse, Robin Christensen, Philip G. Conaghan, Lee S. Simon, Caroline Terwee, David Tovey, George A. Wells, and Paula Williamson

ABSTRACT.   Prior to the Outcome Measures in Rheumatology (OMERACT) 12 meeting in Budapest, Hungary, a workshop was held bringing together individuals from a number of international outcome measure organizations to assess how best to further develop consensus on how pain is conceptualized and measured in trials of musculoskeletal conditions, and how the trials should be reported in systematic reviews. (First Release Feb 1 2015; J Rheumatol 2015;42:1931–3; doi:10.3899/jrheum.141430)

Key Indexing Terms:
PAIN MEASUREMENT   CHRONIC PAIN   OUTCOMES RESEARCH   SYSTEMATIC REVIEW

From Department of Medicine, Faculty of Medicine, University of Ottawa; Ottawa Hospital Research Institute, Clinical Epidemiology Program, University of Ottawa; Department of Epidemiology and Community Medicine, Faculty of Medicine, Institute of Population Health, Ottawa; Institute for Work and Health, Toronto; Department of Anesthesia, Department of Clinical Epidemiology and Biostatistics, The Michael G. DeGroote Institute for Pain Research and Care, McMaster University, Hamilton, Canada; Musculoskeletal Statistics Unit, Department of Rheumatology, The Parker Institute, Bispebjerg and Frederiksberg Hospital, University of Copenhagen, Copenhagen, Denmark; Leeds Institute of Rheumatic and Musculoskeletal Medicine, University of Leeds and the UK National Institute for Health Research (NIHR) Leeds Musculoskeletal Biomedical Research Unit, Leeds, UK; SDG LLC, Cambridge, Massachusetts, USA; VU University Medical Center, Department of Epidemiology and Biostatistics, The Netherlands; Cochrane Editorial Unit, London, UK; Cardiovascular Research Methods Centre, University of Ottawa Heart Institute, Ottawa, Canada; Department of Biostatistics, University of Liverpool, Liverpool, UK.

P.S. Tugwell, MD, MSc, Department of Medicine, Faculty of Medicine, Ottawa Hospital Research Institute, Clinical Epidemiology Program, University of Ottawa, Department of Epidemiology and Community Medicine, Faculty of Medicine, Institute of Population Health; L.J. Maxwell, MSc, Institute of Population Health, University of Ottawa; D.E. Beaton, BScOT, PhD, Institute for Work and Health; J.W. Busse, DC, PhD, Department of Anesthesia, Department of Clinical Epidemiology and Biostatistics, The Michael G. DeGroote Institute for Pain Research and Care, McMaster University; R. Christensen, MSc, PhD, Musculoskeletal Statistics Unit, The Parker Institute, Department of Rheumatology, Bispebjerg and Frederiksberg Hospital, University of Copenhagen; P.G. Conaghan, MBBS PhD, FRACP, FRCP, Leeds Institute of Rheumatic and Musculoskeletal Medicine, University of Leeds and NIHR Leeds Musculoskeletal Biomedical Research Unit, Leeds, UK; L.S. Simon, MD, SDG LLC; C. Terwee, PhD, VU University Medical Center, Department of Epidemiology and Biostatistics; D. Tovey, MBChB, Cochrane Editorial Unit; G.A. Wells, MSc, PhD, Director, Cardiovascular Research Methods Centre, University of Ottawa Heart Institute; Professor, Department of Epidemiology and Community Medicine, University of Ottawa; P. Williamson, PhD, Professor, Department of Biostatistics, University of Liverpool.

Address correspondence to L. Maxwell, Centre for Practice-Changing Research, Ottawa Hospital Research Institute, General Campus, 501 Smyth Road, Office #L1227, Ottawa, Ontario K1H 8L6, Canada; E-mail: lmaxwell@uottawa.ca

Prior to the Outcome Measures in Rheumatology (OMERACT) 12 meeting in Budapest, Hungary, in May 2014, a workshop of 42 individuals was held to assess how best to move toward developing consensus on how pain is conceptualized and measured in trials of musculoskeletal (MSK) conditions, and how the trials should be reported in systematic reviews. The workshop included clinicians, patients, and researchers from 9 countries in the Americas, Australasia, and Europe, from 7 organizations representing the Cochrane Collaboration (6 Cochrane subgroups/entities), COMET (Core Outcome Measures in Effectiveness Trials), COSMIN (COnsensus-based Standards for the selection of health Measurement Instruments), GRADE (Grading of Recommendations Assessment, Development, and Evaluation), IMMPACT/ACTTION (Initiative on Methods, Measurement, and Pain Assessment in Clinical Trials/Analgesic, Anesthetic, and Addiction Clinical Trial Translations, Innovations, Opportunities, and Networks), OMERACT, and VAPAIN (Validation and Application of a core set of patient-relevant outcome domains to assess the effectiveness of multimodal pain therapy).

David Tovey, editor-in-chief of *The Cochrane Library*, noted that Cochrane Systematic Reviews aim to provide trustworthy and interpretable estimates of what works in health and healthcare. Identifying the proper research question using a formulation based on the P (Population), I (Intervention), C (Comparator), and O (Outcomes) framework is a critical first step. The aim of a review focused on treatment is to determine whether, for a given comparison and outcome, there is any effect/difference, the direction of any effect, and the degree of certainty that the calculated

effect estimate lies close to the true effect. To determine the latter, analysis based on the GRADE method and the Summary of Findings is crucial. One of the main characteristics of the GRADE approach is that the unit of analysis is the outcome (e.g., pain or function) — as opposed to a study-by-study approach favored by traditional measures. Another key element is to identify at the outset the main outcomes that will guide decision making, and the minimum important difference for each of these outcomes.

Pain is a commonly reported patient-important outcome in Cochrane Reviews[1], but even for the same clinical condition, reviews often report a whole range of different pain domains that are measured with different instruments and thresholds. This variability poses substantial problems for both the science of evidence synthesis, and the interpretability of any attempt to summarize the evidence to improve healthcare. A survey of outcomes reported in Cochrane reviews demonstrated the need for the development of standardized sets of outcomes to reduce variation and minimize outcome reporting bias[1a]. Those involved in either designing or summarizing MSK trials can make an important contribution if they can achieve consensus, based on methodologic rigor, on which measures of pain should be reported and on how to report them in a way that is most meaningful to patients, clinicians, and policymakers. In this issue of *The Journal*, a series of articles is presented with this aim in mind.

One article summarizes a Cochrane Library Systematic Review survey (covering diseases/conditions from Cochrane review groups for the back; MSK; and pain, palliative, and supportive care) where pain was an outcome; a broad range of conceptualizations of pain outcomes, instruments, and cutoffs have been reported[1]. The authors also report findings from an online survey, supplemented by qualitative interviews, of 36 individuals on key considerations for summary of findings tables (e.g., domains of pain, approaches to presenting results) when expressing the pain response of participants in chronic MSK pain intervention studies[1]. The individuals represented 14 groups (Cochrane Review groups: Back, MSK, Neuromuscular Disease; Pain, Palliative and Supportive Care; Applicability and Recommendations Methods; Patient Reported Outcomes Methods; Cochrane Editorial Unit; COMET; COSMIN; IMMPACT/ACTTION; VAPAIN; National Institutes of Health/National Library of Medicine; OMERACT Patients; OMERACT Researchers).

Four themes emerged as a basis for the working group discussions at the OMERACT 12 preconference meeting: pain domains, clinimetric considerations, thresholds for presenting results, and establishing hierarchies of outcomes.

The first thematic working group on Pain Domains addressed 2 major issues (reported in separate articles). First they tackled the controversy of whether chronic pain is a disease in its own right, as opposed to a "condition," "syndrome," or merely an extremely important symptom[2]. This is an ongoing debate among academic organizations studying pain but was new to many in the MSK community. The advantages and disadvantages are listed in Table 2 in the article. This issue will be taken up for further discussion and debate within the OMERACT Working Group on Pain.

The second article regarding the Pain Domains theme[3] reviewed the range of constructs that should be considered when assessing which aspects of pain (e.g., severity versus interference with activities) are important to patients; this issue of being explicit about such domains is one of the new features emphasized when assessing whether pain outcome assessment instruments meet the requirements of the OMERACT Filter 2.0[4]. The authors proposed a research agenda to establish a process to obtain consensus on standardizing outcome reporting of domains and subdomains of chronic MSK and rheumatologic pain, including a hierarchy of pain subdomains.

The second "clinimetrics" theme group addressed the challenge of achieving consensus on measurement property criteria for deciding whether a pain measurement instrument meets the OMERACT Filter 2.0 requirements and those of other approval agencies, such as the US Food and Drug Administration[5]. The flow chart of OMERACT Filter 2.0 was endorsed with a stepwise approach to selecting the concept first and considering the content validity of potential instruments for the intended context of use, then addressing practical and feasibility issues, and only then moving to a detailed assessment of the measurement properties of the instrument (methodological quality, and potential risk of bias) using an approach such as that developed by COSMIN. If this stepwise approach identifies gaps in the required evidence, additional studies should be proposed to ensure all necessary evidence is available. The research agenda brought forward by this group will complete the last phase of applying the OMERACT Filter 2.0 and provide core outcome developers with a template to ensure that major risk of bias is avoided and an evidence-based decision can be made on choice of instrument.

The third theme of "Thresholds for Presenting Results" was reviewed in the next article in the series[6]. Patient responder analysis, rather than mean results, was strongly recommended, in particular because of the finding, by the Pain, Palliative, and Supportive Care group, that patients in their systematic reviews tended to respond either by a relatively large amount or not at all, so that a mean change is not informative[7,8]. Provisional consensus was reached: Options for individual trials should include reporting of the proportion of patients achieving 1 or more thresholds of improvement from baseline pain (e.g., $\geq 20\%$, $\geq 30\%$, $\geq 50\%$), achievement of a desirable pain state (e.g., no worse than mild pain), and/or a combination of change and state. The research agenda includes (1) evaluating the proposal that when pooling data for metaanalysis, authors should

consider converting all continuous measures for pain to a 10 cm/100 mm visual analog scale (VAS) for pain and use the minimally important difference (MID) of 1 cm/10 mm, and the conventionally used appreciably important differences of 2 cm/20 mm, 3 cm/30 mm, and 5 cm/50 mm, to facilitate interpretation; (2) assessing the loss of discrimination tradeoff with ease of interpretation from the growing practice of combining 20% and 30% improvement from baseline in systematic reviews; (3) assessing whether there is consensus that effect sizes of ≤ 0.5 MID units suggest a small or very small effect, and effects ≥ 2.0 MID units suggest a large effect; (4) assessing whether increased interpretability is achieved by transforming the pooled estimate on the VAS/numerical rating scale to a binary outcome and expressed as a relative risk and risk difference.

A fourth theme, presented in the article by Christensen, *et al*, describes approaches to using a hierarchy for selecting different outcomes to combine in a metaanalysis[9]. Predefining such a hierarchy avoids selective outcome reporting bias in studies with more than 1 pain scale, where the temptation is to use the one ("cherry picking") with the largest effect size[10]. Use of a predefined hierarchy set, initiated by Juni, *et al*[11] and developed by the Cochrane Musculoskeletal Editors, is proving to be popular with the Cochrane Musculoskeletal systematic review authors[12]. Juhl, *et al* proposed using a combination of criteria of frequency of use and effect size[13]. There was agreement on a research agenda, including the need to develop methodology for generation of hierarchical lists of outcome instruments measuring pain to guide metaanalyses. Tools that could be used to steer development of such a prioritized list are the COSMIN checklist and the OMERACT Filter 2.0.

Regarding the topic of Pain Outcomes, Measurement, and Systematic Reviews, the pre-OMERACT meeting provided an important opportunity for representatives from a broad range of groups and organizations to discuss issues face-to-face while using the OMERACT process. Not all individuals participate in more than 1 organization. Different organizations have realized the importance of these issues and have addressed them to different degrees (done in isolation, this may be inefficient, not to mention the danger of competitive ownership). As well, there is too much to be done for any single organization. Our hope is that the participant organizations continue this partnership, while others who wish to participate will join in to address those parts of the research agenda that interest them.

REFERENCES

1. Maxwell LJ, Tugwell P, Wells GA, Simon LS, Conaghan PG, Grosskleg S, et al. Current state of reporting of pain outcomes in Cochrane reviews on chronic musculoskeletal pain conditions and considerations for an OMERACT research agenda. J Rheumatol 2015;42:1936-44.

1a. Smith V, Clarke M, Williamson P, Gargon E. Survey of new 2007 and 2011 Cochrane reviews found 37% of prespecified outcomes not reported. J Clin Epidemiol 2014 Nov 18 (E-pub ahead of print).

2. Taylor AM, Phillips K, Taylor JO, Singh JA, Conaghan PG, Choy EH, et al. Is chronic pain a disease in its own right? Discussions from a pre-OMERACT 2014 workshop on chronic pain. J Rheumatol 2015;42:1947-53.

3. Phillips K, Taylor AM, Mease P, Simon LS, Conaghan PG, Choy E, et al. Harmonizing pain outcome measures: Results of a nominal group technique on pain domains at the Pre-OMERACT meeting on Partnerships for Consensus on Patient-important Outcomes domains. J Rheumatol 2015;42: in press.

4. 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 2015;67:745-53.

5. Beaton DE, Terwee CB, Singh JA, Hawker GA, Patrick DL, Burke LB, et al. A call for evidence-based decision making when selecting outcome measurement instruments for summary of findings tables in systematic reviews: results from an OMERACT working group. J Rheumatol 2015;42:1954-61.

6. Busse JW, Bartlett S, Dougados M, Johnston B, Guyatt GH, Kirwan J. Optimal strategies for reporting pain in clinical trials and systematic reviews: Discussions at the Pre-OMERACT 2014 Meeting. J Rheumatol 2015;42: in press.

7. Moore RA, Straube S, Aldington D. Pain measures and cut-offs — 'no worse than mild pain' as a simple, universal outcome. Anaesthesia 2013;68:400-12.

8. Moore RA, Cai N, Skljarevski V, Tölle TR. Duloxetine use in chronic painful conditions - individual patient data responder analysis. Eur J Pain 2014;18:67-75.

9. Christensen R, Maxwell LJ, Juni P, Tovey D, Williamson PR, Boers M, et al. Consensus on the need for a hierarchical list of patient-reported pain outcomes for metaanalyses of knee osteoarthritis trials: Discussions at the Pre-OMERACT 2014 Meeting. J Rheumatol 2015;42: in press.

10. Hutton JL, Williamson PR. Bias in metaanalysis due to outcome variable selection within studies. Appl Stat 2000;49:359-70.

11. Juni P, Reichenbach S, Dieppe P. Osteoarthritis: rational approach to treating the individual. Best Pract Res Clin Rheumatol 2006;20:721–40.

12. Ghogomu E, Maxwell LJ, Buchbinder R, Rader T, Pardo Pardo J, Johnston RV, et al. Updated method guidelines for Cochrane musculoskeletal group systematic reviews and metaanalyses. J Rheumatol 2014;41:194-205.

13. Juhl C, Lund H, Roos EM, Zhang W, Christensen R. A hierarchy of patient-reported outcomes for meta-analysis of knee osteoarthritis trials: empirical evidence from a survey of high impact journals. Arthritis 2012;2012:136245.