Consensus on the Need for a Hierarchical List of Patient-reported Pain Outcomes for Metaanalyses of Knee Osteoarthritis Trials: An OMERACT Objective

Robin Christensen, Lara J. Maxwell, Peter Jüni, David Tovey, Paula R. Williamson, Maarten Boers, Niti Goel, Rachelle Buchbinder, Lyn March, Caroline B. Terwee, Jasvinder A. Singh, and Peter Tugwell

ABSTRACT. Objective. Although protocol registration for systematic reviews is still not mandatory, reviewers should be strongly encouraged to register the protocol to identify the methodological approach, including all outcomes of interest. This will minimize the likelihood of biased decisions in reviews, such as selective outcome reporting. A group of international experts convened to address issues regarding the need to develop hierarchical lists of outcome measurement instruments for a particular outcome for metaanalyses.

Methods. Multiple outcome measurement instruments exist to measure the same outcome. Metaanalysis of knee osteoarthritis (OA) trials, and the assessment of pain as an outcome, was used as an exemplar to assess how Outcome Measures in Rheumatology (OMERACT), the Cochrane Collaboration, and other international initiatives might contribute in this area. The meeting began with formal presentations of background topics, empirical evidence from the literature, and a brief introduction to 2 existing hierarchical lists of pain outcome measurement instruments recommended for metaanalyses of knee OA trials.

Results. After discussions, most participants agreed that there is a need to develop a methodology for generation of hierarchical lists of outcome measurement instruments to guide metaanalyses. Tools that could be used to steer development of such a prioritized list are the COSMIN checklist (COnsensus-based Standards for the selection of health status Measurement Instruments) and the OMERACT Filter 2.0.

Conclusion. We list meta-epidemiological research agenda items that address the frequency of reported outcomes in trials, as well as methodologies to assess the best measurement properties (i.e., truth, discrimination, and feasibility). (First Release May 1 2015; J Rheumatol 2015;42:1971–5; doi:10.3899/jrheum.141384)

Key Indexing Terms:

| OMERACT | METAANALYSIS | OSTEOARTHRITIS | PAIN |
| PATIENT OUTCOME ASSESSMENT | | COCHRANE COLLABORATION | |

From the Musculoskeletal Statistics Unit, The Parker Institute, Department of Rheumatology, Copenhagen University Hospital; Bispebjerg and Frederiksberg, Denmark; Institute of Population Health, and Department of Medicine, University of Ottawa, Ottawa, Ontario, Canada; Institute of Social and Preventive Medicine and Clinical Trials Unit Bern, Department of Clinical Research, University of Bern, Bern, Switzerland; Cochrane Collaboration, London; Department of Biostatistics, University of Liverpool, Liverpool, UK; departments of Epidemiology and Biostatistics, and Rheumatology, VU University Medical Center, Amsterdam, The Netherlands; Quintiles Inc.; Division of Rheumatology, Department of Medicine, Duke University School of Medicine, Durham, North Carolina, USA; Department of Epidemiology and Preventive Medicine, School of Public Health and Preventive Medicine, Monash University, Melbourne, and Monash Department of Clinical Epidemiology, Cabrini Health, Melbourne; Northern Clinical School, Institute of Bone and Joint Research, University of Sydney, Department of Rheumatology, Sydney, Australia; Department of Epidemiology and Biostatistics and the EMGO Institute for Health and Care Research, VU University Medical Center Amsterdam, Amsterdam, The Netherlands; Birmingham Veterans Affairs (VA) Medical Center and University of Alabama at Birmingham, Birmingham, Alabama, USA.

R. Christensen, BSc, MSc, PhD, Musculoskeletal Statistics Unit, The Parker Institute, Department of Rheumatology, Copenhagen University Hospital; L.J. Maxwell, Institute of Population Health, University of Ottawa; P. Jüni, MD, Institute of Social and Preventive Medicine and Clinical Trials Unit Bern, Department of Clinical Research, University of Bern; D. Tovey, Cochrane Collaboration; P.R. Williamson, Department of Biostatistics, University of Liverpool; M. Boers, MD, PhD, Departments of Epidemiology and Biostatistics, and Rheumatology, VU University Medical Center; N. Goel, Quintiles Inc., Division of Rheumatology, Department of Medicine, Duke University School of Medicine, and Patient Research Partner; R.S. Buchbinder, MBBS (Hons), MSc, PhD, Department of Epidemiology and Preventive Medicine, School of Public Health and Preventive Medicine, Monash University, Melbourne, and

Monash Department of Clinical Epidemiology, Cabrini Health; L. March, MD, PhD, Northern Clinical School, Institute of Bone and Joint Research, University of Sydney, Department of Rheumatology; C.B. Terwee, Department of Epidemiology and Biostatistics and the EMGO Institute for Health and Care Research, VU University Medical Center; J.A. Singh, MBBS, MPH, Birmingham VA Medical Center and University of Alabama at Birmingham; P. Tugwell, MD, MSc, The Department of Medicine, University of Ottawa.

Address correspondence to Prof. R. Christensen, Copenhagen University Hospital at Frederiksberg, Musculoskeletal Statistics Unit, The Parker Institute, Nordre Fasanvej 57, Copenhagen F, DK-2000, Denmark. E-mail: robin.christensen@regionh.dk

The Outcome Measures in Rheumatology (OMERACT) international consensus initiative has successfully developed core outcome sets for trials of many rheumatologic conditions, but its expanding scope called for clarification and updating of its underlying conceptual framework and working process[1]. The selection of appropriate outcomes, and subsequently instruments to measure these, is crucial when designing and interpreting clinical trials, to compare the effects of different interventions directly in ways that minimize bias.

Among patients with osteoarthritis (OA), the knee joint is particularly important given its importance for ambulation and thereby function of the individual. Loss of function is closely related to pain and may be more important to the patient than the actual radiographic signs of OA[2,3]. At the third OMERACT conference (1996), a core domain set was developed for future phase III trials of knee, hip, and hand OA, using a combination of discussion and polling procedures. There was consensus that the following 4 domains should be measured and reported in all clinical trials: pain, physical function, patient global assessment, and for studies of 1 year or longer, joint imaging[4]. However, no consensus was reached on which measurement instruments to apply. Consequently, in clinical trials (for example, in patients with knee problems), pain is being assessed using many different instruments[5].

Systematic reviews of knee OA randomized trials are crucial for making evidence-based decisions regarding available therapies[6,7,8]. The preparation of a protocol for a systematic review starts with a clear formulation of the research question in the PICO search strategy framework[9], i.e., a clinically relevant or policy-relevant question that takes into account Population, Intervention, Comparison, and Outcomes — including both the benefit and harm of the intervention being studied[10]. The choice of the outcome instrument of interest for any given construct, to include in metaanalyses, is generally based on clinical judgment. Because outcomes on a specific domain (e.g., pain) can each be measured using a range of different instruments, it is often necessary to standardize assessments to a common scale using the standardized mean difference[3], before the outcomes across trials can be combined in a metaanalysis.

To prevent bias and ensure transparency in reporting, it is now considered mandatory (for leading journals) that clinical trials be registered, and key protocol elements be made publicly available prior to trial completion[11]. Despite this, outcomes reported in published articles of trial results are frequently incomplete and inconsistent with their protocol where one is available[12], which can lead to bias[13] — perhaps especially for serious adverse events[14]. Following this, systematic reviewers may thereby further introduce bias into their metaanalyses using such data, because any posthoc evaluation can affect the findings of metaanalyses, either in an optimistic or pessimistic direction. Therefore, the validity of systematic reviews is potentially vulnerable if a trial has considered several outcomes but reports only those giving significant results and if the systematic review protocol has not specified outcomes and the hierarchy of outcomes in advance of its conduct[15]. Although protocol registration for systematic reviews is still not mandatory[16] except for Cochrane and Campbell systematic reviews[10], reviewers should be strongly encouraged to prespecify and register all critical issues in their protocol[17]. It follows that a published or registered and publicly available protocol[18] may help restrict the likelihood of biased posthoc decisions in review methods, such as selective outcome reporting[19].

In addition to prespecification of (a hierarchy of) outcomes in clinical trials and systematic reviews, consensus is also needed on instruments that have been used to measure these outcomes. The OMERACT Filter 2.0 explicitly describes a comprehensive conceptual framework and a recommended process to select outcomes (domains) and develop core outcome measurement sets for rheumatology trials[20]. To build on this work, a group of international experts convened to address issues regarding the need to develop hierarchical lists of outcome measurement instruments for reporting these core outcomes in metaanalyses of clinical trials, and to assess how OMERACT might contribute.

MATERIALS AND METHODS

The OMERACT Executive group identified the need to develop a deliberate dialogue between groups interested in achieving consensus on outcomes and outcome measurement instruments and to provide optimal evidence for benefit and harms in systematic reviews of studies on chronic painful musculoskeletal conditions. This article describes results of a workshop evaluating the hierarchy of pain-related continuous outcome measurement instruments recommended for metaanalyses and systematic reviews of knee OA trials; knee OA was selected as a prototype that can be used to develop a template for hierarchies in other musculoskeletal conditions.

Topics for discussion underpinning the need for consensus on the need for hierarchical lists on outcome measurement instruments for pain in metaanalyses of knee OA trials were included. The formal introductory presentations focused on specific issues, including relevant empirical evidence related to the need for a hierarchical list of patient-reported pain outcome measurement instruments.

RESULTS

The workshop participants were clinical epidemiologists, clinicians, patients, biostatisticians, and the editor-in-chief of the Cochrane Collaboration. One of the first tasks of the workshop was to review the objectives to clarify any outstanding issues among the 9 participants. The issues identified for discussion included reporting bias and multi-

plicity, as well as existing hierarchical lists of pain-related outcome measurement instruments for metaanalyses.

Reporting Bias and Multiplicity in Metaanalyses

A synthesis of recent data on this topic was presented. The group agreed that study publication and outcome reporting biases are potential threats to the validity of metaanalysis. In addition, metaanalyses of pain studies may include trials measuring pain using more than 1 instrument, thereby introducing potential bias related to issues of multiplicity[21]. Tendal, *et al* showed that the possible effect of multiple data points in trial reports regarding measurement scales, time-points, or intervention groups on metaanalysis results varies greatly across metaanalyses[21]. Consequently, data in trial reports might lead to biased decisions about which data to include in metaanalyses. Individuals should be aware of these biases as possibly making the readily available evidence unreliable for decision making[22].

After the presentations, participants emphasized that it would be valuable if a systematic review protocol specified major outcomes in terms of core domains (e.g., pain and physical function). A protocol should probably also address potential reporting bias and multiplicity issues by prespecifying a proposed hierarchy of potentially available measurement instruments. Explicit consideration of the timepoint(s) for measurement would also be preferable. One issue raised was whether it was "a given" that metaanalyses are always valuable. Most participants argued that the issues of clinical heterogeneity (e.g., population, interventions, and comparators) are important. Therefore metaanalysis is not always feasible and should always be considered in the context of the contributing studies[23]. However, because good systematic reviews should build on the premise that selected trials fulfill a specific PICO criterion (asking the same research question)[10], it should be possible to combine them, provided there is some degree of clinical homogeneity. The heterogeneity caused by different instruments measuring the same outcome is, on the other hand, not an acceptable reason for excluding evidence synthesis.

Another comment was that any proposed measurement instrument hierarchy needs to be supported by the instruments' documented psychometric properties[24], rather than developed by simply using consensus based on available data from the literature. All participants supported this notion and indicated that, before developing a hierarchy, the psychometric properties of the proposed individual instruments have to be taken into consideration. That led the group to the conceptual question of whether the "simple visual analog scale (VAS)" is valid[25], and/or whether it needs/deserves a degree of psychometric reconsideration.

Existing Hierarchical Lists of Pain-related Outcome Measurement Instruments for Metaanalyses

Jüni, *et al*[26] presented a hierarchical list of pain-related continuous outcome measurement instruments recommended for OA trials. It was suggested that the reported outcome of the instrument highest on this list should be extracted to minimize outcome reporting bias (Table 1)[26]. The list was generated based on input from a small group of OA experts, without any input from patients and without consideration of available empirical data. Currently, this proposed hierarchy is part of the preliminary default template suggested — but not yet endorsed — by the Cochrane Musculoskeletal Group (CMSG) editorial team regarding the selection and inclusion of pain measurement instruments in summary of findings tables for OA interventions.

Jüni, *et al*'s hierarchical list[26] has since been followed by another, as recommended by Juhl, *et al*[27], who concluded that choosing the instrument with the most favorable results for the intervention from individual trials could lead to biased estimates in metaanalyses. As a result, Juhl, *et al* recommended a pragmatic hierarchy based on a prioritized list (Table 1). The creation of this proposed hierarchy involved some degree of empirical evidence on the responsiveness of pain instruments derived from a systematic search conducted in the 20 highest impact factor general and rheumatology journals[27].

After the presentation, most participants agreed that a hierarchy would be desirable, making it "mandatory" for systematic reviewers to follow a prespecified list, rather than having potential selection bias at the metaanalysis stage. Among the participants there was agreement that neither of the existing papers provided compelling data to support any of the proposed hierarchies. A critical element of the discussions identified was that empirical evidence for the frequent use of any particular outcome instrument might not necessarily imply good psychometric properties[24]. Thus we were not able to achieve consensus on the "best approach" for measuring pain outcomes for metaanalyses of knee OA trials. The group was, however, supportive of using both lists (Table 1) as inspiration and a starting point for further research. Two points of view were relevant from these discussions: (1) preference was given to the continuing use of the Jüni, *et al* hierarchy[26] as a consequence of the present CMSG support[10], whereas (2) the recommended hierarchy from Juhl, *et al* has the advantage of encouraging use of the validated Western Ontario and McMaster Universities Arthritis (WOMAC) Index[28] — and thus indirectly also the Knee injury and Osteoarthritis Outcome Score[29,30] — over any VAS assessments with unclear psychometric properties[25].

Voting Results (Plenary)

After the individual workshops, a plenary session was held and all meeting participants were asked to vote on 2 questions related to a hierarchical list of pain instruments applicable for metaanalyses:

(1) Do we need a hierarchical list of instruments for a particular outcome within a particular domain for systematic reviews to prevent selective data extraction?

Table 1. Previously suggested hierarchies of pain measurement instruments applicable for metaanalyses of knee osteoarthritis (OA). Modified from Jüni, *et al*[26]: Best Pract Rheumatol 2006; and from Juhl, *et al*[27]: Arthritis 2012;136245.

Ranking	Jüni, *et al*	Juhl, *et al*
1	Global pain score	WOMAC pain subscale (Likert/100 mm)
2	Pain on walking	Pain during activity (VAS)
3	WOMAC OA index pain subscore	Pain during walking (VAS)
4	Composite pain scores other than WOMAC	Global knee pain (VAS)
5	Pain on activities other than walking (such as stair climbing)	Pain at rest (VAS)
6	WOMAC global score	SF-36 (bodily pain subscale)
7	Lequesne OA index global score	HAQ (pain subscale)
		Lequesne algofunctional index (pain subscale)
		AIMS (pain subscale)
		Knee-Specific Pain Scale
		McGill Pain Questionnaire (pain intensity)
		ASES (pain subscale)
		SES (*Schmerzempfindungsskala*)
8	Other algofunctional composite scores	Pain at night (VAS)
		Pain during activity (NRS)
		Pain on walking (NRS)
		No. painful days (days)
9	Patient's global assessment	
10	Physician's global assessment	

WOMAC: Western Ontario and McMaster Universities Osteoarthritis Index; VAS: visual analog scale; SF-36: Medical Outcome Study Short Form-36; AIMS: Arthritis Impact Measurement Scale; ASES: American Shoulder and Elbow Surgeons Shoulder Score; HAQ: Health Assessment Questionnaire; NRS: numerical rating scale.

(2) Is determining the methodology to develop a hierarchical list of instruments an important research question?

Out of 39 responders, 25 (64%) agreed that there is a need for a hierarchical list of instruments to prevent selective data extraction in metaanalyses and systematic reviews. From the voting it was clear that this topic probably needed further discussion as the number of participants who voted "Don't know" was similar to the proportion voting "No" [7 (18%) in each group]. However, the attendees were less hesitant when asked whether the methodology to develop such hierarchical lists was an important research question: 90% (35/39) voted "Yes" and the remaining 10% voted "No."

By the end of the OMERACT premeeting, there was consensus among the participants of the need to develop a hierarchical list of outcome instruments for outcomes within the pain domain for metaanalyses and systematic reviews. Because choosing the most beneficial result from available outcome measurement instruments from individual trials can over/underestimate the effect compared with a systematic approach[21,22,31], developing a protocol for the metaanalysis or systematic review and using a prioritized list will probably aid reviewers in prespecifying endpoints.

Other Aspects That Need to Be Addressed
The need for a hierarchical list of outcome instruments (to avoid bias) apparently also applies to other critical features of a systematic review, such as: (1) "How should the decision on optimal trial duration be determined?"; (2) "Is there a minimum sample size requirement?"; and (3) "Is there a minimum baseline pain intensity/frequency, etc.?"

Judging from the intensive discussions, the need for hierarchies of instruments for available core outcome measurement sets is likely to be relevant for any outcome in any health condition for which systematic reviewers might do evidence synthesis — not just pain[21]. This underlines the potential use of the existing OMERACT core domain set for guidance on possible endpoints[4]. It was also clear from the discussions and the voting results, that determining the methodology to develop hierarchical lists of instruments is an important research question.

Research Agenda
Following these valuable discussions, we suggest the following for working groups who want to propose hierarchical lists for outcome measurement instruments in metaanalyses:

- Combine meta-epidemiology[32], i.e., evaluation of the frequency of use of different instruments in the published trial literature[33];
- Examine the performance of systematic reviews of studies on different instruments' measurement properties[24]; and
- Consider the suggestion that, among outcomes frequently reported in trials based upon meta-epidemiology studies, those with the best measurement properties as assessed by frameworks such as "OMERACT Filter 2.0"[20] or the COSMIN (COnsensus-based Standards for the selection of health status Measurement Instruments) checklist[34], would be ranked high on a hierarchical list for selection of outcome measurement instruments in metaanalyses.

REFERENCES

1. Boers M, Idzerda L, Kirwan JR, Beaton D, Escorpizo R, Boonen A, et al. Toward a generalized framework of core measurement areas in clinical trials: a position paper for OMERACT 11. J Rheumatol 2014;41:978-85.
2. Altman R, Asch E, Bloch D, Bole G, Borenstein D, Brandt K, et al. Development of criteria for the classification and reporting of osteoarthritis. Classification of osteoarthritis of the knee. Diagnostic and Therapeutic Criteria Committee of the American Rheumatism Association. Arthritis Rheum 1986;29:1039-49.
3. Bliddal H, Christensen R. The treatment and prevention of knee osteoarthritis: a tool for clinical decision-making. Expert Opin Pharmacother 2009;10:1793-804.
4. Bellamy N, Kirwan J, Boers M, Brooks P, Strand V, Tugwell P, et al. Recommendations for a core set of outcome measures for future phase III clinical trials in knee, hip, and hand osteoarthritis. Consensus development at OMERACT III. J Rheumatol 1997;24:799-802.
5. Garratt AM, Brealey S, Gillespie WJ. Patient-assessed health instruments for the knee: a structured review. Rheumatology 2004;43:1414-23.
6. Zhang W, Nuki G, Moskowitz RW, Abramson S, Altman RD, Arden NK, et al. OARSI recommendations for the management of hip and knee osteoarthritis: part III: Changes in evidence following systematic cumulative update of research published through January 2009. Osteoarthritis Cartilage 2010;18:476-99.
7. Hochberg MC, Altman RD, April KT, Benkhalti M, Guyatt G, McGowan J, et al. American College of Rheumatology 2012 recommendations for the use of nonpharmacologic and pharmacologic therapies in osteoarthritis of the hand, hip, and knee. Arthritis Care Res 2012;64:455-74.
8. McAlindon TE, Bannuru RR, Sullivan MC, Arden NK, Berenbaum F, Bierma-Zeinstra SM, et al. OARSI guidelines for the non-surgical management of knee osteoarthritis. Osteoarthritis Cartilage 2014;22:363-88.
9. Guyatt GH, Oxman AD, Kunz R, Atkins D, Brozek J, Vist G, et al. GRADE guidelines: 2. Framing the question and deciding on important outcomes. J Clin Epidemiol 2011;64:395-400.
10. Ghogomu EA, Maxwell LJ, Buchbinder R, Rader T, Pardo PJ, Johnston RV, et al. Updated method guidelines for Cochrane Musculoskeletal Group systematic reviews and metaanalyses. J Rheumatol 2014;41:194-205.
11. Dickersin K, Rennie D. Registering clinical trials. JAMA 2003;290:516-23.
12. Dwan K, Altman DG, Clarke M, Gamble C, Higgins JP, Sterne JA, et al. Evidence for the selective reporting of analyses and discrepancies in clinical trials: a systematic review of cohort studies of clinical trials. PLoS Med 2014;11:e1001666.
13. Chan AW, Hrobjartsson A, Haahr MT, Gotzsche PC, Altman DG. Empirical evidence for selective reporting of outcomes in randomized trials: comparison of protocols to published articles. JAMA 2004;291:2457-65.
14. Riveros C, Dechartres A, Perrodeau E, Haneef R, Boutron I, Ravaud P. Timing and completeness of trial results posted at ClinicalTrials.gov and published in journals. PLoS Med 2013;10:e1001566.
15. Copas J, Dwan K, Kirkham J, Williamson P. A model-based correction for outcome reporting bias in meta-analysis. Biostatistics 2014;15:370-83.
16. Booth A, Clarke M, Ghersi D, Moher D, Petticrew M, Stewart L. An international registry of systematic-review protocols. Lancet 2011;377:108-9.
17. Booth A, Clarke M, Dooley G, Ghersi D, Moher D, Petticrew M, et al. The nuts and bolts of PROSPERO: an international prospective register of systematic reviews. Syst Rev 2012;1:2.
18. Booth A, Clarke M, Dooley G, Ghersi D, Moher D, Petticrew M, et al. PROSPERO at one year: an evaluation of its utility. Syst Rev 2013;2:4.
19. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gotzsche PC, Ioannidis JP, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. J Clin Epidemiol 2009;62:e1-34.
20. 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.
21. Tendal B, Nuesch E, Higgins JP, Juni P, Gotzsche PC. Multiplicity of data in trial reports and the reliability of meta-analyses: empirical study. BMJ 2011;343:d4829.
22. Dwan K, Gamble C, Williamson PR, Kirkham JJ. Systematic review of the empirical evidence of study publication bias and outcome reporting bias - an updated review. PLoS One 2013;8:e66844.
23. Gagnier JJ, Morgenstern H, Altman DG, Berlin J, Chang S, McCulloch P, et al. Consensus-based recommendations for investigating clinical heterogeneity in systematic reviews. BMC Med Res Methodol 2013;13:106.
24. Terwee CB, Mokkink LB, Knol DL, Ostelo RW, Bouter LM, de Vet HC. Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist. Qual Life Res 2012;21:651-7.
25. Huskisson EC. Measurement of pain. J Rheumatol 1982;9:768-9.
26. Jüni P, Reichenbach S, Dieppe P. Osteoarthritis: rational approach to treating the individual. Best Pract Res Clin Rheumatol 2006; 20:721-40.
27. 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.
28. Bellamy N, Buchanan WW, Goldsmith CH, Campbell J, Stitt LW. Validation study of WOMAC: a health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip or knee. J Rheumatol 1988;15:1833-40.
29. Roos EM, Roos HP, Lohmander LS, Ekdahl C, Beynnon BD. Knee Injury and Osteoarthritis Outcome Score (KOOS)—development of a self-administered outcome measure. J Orthop Sports Phys Ther 1998;28:88-96.
30. Roos EM, Toksvig-Larsen S. Knee injury and Osteoarthritis Outcome Score (KOOS) - validation and comparison to the WOMAC in total knee replacement. Health Qual Life Outcomes 2003;1:17.
31. Dechartres A, Altman DG, Trinquart L, Boutron I, Ravaud P. Association between analytic strategy and estimates of treatment outcomes in meta-analyses. JAMA 2014;312:623-30.
32. Sterne JA, Juni P, Schulz KF, Altman DG, Bartlett C, Egger M. Statistical methods for assessing the influence of study characteristics on treatment effects in 'meta-epidemiological' research. Stat Med 2002;21:1513-24.
33. Hansen JB, Juhl CB, Boutron I, Tugwell P, Ghogomu EA, Pardo PJ, et al. Assessing bias in osteoarthritis trials included in Cochrane reviews: protocol for a meta-epidemiological study. BMJ Open 2014;4:e005491.
34. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. J Clin Epidemiol 2010;63:737-45.