

Updating the OMERACT Filter: Core Areas as a Basis for Defining Core Outcome Sets

John R. Kirwan, Maarten Boers, Sarah Hewlett, Dorcas Beaton, Clifton O. Bingham III, Ernest Choy, Philip G. Conaghan, Maria-Antonietta D'Agostino, Maxime Dougados, Daniel E. Furst, Francis Guillemain, Laure Gossec, Désirée M. van der Heijde, Margreet Kloppenburg, Tore K. Kvien, Robert B.M. Landewé, Sarah L. Mackie, Eric L. Matteson, Philip J. Mease, Peter A. Merkel, Mikkel Ostergaard, Lesley Ann Saketkoo, Lee Simon, Jasvinder A. Singh, Vibeke Strand, and Peter Tugwell

ABSTRACT. Objective. The Outcome Measures in Rheumatology (OMERACT) Filter provides guidelines for the development and validation of outcome measures for use in clinical research. The “Truth” section of the OMERACT Filter presupposes an explicit framework for identifying the relevant core outcomes that are universal to all studies of the effects of intervention effects. There is no published outline for instrument choice or development that is aimed at measuring outcome, was derived from broad consensus over its underlying philosophy, or includes a structured and documented critique. Therefore, a new proposal for defining core areas of measurement (“Filter 2.0 Core Areas of Measurement”) was presented at OMERACT 11 to explore areas of consensus and to consider whether already endorsed core outcome sets fit into this newly proposed framework.

Methods. Discussion groups critically reviewed the extent to which case studies of current OMERACT Working Groups complied with or negated the proposed framework, whether these observations had a more general application, and what issues remained to be resolved.

Results. Although there was broad acceptance of the framework in general, several important areas of construction, presentation, and clarity of the framework were questioned. The discussion groups and subsequent feedback highlighted 20 such issues.

Conclusion. These issues will require resolution to reach consensus on accepting the proposed Filter 2.0 framework of Core Areas as the basis for the selection of Core Outcome Domains and hence appropriate Core Outcome Sets for clinical trials. (First Release March 15 2014; *J Rheumatol* 2014;41:994–9; doi:10.3899/jrheum.131309)

Key Indexing Terms:

OUTCOME MEASURES IN RHEUMATOLOGY OUTCOME AND PROCESS ASSESSMENT
CLINICAL TRIALS CONSENSUS DEVELOPMENT CONFERENCE

From the University of Bristol Academic Rheumatology Unit, Bristol Royal Infirmary, Bristol, UK; Departments of Epidemiology and Biostatistics, and Rheumatology, VU University Medical Center, Amsterdam, The Netherlands; University of the West of England, Academic Rheumatology Unit, Bristol Royal Infirmary, Bristol, UK; Department of Occupational Sciences and Occupational Therapy, Institute for Health Policy Management and Evaluation, University of Toronto, Toronto, Ontario, Canada; Division of Rheumatology, Johns Hopkins University, Baltimore, Maryland, USA; Section of Rheumatology, Cardiff University School of Medicine, Cardiff, UK; Division of Musculoskeletal Disease, University of Leeds; UK National Institute for Health Research (NIHR) Leeds Musculoskeletal Biomedical Research Unit, Leeds, UK; Versailles-Saint Quentin En Yvelines University, Department of Rheumatology, Ambroise Paré Hospital, APHP, Boulogne-Billancourt; Paris-Descartes University, Medicine Faculty, APHP, Cochin Hospital, Rheumatology B, Paris, France; University of California, Geffen School of Medicine, Los Angeles, California, USA; Université de Lorraine, Université Paris Descartes, Nancy; Université Pierre et Marie Curie (UPMC), Paris; AP-HP Pitié Salpêtrière Hospital, Department of Rheumatology, Paris, France; Departments of Rheumatology and Clinical Epidemiology, Leiden University Medical Center, Leiden, The Netherlands; Department of Rheumatology, Diakonhjemmet Hospital, Oslo, Norway; Department of Clinical Immunology and Rheumatology, Academic Medical Center, University of Amsterdam and Atrium Medical Center Heerlen, the Netherlands;

NIHR-Leeds Musculoskeletal Biomedical Research Unit, Division of Rheumatic and Musculoskeletal Diseases, University of Leeds, Leeds, UK; Division of Rheumatology, Mayo Clinic College of Medicine, Rochester, Minnesota, USA; Seattle Rheumatology Associates, Swedish Medical Center Rheumatology Research Division, University of Washington School of Medicine, Seattle, Washington; Division of Rheumatology, University of Pennsylvania, Philadelphia, Pennsylvania, USA; Department of Rheumatology, Copenhagen University Hospital at Glostrup, Copenhagen, Denmark; Section of Rheumatology, Louisiana State University Health Sciences Center, New Orleans, Louisiana; SDG LLC, Cambridge, Massachusetts; Birmingham Veterans Affairs Medical Center and University of Alabama at Birmingham, Birmingham, Alabama; Division of Immunology/Rheumatology, Stanford University School of Medicine, Palo Alto, California, USA; Department of Medicine, University of Ottawa, Ottawa, Ontario, Canada.

Dr. Matteson is supported in part by grants from Novartis, the American College of Rheumatology, and the European League Against Rheumatism; and was a consultant to Horizon Pharmaceuticals; Dr. Singh received research grants from Takeda and Savient, consultant fees from Savient, Takeda, Ardea, Regeneron, Allergan, URL pharmaceuticals and Novartis; is a member of the OMERACT executive, which receives arms-length funding from 36 companies; and is a member of the American College of Rheumatology's Guidelines Subcommittee of the Quality of Care Committee and the Veterans Affairs Rheumatology Field Advisory Committee.

J.R. Kirwan, MD, University of Bristol Academic Rheumatology Unit, Bristol Royal Infirmary; M. Boers, MD, PhD, Departments of Epidemiology and Biostatistics, and Rheumatology, VU University Medical Center; S. Hewlett, PhD, RN, University of the West of England, Academic Rheumatology Unit, Bristol Royal Infirmary; D.E. Beaton, BScOT, PhD, Department of Occupational Sciences and Occupational Therapy, Institute for Health Policy Management and Evaluation, University of Toronto; C.O. Bingham III, MD, Division of Rheumatology, Johns Hopkins University; E. Choy, MD, Section of Rheumatology, Cardiff University School of Medicine; P.G. Conaghan, MB, BS, PhD, FRACP, FRCP, Division of Musculoskeletal Disease, University of Leeds, and NIHR Leeds Musculoskeletal Biomedical Research Unit; M-A. D'Agostino, MD, Versailles-Saint Quentin En Yvelines University, Department of Rheumatology, Ambroise Paré Hospital, APHP; M. Dougados, MD, Paris-Descartes University, Medicine Faculty, APHP, Cochin Hospital, Rheumatology B; D.E. Furst, MD, University of California, Geffen School of Medicine; F. Guillemin, MD, PhD, Université de Lorraine, Université Paris Descartes; L. Gossec, MD, PhD, UPMC—Paris 6, Groupe de Recherche Clinique-UPMC 08 (EEMOIS) and AP-HP Pitié Salpêtrière Hospital, Department of Rheumatology; D.M. van der Heijde, MD, PhD, Department of Rheumatology, Leiden University Medical Center; M. Kloppenburg, MD, PhD, Departments of Rheumatology and Clinical Epidemiology, Leiden University Medical Center; T.K. Kvien, MD, Department of Rheumatology, Diakonhjemmet Hospital; R.B. Landewé, MD, Department of Clinical Immunology and Rheumatology, Academic Medical Center, University of Amsterdam and Atrium Medical Center Heerlen; S.L. Mackie, PhD, MD, NIHR-Leeds Musculoskeletal Biomedical Research Unit, Division of Rheumatic and Musculoskeletal Diseases, University of Leeds; E.L. Matteson, MD, MPH, Division of Rheumatology, Mayo Clinic College of Medicine; P.J. Mease, MD, Seattle Rheumatology Associates, Chief, Swedish Medical Center Rheumatology Research Division, Clinical Professor, University of Washington School of Medicine; P. Merkel, MD, MPH, Division of Rheumatology, University of Pennsylvania; M. Østergaard, MD, DMSc, Department of Rheumatology, Copenhagen University Hospital at Glostrup; L.A. Saketkoo, MD, MPH, Section of Rheumatology, Louisiana State University Health Sciences Center; L.S. Simon, MD, SDG LLC; J.A. Singh, MBBS, MPH, Birmingham Veterans Affairs Medical Center and University of Alabama at Birmingham; V. Strand, MD, Division of Immunology/Rheumatology, Stanford University School of Medicine; P. Tugwell, MD, Department of Medicine, University of Ottawa.

Address correspondence to Prof. Kirwan, University of Bristol Academic Rheumatology Unit, Bristol Royal Infirmary, Bristol, UK.
E-mail: john.kirwan@bristol.ac.uk

Outcome Measures in Rheumatology (OMERACT) strives to develop core outcome sets for rheumatologic conditions. Such core sets specify, for each condition, the minimum set of outcomes (and associated measurement instruments) necessary to provide the best estimate of benefits of an intervention. After adoption of a core set, OMERACT recommends that all studies of a health intervention in that condition report the results of these outcomes, regardless of the primary study question and the intended primary outcome measure. The original OMERACT Filter¹ describes the procedure of consensus building regarding core outcome sets, and the filter's components are summarized in 3 words: truth, discrimination, and feasibility. "Truth" is the notion that a core set measures what is intended and is unbiased and relevant. However, as OMERACT members have applied the filter in a wider range of conditions and have become associated with a broader movement to identify core outcome sets in medicine as a whole [the Core Outcome Measures in Effectiveness Trials (COMET) group²], it has become clear that this

definition presupposes an explicit way of identifying the relevant core outcomes that are universal to all studies. That this supposition may be inadequately elucidated has been highlighted since patients began to be included in the OMERACT process³. For example, as presented by S. Hewlett, in 2002 fatigue was identified by OMERACT meeting patient participants as a relevant outcome not included in the original core set³ that was subsequently found to add to our understanding of rheumatoid arthritis (RA)^{5,6,7} and finally recommended as an additional core set item in 2007⁸.

To address the question of "truth," and in particular the basis upon which core areas of outcome can be identified, a discussion paper⁹ and a literature review¹⁰ were prepared for this OMERACT 11 session. These recognized several proposals to identify essential areas of outcome assessment [e.g., the International Classification of Functioning, Disability and Health (ICF)¹¹ and its predecessors^{12,13}]. However, no proposal was found that explicitly aimed at measuring outcome as a consequence of an intervention. In addition, the development process of previous proposals was insufficiently documented; therefore, based on input from experts in the field and repeated consultations with and surveys of OMERACT and COMET attendees for more than 1 year, a new system for defining core areas of measurement was proposed specifically for discussion and possible adoption at OMERACT 11⁹.

The new proposal was laid out in detail in a pre-conference paper⁹ and presented by M. Boers. It stated that 4 core areas of outcome should be included in some manner in every clinical trial: Death, Life Impact, Resource Use, and Pathophysiological Manifestations. Under these headings, disease-specific Core Domains would be specified by groups developing core sets. In addition, contextual factors should be specified that could influence the interpretation of outcomes in the setting in which they are applied. In this OMERACT session, small discussion ("breakout") groups were presented with case studies drawn from working groups across the spectrum of OMERACT activity and invited to critically review the core area proposal in the light of the case study. Further formal and informal discussions during the OMERACT 11 meeting provided opportunities for clarifications and resolution of many areas of uncertainty before a final plenary vote at the last conference session.

Case Studies and Breakout Discussions

Five illustrative case studies were reported, each to 2 breakout groups before a discussion among OMERACT 11 delegates. Case presenters addressed specific questions on their current work (Table 1A, B). Breakout discussion groups with about 20 participants, including 2 patient partners, were first asked to consider the match between the presented case study and the proposed core area framework,

Table 1A. Summary of case studies.

Case Study	Full Title	What are the outcome domains you are currently working with?	Why have you chosen them?	What contextual factors did you consider?	How did you make these decisions?
CTD-ILD	Connective tissue diseases — interstitial lung disease	Dyspnea; health related quality of life; lung imaging; lung physiology and function; survival; cough	These domains have been identified as the most relevant and measureable for clinical trials in CTD-ILD and interstitial pulmonary fibrosis by medical and patient experts	Primary context is the randomized clinical trial; also considered was the context of clinical relevance of the domains to patient outcome in usual clinical care	Informed by expert Delphi involving rheumatologists and pulmonologists; a 3-round Delphi identified potential domains and measures; patient perspective solicited via survey and patient focus groups
PMR	Polymyalgia rheumatica	Pain; stiffness; function; systemic inflammation	Candidate outcome measures chosen for a postulated future interventional trial of an alternative to prednisolone for PMR	Age; gender; cultural background; time of day specified in the patient-reported outcome measures	Informal patient consultation; systematic literature review; work in progress
Vasculitis	ANCA-associated vasculitis	Disease activity; disease damage; patient-reported outcome; mortality	These domains have been measured in many trials and have been considered critical for both evaluating efficacy and guiding evaluation and treatment	Primary context was the randomized clinical trial; trials have often been modeled on standard of clinical care	Extrapolation from clinical trials; expert opinion; patient surveys

Table 1B. Summary of case studies.

Case Study	Full Title	What are the outcome domains you are currently working with?	Why did you decide not to use the existing Core Set?	What contextual factors did you consider?	How did you make these decisions?
OA Hand	Osteoarthritis of the hand	Domains identified in OMERACT 3; depending on the setting: pain, (physical) function and patient global assessment for symptom modifying trials and extra imaging for structure modifying trials	In the selection process patients were not involved and there was limited attention for different settings, study populations, different hand OA subsets and the co-occurrence of OA at other joint sites	Different hand OA subsets (thumb base OA, interphalangeal OA, erosive OA) and generalized OA; additionally: esthetic damage, inflammation, and thumb base prosthesis suggested	Discussions and Delphi exercises within a group of hand OA experts; work in progress;
Flare in RA	Flares in rheumatoid arthritis	Pain; physical function; tender joints; swollen joints; patient global; assessor global; laboratory measures fatigue; Stiffness; participation; self-management	A priori decision to use “bottom-up” approach to identify domains [literature review, qualitative studies (Ref. 14) with patients, expert input] to ensure all potential domains considered; domains ultimately recommended: existing RA core set plus domains initially identified by patients and then prioritized by both groups	Area of application (RCT, LOS, clinical practice), disease duration, duration in current (low) disease activity state, comorbid conditions, knowledge/experience with self-management strategies, individual context, environment	Domains identified through iterative 3 stage Delphi process of > 200 patients and health care providers (Ref. 15); Domains with > 70% agreement as important or essential in measuring flare by either patients or HCP were considered “Core”

ANCA: antineutrophil cytoplasmic antibodies; RCT: randomized controlled trials; LOS: longitudinal observational studies; HCP: healthcare professional.

its illustration of the importance of contextual factors, and to list any elements of the framework the case study had not addressed. Subsequently the groups considered in more general terms how outcome measurement sets can be

developed and addressed the question, “In the light of these considerations, do you think the proposed concepts of core areas and core domains with contextual factors offer a useful model for core domain set development?”

Report Back to Plenary and Discussion

Each breakout group reported the main points from its discussion to a plenary session of all participants. While the case studies each brought to light specific points related to particular areas of work (helpful for the OMERACT group working in that area to consider further), 20 common issues emerged requiring clarification and resolution. These themes and the broad areas where existing work was entirely compatible with the new proposal were further explored during a highly participative plenary discussion session and are summarized in Table 2.

Many participants had difficulty using death as an outcome in all circumstances. In many case studies, death was not a direct outcome of interest: it was not expected that any deaths would be related to the condition (e.g., hand osteoarthritis) or treatment (e.g., physiotherapy) or would occur in the time window under investigation (e.g., the short-term response to intramuscular glucocorticoids). Participants did recognize that any death occurring during a clinical trial would need to be reported regardless of perceived causality. Other participants raised the possibility that “states worse than death” may be experienced by patients and wondered how this would be dealt with within the framework.

The concept and importance of effect on life as a core area (including function, quality of life, the ICF domains¹¹, patient perception of health, etc.) was widely accepted. Debate centered on whether, at a core area level, life impact should be subdivided further. Several different suggestions were made, such as work-related problems, mobility and

independence, and social interactions, but these were often relevant in only 1 disease group or 1 particular context. No clear consensus about further subdivisions of life impact emerged.

“Considering Resource Use as a Core Area” produced the greatest discussion. Many participants saw this as an economic evaluation that was worth undertaking only in studies designed for that purpose. Some felt that resources included family support, support at work to continue working, personal time and effort of the patient, opportunity costs to the healthcare system, etc. Other remarks addressed the costs/feasibility of adequately capturing resource use; and the early development phase of a therapy where true resource costs may not be relevant to the question (e.g., in proof-of-principle studies), or not even calculable (e.g., the final cost of the therapy might depend on technical manufacturing issues and market forces). In sum, this area, while very important and relevant in many circumstances, was felt to require additional discussion and delineation before it could be considered a definite core area.

The concept of pathophysiological manifestations also produced some debate as a core area. There was general recognition that some information concerning the underlying disease process and its activity was needed to measure the effects of any treatment, and that most of our existing outcome measures focused on evaluating this core area. For example, most of the current RA core set instruments (joint counts, acute-phase reactants, and imaging) measure a pathophysiologic manifestation of the underlying disease process. There was some confusion as to the way this area

Table 2. Issues emerging from breakout groups requiring clarification and resolution before the Core Area model could be fully accepted.

Topic	Issue
Death	Death may not be an outcome of interest Should states worse than death be mentioned?
Life impact	Should life impact be subdivided further?
Resource use	What does “resource use” mean? Are there any surrogates? What point of view is considered (patient, health system, society)? Will measurement of resource use be impractical in many trials?
Pathophysiological manifestations	Can clinical signs (and sometimes symptoms) also indicate pathophysiological status? Need to be flexible about how this is defined
Contextual factors	Can we better define what these contextual factors are? Can we provide a list? Can we better distinguish between factors? Who decides what is required?
Some general issues	Can we provide more concrete examples? Are adverse effects a core area in themselves? Difference between domains and instruments unclear Will instruments crossing domains be a problem?
Process issues	Difference between core areas and primary and secondary outcomes Does core set development come to a stop if one or more core domains does not have a validated instrument? There should be provision for updating or revision of core outcome sets as further data accumulate

might be assessed — could symptoms of pain and swelling in a joint be used as a measure of pathophysiology? There was a feeling that this needs to be defined in a flexible way.

There was wide recognition of a strong conceptual need to consider contextual factors. Confounding factors, comorbidities, variation in healthcare systems, and factors related to psychological status were all identified as potential contextual factors. However, it was unclear to many participants how these factors would be identified and which would be labeled as “core” to particular investigations. A number of general issues emerged from the breakout group reports and the plenary discussion. A recurrent theme was the request to provide concrete examples for the theoretical framework. Whether adverse effects should be a core area was also a topic of disagreement and uncertainty. All agreed that adverse effects should be reported in trials, and it was recognized that this often constitutes a specific section within a clinical trial report. However, it was also agreed that any given adverse effect would occur under the umbrella of one of the proposed core areas.

There were many points raised in which issues related to choosing, testing, and developing specific instruments became entangled with questions of whether a core area or a particular core domain would be adequately addressed and hence whether a core set of outcome measures would then be achievable within the proposed framework. There was also a recognition that many existing instruments, such as questionnaires, relate to more than 1 core area, and participants were unclear whether this would allow separate assessment of different core areas.

Some participants were unsure of the difference between core areas, primary outcomes, and secondary outcomes, and wondered whether core areas were intended to be the primary or secondary outcome measures. They feared this might override the intention of trial designers in setting up the study protocol. There was also concern that the work of core outcome development might come to a halt if, in relation to a particular condition, a core domain was identified but no valid assessment instrument existed in that domain. This led to a fruitful exploration of the difference between core domains and core outcome sets, and a clearer understanding that there is a 2-step process in defining first core domains within the core areas, and second identifying (or devising) instruments to include in the core outcome set. Finally, and perhaps unsurprisingly, there was a strong call for inclusion of a review process to ensure that as data accumulate, the whole philosophy of the emerging Filter 2.0 framework would be regularly scrutinized and updated.

This OMERACT session was deliberately constructed to test the proposed Filter 2.0 framework of core areas, core domains and contextual factors, which had already been subject to discussion, debate, and extensive development before the meeting. Using case studies from different working groups allowed participants to probe the theoretical

and practical implications of the framework, and to look for areas of strength and weakness. There was a broad agreement with the need to formalize an overarching structure to justify the subsequent selection of core domains. Until challenged by the introduction of the patient perspective and the emergence of the COMET initiative, the OMERACT community has, in effect, been relying on clinicians’ common understanding of the disease areas in which they are working. This workshop, which took place at the start of the OMERACT 11 conference, concluded from case studies and discussions that most of the current work of the OMERACT participants already fits into the principles of the new framework, but several important areas of uncertainty emerged, as described. If sufficient consensus was to be achieved in time for the plenary session at the end of the conference¹⁶, these areas would need to be clarified and addressed further by the Filter 2.0 development group.

REFERENCES

1. Boers M, Brooks P, Strand V, Tugwell P. The OMERACT Filter for outcome measures in rheumatology. *J Rheumatol* 1998;25:198-9.
2. Core Outcome Measures in Effectiveness Trials (COMET) initiative. [Internet. Accessed January 24, 2014.] Available from: www.comet-initiative.org/
3. Kirwan J, Heiberg T, Hewlett S, Hughes R, Kvien T, Ahlmén M, et al. Outcomes from the Patient Perspective workshop at OMERACT 6. *J Rheumatol* 2003;30:868-72.
4. Boers M, Tugwell P, Felson DT, van Riel PLCM, Kirwan JR, Edmonds JP, et al. World Health Organization and International League of Associations for Rheumatology core endpoints for symptom modifying antirheumatic drugs in rheumatoid arthritis clinical trials. *J Rheumatol Suppl.* 1994 Sept;41:86-9.
5. Hewlett S, Cockshot Z, Byron M, Kitchen K, Tipler S, Pope D, et al. Patients’ perceptions of fatigue in rheumatoid arthritis: overwhelming, uncontrollable, ignored. *Arthritis Rheum* 2005;53:697-702.
6. Ahlmen M, Nordenskiöld U, Archenholtz B, Thyberg I, Rönnqvist R, Lindén L, et al. Rheumatology outcomes: the patient’s perspective. A multicentre focus group interview study of Swedish rheumatoid arthritis patients. *Rheumatology* 2005;44:105-10.
7. Hewlett S, Hehir M, Kirwan J. Measuring fatigue in rheumatoid arthritis: a systematic review of scales in use. *Arthritis Rheum* 2007;57:429-539.
8. Kirwan JR, Minnock P, Adebajo A, Bresnihan B, Choy E, de Wit M, et al. Patient perspective: fatigue as a recommended patient centered outcome measure in rheumatoid arthritis. *J Rheumatol* 2007;34:1174-7.
9. Boers M, Idzerda L, Kirwan JR, Beaton D, Escorpizo R, Boonen A, et al. Towards a generalized framework of core measurement areas in clinical trials: a position paper for OMERACT 11. *J Rheumatol* 2014;41:978-85.
10. Idzerda L, Rader T, Tugwell P, Boers M. Can we decide which outcomes should be measured in every clinical trial? A scoping review of the existing conceptual frameworks and processes to develop core outcome sets. *J Rheumatol* 2014;41:986-93.
11. World Health Organization. International Classification of Functioning, Disability and Health. [Internet. Accessed January 24, 2014.] Available from: www.who.int/classifications/icf/en/
12. World Health Organization. International Classification of Diseases (ICD-10). [Internet. Accessed December 17, 2013.] World Health Organization: 1990. Available from: www.who.int/classifications/icd/en

13. World Health Organization. International Classification of Impairments, Disabilities and Handicaps. Geneva, 1980. [Internet. Accessed January 24, 2014.] Available from: http://whqlibdoc.who.int/publications/1980/9241541261_eng.pdf
14. Hewlett SE, Sanderson T, May JE, Alten R, Bingham CO 3rd, March L, et al. "I'm hurting, I want to kill myself": rheumatoid arthritis flare is more than a high joint count — an international patient perspective on flare where medical help is sought. *Rheumatology* 2012;51:69-76.
15. Bartlett SJ, Hewlett SE, Bingham CO 3rd, Woodworth TG, Alten R, Pohl C, and the OMERACT Flare Working Group. Identifying core domains to assess flare in rheumatoid arthritis: an OMERACT international patient and provider combined Delphi consensus. *Ann Rheum Dis* 2012;71:1855-60.
16. Boers M, Kirwan JR, Gossec L, Conaghan PG, D'Agostino M-A, Bingham III CO, et al. How to choose core outcome measurement sets for clinical trials: OMERACT 11 approves Filter 2.0. *J Rheumatol* 2014;1025-30.