Instrument Selection Using the OMERACT Filter 2.1: The OMERACT Methodology

Dorcas E. Beaton, Lara J. Maxwell, Beverley J. Shea, George A. Wells, Maarten Boers, Shawna Grosskleg, Clifton O. Bingham III, Philip G. Conaghan, Maria Antonietta D'Agostino, Maarten P. de Wit, Laure Gossec, Lyn M. March, Lee S. Simon, Jasvinder A. Singh, Vibeke Strand, and Peter Tugwell

ABSTRACT. ***Objective.*** Outcome Measures in Rheumatology (OMERACT) Filter 2.1 revised the process used for core outcome measurement set selection to add rigor and transparency in decision making. This paper describes OMERACT's methodology for instrument selection.

Methods. We presented instrument selection processes, tools, and reporting templates at OMERACT 2018, introducing the concept of "3 pillars, 4 questions, 7 measurement properties, 1 answer." Truth, discrimination, and feasibility are the 3 original OMERACT pillars. Based on these, we developed 4 signaling questions. We introduced the Summary of Measurement Properties table that summarizes the 7 measurement properties: truth (domain match, construct validity), discrimination [test-retest reliability, longitudinal construct validity (responsiveness), clinical trial discrimination, thresholds of meaning], and feasibility. These properties address a set of standards which, when met, answer the one question: Is there enough evidence to support the use of this instrument in clinical research of the benefits and harms of treatments in the population and study setting described? The OMERACT Filter 2.1 was piloted on 2 instruments by the Psoriatic Arthritis Working Group.

Results. The methodology was reviewed in a full plenary session and facilitated breakout groups. Tools to facilitate retention of the process (i.e., "The OMERACT Way") were provided. The 2 instruments were presented, and the recommendation of the working group was endorsed in the first OMERACT Filter 2.1 Instrument Selection votes.

Conclusion. Instrument selection using OMERACT Filter 2.1 is feasible and is now being implemented. (First Release June 1 2019; J Rheumatol 2019;46:1028–35; doi:10.3899/jrheum.181218)

Key Indexing Terms:

OUTCOME MEASURES | OUTCOME ASSESSMENT | HEALTH STATUS INDICATOR
OMERACT | REPRODUCIBILITY OF RESULTS | RELIABILITY AND VALIDITY

From the Institute for Work & Health and Institute for Health Policy Management and Evaluation, University of Toronto, Toronto; Clinical Epidemiology Program, and Centre for Practice-Changing Research, Ottawa Hospital Research Institute; Cardiovascular Research Methods Centre, University of Ottawa Heart Institute; Division of Rheumatology, Department of Medicine, and School of Epidemiology and Public Health, Faculty of Medicine, University of Ottawa, Ontario, Canada; Department of Epidemiology and Biostatistics, and Department of Medical Humanities, Amsterdam UMC, Vrije Universiteit, Amsterdam, the Netherlands; Division of Rheumatology, Department of Medicine, Johns Hopkins University, Baltimore, Maryland; Medicine Service, VA Medical Center; Department of Medicine, School of Medicine, University of Alabama; Division of Epidemiology, School of Public Health, University of Alabama, Birmingham, Alabama; SDG LLC, Cambridge, Massachusetts, USA; Leeds Institute of Rheumatic and Musculoskeletal Medicine, University of Leeds; UK National Institute for Health Research (NIHR) Leeds Biomedical Research Centre, Leeds, UK; Hôpital Ambroise Paré, Rheumatology Department, Boulogne-Billancourt; INSERM U1173, Laboratoire d'Excellence INFLAMEX, UFR Simone Veil, Versailles-Saint-Quentin University, Saint-Quentin en Yvelines; Sorbonne Université; Pitié Salpêtrière hospital, AP-HP, Rheumatology Department, Paris, France; Sydney Medical School, Institute of Bone and Joint Research and Department of Rheumatology, Royal North Shore Hospital, St. Leonards, Australia.

PGC is funded in part by the NIHR Leeds Biomedical Research Centre. The views expressed are those of the authors and not necessarily those of the UK National Health Service, the NIHR, or the Department of Health. JAS is supported by the resources and the use of facilities at the VA Medical Center at Birmingham, Alabama. LMM is a Principal Investigator on the Australian Rheumatology Association Database, which has received arms-length funding from AbbVie Australia, Pfizer Australia, Janssen Australia, and Eli Lilly Australia. OMERACT is a registered nonprofit independent medical research organization whose goal is to improve and advance the health outcomes for patients with musculoskeletal conditions. OMERACT receives unrestricted educational grants from the American College of Rheumatology, the European League of Rheumatology, and several pharmaceutical companies. The grants are used to support fellows, international patient groups, and a major international biennial conference that results in many peer-reviewed publications. The views expressed in this article are those of the authors and do not necessarily reflect the position of the US Department of Veterans Affairs or the US government.

D.E. Beaton, PhD, Senior Scientist, Institute for Work & Health, Department of Occupational Science and Occupational Therapy, Rehabilitation Sciences Institute, and Associate Professor, Institute Health Policy, Management and Evaluation, University of Toronto; B.J. Shea, PhD, Senior Scientist, Ottawa Hospital Research Institute, Clinical Epidemiology Program, and School of Epidemiology, Public Health and Preventive Medicine, Faculty of Medicine, University of Ottawa; L.J. Maxwell, PhD, Methodological Advisory to the Technical Advisory Group, OMERACT, University of Ottawa, and Centre for Practice-Changing Research, Ottawa Hospital Research Institute; G.A. Wells, MSc,

Personal non-commercial use only. The Journal of Rheumatology Copyright © 2019. All rights reserved.

Downloaded on April 18, 2024 from www.jrheum.org

Core outcome sets (COS) are increasingly recognized as a minimum set of outcomes that will be measured across all clinical trials in a given field to facilitate comparisons of interventions and metaanalyses, and to avoid selective outcome reporting bias[1]. The Outcome Measurement in Rheumatology (OMERACT) has promoted and supported the development of COS since its inception in 1992[2]. Although the main focus has been in the area of musculoskeletal disorders and rheumatologic conditions[3], it has also found application in other fields[4,5].

OMERACT divides the task of creating a COS into 2 components: first, determining what needs to be measured (core domain sets), and second, deciding how to measure each of the domains, also referred to as "instrument selection." This in turn leads to a core outcome measurement set, when there is at least 1 outcome measurement instrument identified for each domain. In 2012, OMERACT voted to revise its processes to recognize both the growth of the organization and of the literature available on measurement properties of any given outcome measurement instrument. The creation of a core domain set was outlined by Boers, *et al*[6] and is expanded on in this issue in 2 companion papers[7,8]. The purpose of this paper is to describe a data-driven, evidence-based process for the instrument selection process and the OMERACT Filter 2.1 methodology.

MATERIALS AND METHODS

Foundations of the OMERACT Filter: 3 pillars, 4 questions, 7 measurement properties, 1 answer. Truth, discrimination, and feasibility are the pillars of the OMERACT Filter[9]. Truth refers to whether the measure's scores can be shown to be truthful, measuring what was intended. Discrimination asks whether the measure discriminates between situations of interest, such as between treatment arms in a clinical trial. Finally, feasibility answers questions about the practicality of using the tool: time, cost, and burden. Together, these 3 pillars describe a set of standards which, when met, answer one question: Is there enough evidence to support the use of this instrument in clinical research of the benefits and harms of treatments in the population and study setting described?

In OMERACT Filter 2.1, we recognized that the 3 pillars of the original OMERACT Filter are best represented by 4 signaling questions (Figure 1A). Two questions split the truth pillar into a practical appraisal of the instrument and its content with "Is it a match with the target domain?", and a more data-driven, hypothesis-testing assessment of the instrument's scores with "Do the numeric scores make sense (i.e., are the scores relating to other measures or the testing situation in a way it should if it measures the domain well)?" The question reflecting the discrimination pillar is "Can it discriminate between groups of interest?", assessing whether the instrument identifies differences between treatment and control groups found in clinical trials. The signaling question, "Is it practical to use?" i.e., in cost, burden, and access, reflects the feasibility pillar.

In practice, when this method is used to assess an instrument, the signaling questions are slightly reordered, putting practical appraisals of concept match and feasibility ahead of the review of the evidence available on the more data-driven features of testing truth and discrimination. This saves time and resources because it allows instruments to be set aside if they are not identifying the target domain concept or are not feasible for use in the target application. This reordering is seen in the bottom of Figure 1B.

The 4 signaling questions and the traffic light ratings they received are linked on the OMERACT Filter 2.1 Instrument Selection Algorithm (Figure 2). Ratings are completed for each question and then combined into an overall rating for the instrument. Red always means "stop, do not continue," amber means "a caution is raised, but you can continue," and green means "go, this question is definitely answered affirmatively." White circles indicate an absence of evidence, leaving working groups to decide whether they wish to create the evidence needed or consider it as a gap, so further evaluation should stop because evidence is missing. Once all 4 questions are answered, based on this evidence the working group recommends an overall level of endorsement (Figure 2 bottom panel).

Instrument selection using the OMERACT methodology. The step-by-step process of OMERACT's instrument selection methodology will be described briefly here following the steps illustrated in Figure 3, "How to choose an instrument the OMERACT Way." A detailed description of these steps is available in The OMERACT Handbook[10].

1. Revisit the domain definition. Prior to embarking on any instrument selection process, working groups should review the domain(s) each instrument is trying to identify. This is done making use of the definitions described in the OMERACT Onion document[7] and the OMERACT Filter 2.1 Framework[8].

2. Find candidate instruments. Creating a new instrument is a difficult task, and groups often can identify an existing instrument(s) by searching the literature[11,12,13] or speaking to experts in the field.

3. Is the instrument a match for the target domain? Working groups then address the signaling questions described above. Armed with the domain definition and the candidate instrument, working groups can identify whether the instrument or outcome measure (terms used here interchangeably) matches the intended target domain. This is done by seeking the experiences of those who will respond to the instrument. Working groups should talk to people, particularly those with the lived experience of the disease and domain, to see whether the instrument identifies the breadth and depth of the experience. Templates for surveying respondents are provided in the OMERACT Instrument Selection Workbook (www.omeract.org/resources). Available data can be used to examine whether the response distribution for

PhD, Director, Cardiovascular Research Methods Centre, University of Ottawa Heart Institute, and Professor, Department of Epidemiology and Community Medicine, University of Ottawa; M. Boers, MD, PhD, Professor of Clinical Epidemiology, Amsterdam UMC, Vrije Universiteit Amsterdam; S. Grosskleg, BS, Secretariat, OMERACT, and University of Ottawa; C.O. Bingham III, MD, Professor of Medicine, Division of Rheumatology, Johns Hopkins University; P.G. Conaghan, MD, PhD, Director, Leeds Institute of Rheumatic and Musculoskeletal Medicine, University of Leeds, and Deputy Director, NIHR Leeds Biomedical Research Centre; M.A. D'Agostino, MD, PhD, AP-HP, Hôpital Ambroise Paré, Rheumatology Department, and INSERM U1173, Laboratoire d'Excellence INFLAMEX, UFR Simone Veil, Versailles-Saint-Quentin University; M.P. de Wit, PhD, OMERACT Patient Research Partner, and Amsterdam University Medical Centre, Department Medical Humanities, Amsterdam Public Health; L. Gossec, MD, PhD, Sorbonne Université, and AP-HP, Pitié Salpêtrière Hospital, Department of Rheumatology; L.M. March, MBBS, PhD, Liggins Professor of Rheumatology and Musculoskeletal Epidemiology, Sydney Medical School, Institute of Bone and Joint Research, and Head of Department of Rheumatology, Royal North Shore Hospital; J.A. Singh, MBBS, MPH, Professor of Medicine and Epidemiology, Department of Medicine at the School of Medicine, University of Alabama at Birmingham; L.S. Simon, MD, Co-managing director of SDG LLC; V. Strand, MD, MACR, FACP, Biopharmaceutical Consultant; P. Tugwell, MD, MSc, Professor, Division of Rheumatology, Department of Medicine, and School of Epidemiology and Public Health, Faculty of Medicine, University of Ottawa, and Clinical Epidemiology Program, Ottawa Hospital Research Institute.

Address correspondence to D.E. Beaton, Senior Scientist, Institute for Work & Health, 481 University Ave., Toronto, Ontario M5G 2E9, Canada. E-mail: dbeaton@iwh.on.ca

Accepted for publication January 24, 2019.

Personal non-commercial use only. The Journal of Rheumatology Copyright © 2019. All rights reserved.

Downloaded on April 18, 2024 from www.jrheum.org

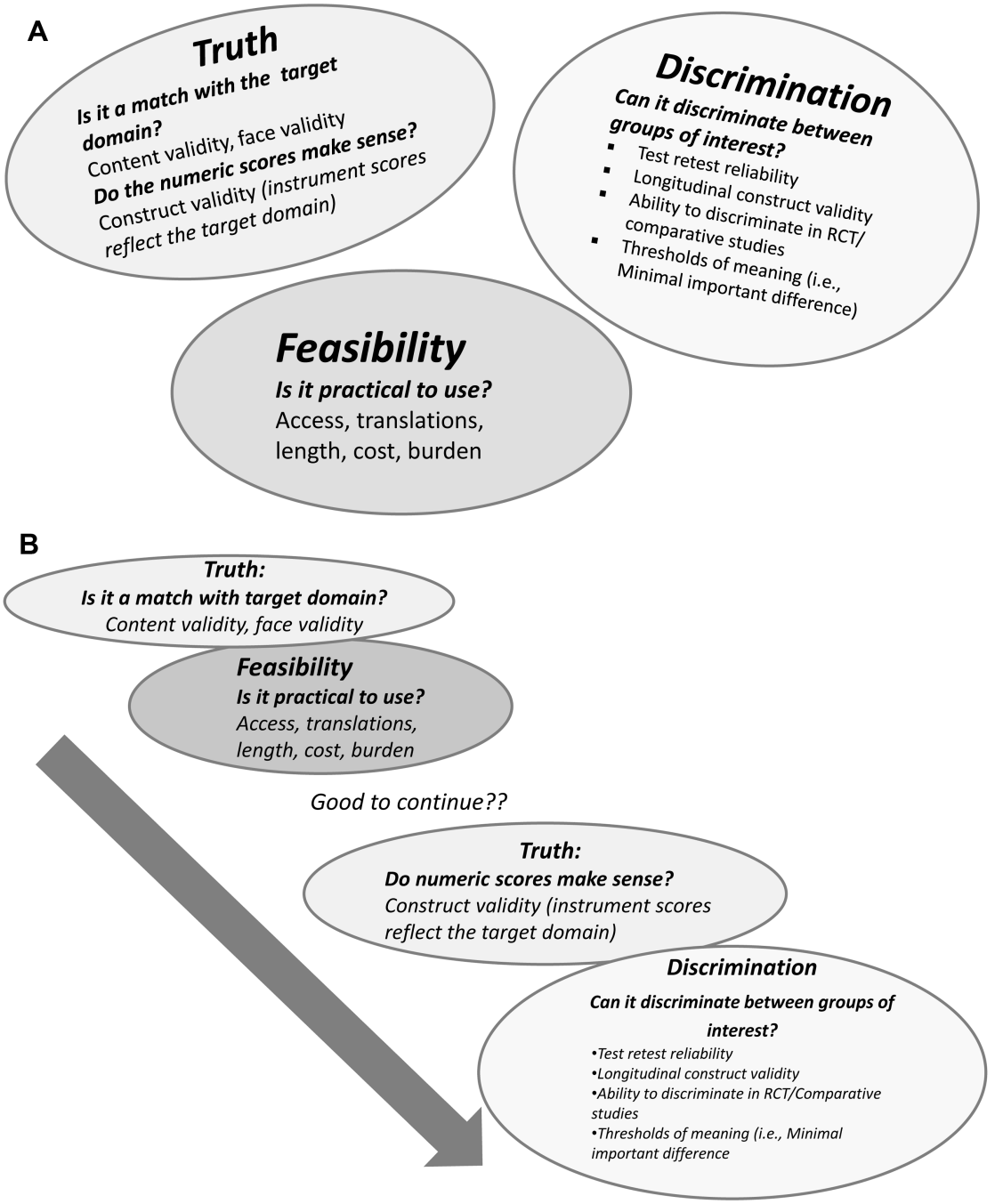

Figure 1. (A) The 3 OMERACT Filter Pillars of Truth, Discrimination, and Feasibility (circles) and the signaling questions involved in each. Measurement properties required to answer each signaling question (7 in total) are listed. (B) Pragmatic reordering of signaling questions, separating the 2 "Truth" questions and inserting "Is it practical to use? (Feasibility)" between. This order now reflects increasing investment of time and effort, and reflects decision-making nodes. If there is a "no" to either domain match or feasibility, there is no need to continue to the more difficult stage of finding or creating evidence of the other properties. OMERACT: Outcome Measure in Rheumatology; RCT: randomized controlled trials.

the scale is appropriate. High ceiling or floor effects in people experiencing the domain (i.e., physical limitation) could flag that the scale will not detect the differences of interest in the relevant population or could also reflect an expected level for certain indices or aggregate scores[14]. Cognitive interviews can be used at this stage to examine how items are interpreted; for example, whether people, particularly those with the lived experience of the disease and domain, would prefer different question stems, anchors, or response options[15].

4. Is it feasible to use this outcome measure? Feasibility is a practical assessment of the burden of use, where burden could be cost, time, equipment, personal burden for the respondent (e.g., language, health literacy) or admin-

Personal non-commercial use only. The Journal of Rheumatology Copyright © 2019. All rights reserved.

Downloaded on April 18, 2024 from www.jrheum.org

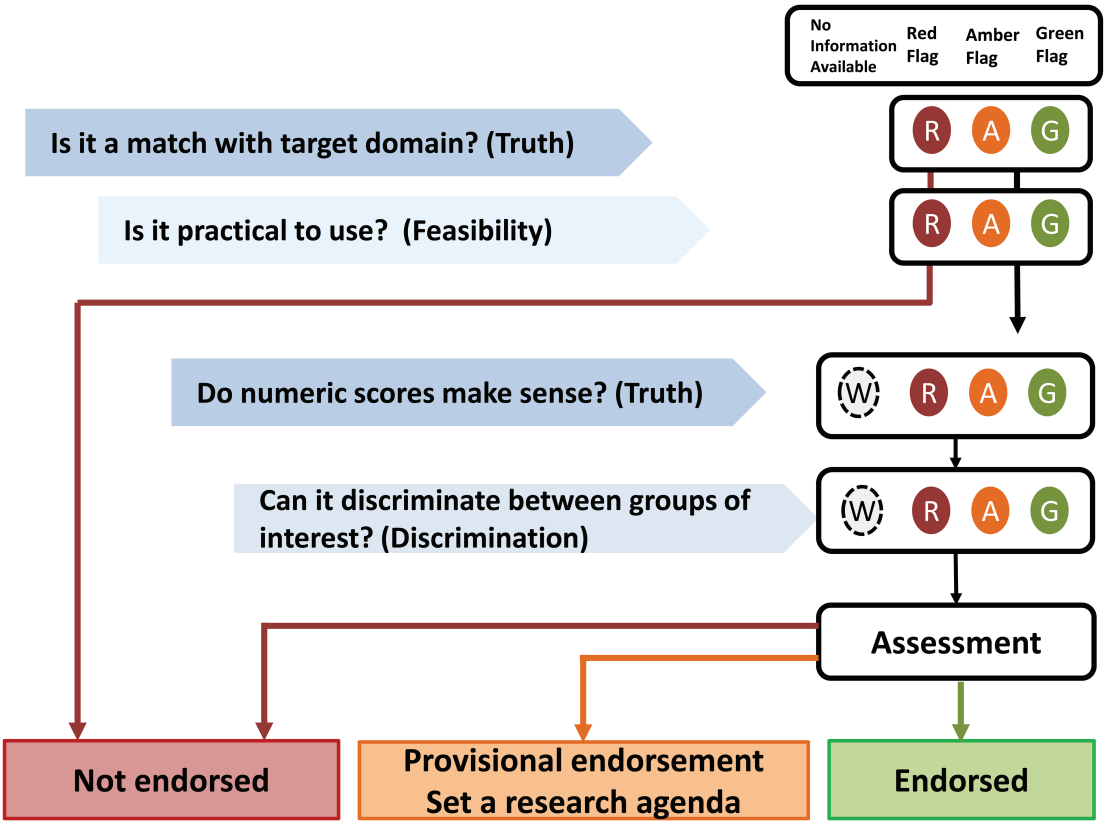

Figure 2. OMERACT Filter 2.1 Instrument Selection Algorithm (OFISA). The 4 signaling questions are linked to a results column (traffic light ratings), and a renewed emphasis on the setting aside of instruments that receive a red rating for either of the first 2 questions. Amber and green continue to the last 2 signaling questions, though the former are to be used with care and caution. OMERACT: Outcome Measure in Rheumatology; W: white; R: red; A: amber; G: green.

istrator (e.g., required training), the interpretability of the scores, and other similar considerations[16]. Some of these features can be assessed using surveys or checklists compiled with working group and other input (see OMERACT Instrument Selection Workbook; www.omeract.org/resources) or through other structured techniques in focus groups or nominal group processes. Occasionally assessments of feasibility (time to complete the assessment or survey, complexity of language, or technical demands of interpreting imaging results) are published in the literature; however, OMERACT will also accept the appraisal of the working group for the answer to this question.

5. Narrowing the number of candidate measures. At the next stage, the working group determines whether there is a clear match of an outcome measurement instrument with the target domain and whether the instrument is feasible for use in the intended setting. An instrument that is not a good match to the target domain definition or is not feasible should be set aside, because these shortcomings are unlikely to be easily addressable. This is a key step in the process and often leads to a shortened list of candidate instruments. Working groups are asked to record the level of agreement within their working group and any comments made when either proceeding or setting aside an outcome measurement instrument at this point.

6. Gather evidence for the next 2 signaling questions. The last 2 questions ("Do the numeric scores make sense?" and "Can it discriminate between groups of interest?") are represented by 5 additional measurement properties that require data-oriented answers: construct validity (scores relate to other known measures in a way that is consistent with the underlying domain of interest), test-retest reliability (no change in score when patients are stable, estimate of day-to-day variability), longitudinal construct validity (responsiveness; ability to detect change when it has occurred), ability to discriminate in a clinical trial (specific ability to detect change between arms in a clinical trial), and thresholds of meaning (benchmarking scores and changes in score for interpretation; as seen in Figure 1B). The evidence to support performance of an instrument on each of these properties is based on the growing body of literature on measurement properties[17,18]. In response to this, the OMERACT Filter 2.1 has adopted standard systematic review techniques as described by Slavin[19] to identify and process available literature. Slavin describes the stages of such a review as (i) gathering the evidence, (ii) appraisal of quality of the evidence, (iii) data extraction, and (iv) synthesis of findings. The result is parallel systematic reviews, one for each of the measurement properties of interest. The process is described briefly here and in more detail in The OMERACT Handbook[10].

i) Gathering the evidence on the measurement properties. Systematic literature searches are conducted with the support of library scientists and standard search term templates available to working groups. The search terms focus on the measurement properties and the relevant patient population for the outcome measure. Searches are run often by a librarian or information scientist; the working group screens the titles and abstracts to see if they match the instrument and to ensure they are about measurement properties. Positive or possible articles are obtained for full–text review of their relevance, and to see which measurement properties are addressed in that article. Working groups at this point begin building their Summary of Measurement Property (SOMP) table, where the relevant articles are listed (Table 1) and the measurement properties covered are recorded. Importantly,

Personal non-commercial use only. The Journal of Rheumatology Copyright © 2019. All rights reserved.

Downloaded on April 18, 2024 from www.jrheum.org

Figure 3. The Outcome Measure in Rheumatology (OMERACT) Way flowchart describing the step-by-step process of OMERACT's instrument selection methodology.

Personal non-commercial use only. The Journal of Rheumatology Copyright © 2019. All rights reserved.

Downloaded on April 18, 2024 from www.jrheum.org

only the 7 measurement properties relevant to the application of an existing measure in a clinical trial are reviewed. Tracking of the yield and selection of articles should be rigorous and reported in a PRISMA flow chart (http://www.prisma-statement.org/).

Evidence for OMERACT endorsement can also be created by the working groups by conducting a study to address any gaps found in the SOMP table. The methods and results of these studies are independently reviewed by at least 2 members of the Technical Advisory Group of OMERACT (https://omeract.org/tag) before they are considered for inclusion.

ii) Quality assessment. All evidence, both that found in the literature and new evidence created by the working group, undergoes quality assessment. Several quality assessment tools exist in the literature, though few specifically address our goal of looking to exclude those with critical flaws that could lead to a risk of bias in the estimation of the measurement property performance. COSMIN (4-point checklist version) is one frequently used critical appraisal tool for measurement studies[20]. Only certain items in the checklist offer a "poor" response category. This rating is reserved to indicate the situations in which the methods reported are flawed enough that this evidence should not be included in the review owing to risk of bias. In 2015, we worked with the COSMIN and reworded these specific items into a positive, dichotomized response to identify whether the study reported good methods, and had successfully avoided a risk of bias as indicated in that poor rating. Focusing only on measurement properties needed for OMERACT Filter 2.1, we added 2 measurement properties important to OMERACT that were not in COSMIN (clinical trial discrimination and thresholds of meaning), to produce the COSMIN-OMERACT Good Methods Checklist found in our current OMERACT Handbook[10].

The Good Methods Checklist items are assessed independently by 2 persons, and agreement is sought between them. Any newly created evidence has the Good Methods check done by 2 members of the technical advisory group independent of the working group. This is rated in traffic light format again and the color entered in the cells of the SOMP table (Table 1), with green or amber indicating good methods, and red indicating a high risk of bias. Only studies that have passed the Good Methods Check move to the next stage of extracting information and the results of the measurement property tests.

Table 1. OMERACT Summary of Measurement Properties table. A summary of the work leading up to the working group's decision about an instrument and whether it has passed the OMERACT Filter of truth, discrimination, and feasibility. In this table, fictitious studies are shown for demonstration only. All selected articles are listed and the measurement properties they studied noted. Color reflects the "good methods check" with green saying good avoidance of risk of bias, amber meaning some concerns, and red being a red flag for risk of bias. Only amber and green cells are used in synthesis. OMERACT: Outcome Measure in Rheumatology; RCT: randomized controlled trials; +: surpasses standard for good performance; -: does not surpass performance; +/-: equivocal findings.

Author/Year	Truth	Feasibility	Truth	Discrimination			
	Domain Match		Construct Validity	Test retest Reliability	Longitudinal Construct Validity (Responsiveness)	Clinical Trial Discrimination	Thresholds of Meaning
Lennon 1991			+				
McCartney 2004					+		
Harrison 2004					+	+/-	
Starr 2005				+	+/-	+	+
Best 2006					+		+
Sutcliffe 2006							+
Boers 2007					+/-		+/-
Tugwell 2009							+
Strand 2010	+						
Simon 2010				+	-		+
Brooks 2015	+						
Total available studies for each property	2	0	2	2	6	2	6
Total studies available for synthesis	2	0	1	2	6	2	6
Rating (RAGW) [put on Master Checklist]	Green	Green	Amber	Green	Green	Amber	Green
Overall rating for instrument across properties [Options: Endorsed, Provisional Endorsement, Not endorsed]	Provisional endorsement: needs additional construct and RCT discrimination						

Personal non-commercial use only. The Journal of Rheumatology Copyright © 2019. All rights reserved.

Downloaded on April 18, 2024 from www.jrheum.org

iii) Data extraction. The results of the testing of measurement properties are extracted from the publications and placed into a narrative summary of the testing procedures, study characteristics, and results. Enough detail is provided in a data extraction table to allow a user of the data to follow the logic and rationale for the decisions made. Results are compared to international recommendations for acceptable performance in terms of results of a measurement property study. In the SOMP table, a "+" is placed for a positive performance, "+/–" for equivocal, and "–" for inadequate performance.

iv) Synthesis. The next step is the synthesis of evidence that has been appraised as at least adequate-quality evidence (green or amber color), into a rating of the performance of the instrument for each of the 7 measurement properties in the SOMP. Both published and new studies are considered. Our synthesis methods are based on the practices of several groups in different fields[4,20,21], which emphasize the importance of having Quality information (using studies with good methods); Quantity (at least 2 good methods studies), showing Consistency of the findings across these pieces of evidence; and adequate Performance in the tests of that measurement property. Combining these elements, Quality, Quantity, Consistency, and Performance (QQC-P), a synthesis statement is made for each measurement property. The working group then decides on a recommendation based on their good quality evidence.

3.7 Identify the "winners" (best instruments). In the last row of the SOMP, the working group identifies the instrument(s) that have passed the Filter 2.1 requirements with either a green (endorsed) or amber (provisionally endorsed) rating at the instrument level. All amber-rated instruments must have a clearly defined research agenda of what additional work is needed to bring this instrument to a green for full endorsement.

3.8 Bring it to a vote. Core to the OMERACT decision-making process is engaging the OMERACT community in evaluating the results of the instrument selection process and seeking a vote of support from that community regarding the rigor and conclusions of that process. When evidence about an instrument is gathered, and a decision is made as to the level of endorsement the working group thinks it should receive, the group will bring this to the OMERACT Technical Advisory Group for review. If the evidence is deemed to be of sufficient quality, the group may have an opportunity to present its findings at a full plenary session, called a workshop, during a face-to-face OMERACT biennial meeting. Seventy percent agreement by the OMERACT community (voting at that session) will be considered support for the endorsement.

In addition to the guidance in Instrument Selection chapter The OMERACT Handbook[10], the OMERACT Master Checklist and Workbook for Instrument Selection have been developed to help working groups keep track of their progress and to ensure full and transparent reporting. These resources are available on the OMERACT Website (https://omeract.org/resources). No ethics approval was required for this work because it did not involve human subjects.

RESULTS

Results of the initial application of the OMERACT Filter 2.1 Instrument Selection Algorithm. At OMERACT 2018, a presentation was given in the opening plenary to describe the instrument selection process delineated above, and in The OMERACT handbook. The OMERACT methods for instrument selection figure, known as the "The OMERACT Way," and the OMERACT Filter 2.1 Instrument Selection Algorithm were provided for reference throughout the meeting. The Psoriatic Arthritis Working Group presented 2 instruments for endorsement by the OMERACT community, becoming the first group to move through the Filter 2.1 Instrument Selection process. The first was the 66-joint swollen joint count and 68-joint tender joint count (SJC66/TJC68 joint counts) as instruments to reflect the domain of musculoskeletal disease activity in the peripheral joints. The second was the Psoriatic Arthritis Impact of Disease questionnaire (PsAID12) for the measurement of the core domain psoriatic arthritis-specific health-related quality of life. The final recommendations of the working group were presented at the plenary session, where they highlighted strengths and weaknesses of the 2 candidate instruments. Both the SJC66/TJC68 and PsAID12 achieved consensus (i.e., 70% or greater vote) by the OMERACT community and were the first instruments to be passed through OMERACT Filter 2.1 as fully and provisionally endorsed measures, respectively[22,23].

DISCUSSION

The OMERACT Filter 2.1 revisions address instrument selection within an evolving paradigm of measurement instrument assessment. These methods emphasize the increasing need for an outcome measure's scores to have enough evidence to engender confidence in its use in a particular setting. The process has its foundation in the original OMERACT pillars of truth, discrimination, and feasibility that are still critical requirements for instruments to meet, and adds systematic approaches to gathering, appraising, and synthesizing evidence on the performance of the instrument. The OMERACT Technical Advisory Group will continue to work with OMERACT working groups to operationalize the instrument selection process to ensure we are achieving the goal of transparent, rigorous, evidence-based instrument selection for core outcome measurement sets.

ACKNOWLEDGMENT

The authors thank Dr. Caroline Terwee for her valuable input to the development of the COSMIN-OMERACT good methods checklist. In-kind support from the Knowledge Transfer and Exchange team and the Measurement Sciences group at the Institute for Work & Health is acknowledged for the development of the OMERACT Way graphics and input into the methods used.

REFERENCES

1. Williamson PR, Altman DG, Blazeby JM, Clarke M, Devane D, Gargon E, et al. Developing core outcome sets for clinical trials: issues to consider. Trials 2012;13-132.
2. Tugwell P, Boers M. OMERACT conference on outcome measures in rheumatoid arthritis clinical trials: introduction. J Rheumatol 1993;20:528-30.
3. Kirkham JJ, Clarke M, Williamson PR. A methodological approach for assessing the uptake of core outcome sets using ClinicalTrials.gov: findings from a review of randomised controlled trials of rheumatoid arthritis. BMJ 2017;357:j2262.
4. Schmitt J, Apfelbacher C, Spuls PI, Thomas KS, Simpson EL, Furue M, et al. The harmonizing outcome measures for eczema (home) roadmap: a methodological framework to develop core sets of outcome measurements in dermatology. J Invest Dermatol 2015;135:24-30.
5. Tong A, Manns B, Hemmelgarn B, Wheeler DC, Evangelidis N, Tugwell P, et al; SONG-HD Investigators. Establishing core outcome domains in hemodialysis: report of the Standardized Outcomes in Nephrology-Hemodialysis (SONG-HD) consensus workshop. Am J Kidney Dis 2017;69:97-107.

Personal non-commercial use only. The Journal of Rheumatology Copyright © 2019. All rights reserved.

Downloaded on April 18, 2024 from www.jrheum.org

6. 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.
7. Maxwell LJ, Beaton DE, Shea BJ, Wells GA, Boers M, Grosskleg S, et al. Core domain set selection according to OMERACT Filter 2.1: the OMERACT methodology. J Rheumatol 2019;46:1014-20.
8. Boers M, Beaton DE, Shea BJ, Maxwell LJ, Bartlett SJ, Bingham III CO, et al. OMERACT Filter 2.1: elaboration of the conceptual framework for outcome measurement in health intervention studies. J Rheumatol 2019;46:1021-7.
9. Boers M, Brooks P, Strand CV, Tugwell P. The OMERACT Filter for outcome measures in rheumatology [editorial]. J Rheumatol 1998;25:198–9.
10. Boers M, Kirwan JR, Tugwell P, Beaton D, Bingham CO III, Conaghan PG, et al. The OMERACT Handbook. [Internet. Accessed August 8, 2018.] Available from: https://omeract.org/resources
11. Page MJ, Huang H, Verhagen AP, Gagnier JJ, Buchbinder R. Outcome reporting in randomized trials for shoulder disorders: literature review to inform the development of a core outcome set. Arthritis Care Res 2018;70:252-9.
12. Page MJ, McKenzie JE, Green SE, Beaton DE, Jain NB, Lenza M, et al. Core domain and outcome measurement sets for shoulder pain trials are needed: Systematic review of physical therapy trials. J Clin Epidemiol 2015;68:1270-81.
13. COSMIN. COSMIN database of systematic reviews. [Internet. Accessed March 6, 2019.] Available from: https://www.cosmin.nl/tools/database-systematic-reviews/
14. McHorney CA. Health status assessment methods for adults: past accomplishments and future challenges. Annu Rev Public Health 1999;20:309-35.
15. Collins D. Pretesting survey instruments: An overview of cognitive methods. Qual Life Res 2003;12:229-38.
16. Auger C, Demers L, Desrosiers J, Giroux F, Ska B, Wolfson C. Applicability of a toolkit for geriatric rehabilitation outcomes. Disabil Rehabil 2007;29:97-109.
17. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. Qual Life Res 2010;19:539-49.
18. Terwee CB, Jansma EP, Riphagen II, de Vet HC. Development of a methodological PubMed search filter for finding studies on measurement properties of measurement instruments. Qual Life Res 2009;18:1115-23.
19. Slavin RE. Best evidence synthesis: An intelligent alternative to meta-analysis. J Clin Epidemiol 1995;48:9-18.
20. 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.
21. 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.
22. Orbai AM, Holland R, Leung YY, Tillet W, Goel N, Christensen R, et al. PSAID12 provisionally endorsed at OMERACT 2018 as core outcome measure to assess psoriatic arthritis-specific health-related quality of life in clinical trials. J Rheumatol 2019;46:990-6.
23. Duarte-García A, Leung YY, Coates LC, Beaton D, Christensen R, Craig ET, et al. Endorsement of the 66/68 joint count for the measurement of musculoskeletal disease activity: OMERACT 2018 psoriatic arthritis workshop report. J Rheumatol 2019;46:996-1005.

Personal non-commercial use only. The Journal of Rheumatology Copyright © 2019. All rights reserved.

Downloaded on April 18, 2024 from www.jrheum.org