Examining the Similarities and Differences of OMERACT Core Sets Using the ICF: First Step Towards an Improved Domain Specification and Development of an Item Pool to Measure Functioning and Health

REUBEN ESCORPIZO, MAARTEN BOERS, GEROLD STUCKI, and ANNELIES BOONEN

ABSTRACT. Objective. To contribute to the discussion on a common approach for domain selection in the Outcomes in Rheumatology Clinical Trials (OMERACT) process. First, this article reports on the consistency in the selection and names of the domains of the current OMERACT core set, and next on the comparability of the specifications of concepts that are relevant within the domains. For this purpose, a convenience sample of 4 OMERACT core sets was used: rheumatoid arthritis (RA), psoriatic arthritis (PsA), longitudinal observational studies (LOS) in rheumatology, and ankylosing spondylitis (AS). Domains from the different core sets were compared directly. To be able to compare the specific content of the domains, the concepts contained in the questionnaires that were considered or proposed to measure the domains were identified and linked to the category of the International Classification of Functioning, Disability, and Health (ICF) that best fit that construct. Large differences in the domains, and lack of domain definitions, were noted among the 4 OMERACT core sets. When comparing the concepts in the questionnaires that represent the domains, core sets differed also in the number and type of constructs that were addressed within each of the domains. Especially for the specification of the concepts within the domains Discomfort and Disability, the ICF proved to be useful as external reference to classify the different constructs. Our exercise suggests that the OMERACT process could benefit from a standardized approach to select, define, and specify domains, and demonstrated that the ICF is useful for further classification of the more specific concepts of “what to measure” within the domains. A clear definition and classification of domains and their specification can be useful as a starting point to build a pool of items that could then be used to develop new instruments to assess functioning and health for rheumatological conditions. (J Rheumatol 2011;38:1739–44; doi:10.3899/jrheum.110395)

Key Indexing Terms:
PHYSICAL FUNCTION OUTCOME DOMAINS DISABILITY
INTERNATIONAL CLASSIFICATION OF FUNCTIONING, DISABILITY, AND HEALTH

Since its foundation, Outcome Measures in Rheumatology Clinical Trials (OMERACT) has proposed several core sets for outcome measurement in specific rheumatologic diseases, and other core sets are being prepared. OMERACT core sets refer to the minimum number of domains and instruments that are vital to describe outcomes in clinical studies or clinical practice. While “domains” refer to “what” should be measured, instruments specify “how” those domains should be measured. The typical process is that, first, the relevant domains are selected while, second, instruments are identified that measure or assess these domains. Domains are selected by a nominal group consensus process, usually preceded by a Delphi exercise1. Instruments should satisfy the OMERACT filter of validity, and candidate instruments are usually retrieved after systematic review of the literature. For domain selection, OMERACT has suggested consideration of 5 main

---

From the Department of Health Sciences and Health Policy, University of Lucerne, Lucerne; Swiss Paraplegic Research, Nottwil; and ICF Research Branch of WHO Collaborating Centre for the Family of International Classifications in German, Nottwil, Switzerland; Department of Epidemiology and Biostatistics, VU University Medical Center, Amsterdam, The Netherlands; Division of Rheumatology, Maastricht University Medical Center, Maastricht; and Caphri Research Institute, Maastricht, The Netherlands.

R. Escorpizo, PT, DPT, MSc, Research Scientist, Department of Health Sciences and Health Policy, University of Lucerne; Swiss Paraplegic Research; and ICF Research Branch of WHO Collaborating Centre for the Family of International Classifications in German; G. Stucki, MD, MS, Professor, Chair, Department of Health Sciences and Health Policy, University of Lucerne and Swiss Paraplegic Research; Director, ICF Research Branch of WHO Collaborating Centre for the Family of International Classifications in German; M. Boers, MD, PhD, Professor of Clinical Epidemiology, Department of Epidemiology and Biostatistics, VU University Medical Center; A. Boonen, MD, PhD, Department of Internal Medicine, Division of Rheumatology, Maastricht University Medical Center, and Caphri Research Institute.

Address correspondence to Dr. Boonen; E-mail: a.boonen@mumc.nl

---

1739

Personal non-commercial use only. The Journal of Rheumatology Copyright © 2011. All rights reserved.
dimensions (the five “D’s”) of outcome that were proposed by Fries, et al, representing Discomfort, Disability, Damage, Drug toxicity, and Dollars. New insights in clinimetric research identified advantages of improving the domain selection by specifying the large domains into the most specific units of “what to measure.” The main reason to revisit domain definition and content specification is the increasing insight that the instruments used at present, although proven useful, have some shortcomings than can be addressed by new methods. Existing instruments often contain items that cover different dimensions (or domains), and more importantly, the individual items often address several concepts, which can result in disordered responses. The first step in instrument development is to again answer the question, “what should be measured.”

Although several approaches for domain definition and domain specification have been developed, the World Health Organization endorsed the International Classification of Functioning, Disability and Health (ICF) as the universal framework and classification system. The ICF framework supports the biopsychosocial model of functioning and health by recognizing the influence of contextual factors (environmental and personal factors) on functioning and disability. In addition to the framework, the ICF also proposes a classification system and definition of functioning by means of the so-called ICF categories. A total of 1454 categories in the ICF belong to one of the ICF components, namely, Body functions (e.g., pain, energy, sleep, emotion, muscle power); Body structures (e.g., hip joint, ligaments, cervical vertebral column); Activities and participation (such as reaching, walking, self-care, housework, remunerative employment); and Environmental factors (such as drugs, health services, attitudes of family members). Within each component, the categories are grouped within chapters and are further specified as 2nd, 3rd, and 4th level categories, in which 3rd and 4th level are specifications of the 2nd level categories. As such, ICF categories are units of functioning within health-related domains. Applying this to the OMERACT domains, the ICF categories can be seen as specifications of some of the OMERACT health-related domains, and refine “what can be measured” when addressing functioning and health. Although Personal Factors are included as a component in the ICF framework, they are as yet not defined by a series of ICF categories.

Terminology to understand the relation between OMERACT domains and ICF classification:

**OMERACT domain**
- vaguely described concept referring to relevant areas of outcome in rheumatology outcome studies. Each area likely consists of a number of (hierarchically ordered) subdomains that specify the domain further

**ICF categories**
- the units that allow us to specify and classify health-related domains (areas)

- ICF categories are assigned an alphanumeric code, a title, and an operational definition. Categories can be target (outcome) or explanatory variables
- each ICF category belongs to one of the ICF components: Body Functions, Body Structures, Activities and Participation, and Environmental factors
- ICF categories are structured hierarchically, lower levels specifying higher levels

**Concept**
- the idea or abstract principle relating to a specific subject (in the context of this article, ‘how to measure’ the specific concept of the domains (the idea or principle it represents) is often represented in the content of items from questions or questionnaires

The objective of this article is to find out to what extent the domains and concepts (“what to measure”) within the OMERACT core sets for musculoskeletal disease (i.e., arthritis) can be compared, and whether the ICF can be useful in the process of specification of OMERACT domains by providing an external and universal framework of all aspects of functioning.

**METHODS**

Four OMERACT core sets, psoriatic arthritis (PsA), rheumatoid arthritis (RA), ankylosing spondylitis (AS), and longitudinal observational studies (LOS) in rheumatology, were conveniently selected and compared. First, the selected domains for each were compared directly. Next, concepts contained in the instruments that were considered or selected to measure the OMERACT domains were linked to the ICF categories, which served as the external reference for comparison. Linking refers to the procedure that follows established rules and aims to match a domain or concept to the most precise ICF category that represents that concept or domain. Comparability of OMERACT core sets is discussed in view of the difficulties encountered in comparing the domains and concepts.

**RESULTS**

Table 1 presents the comparison of the “core domains” of the OMERACT core sets for PsA, RA, AS, and LOS in rheumatology. First, core sets differed in the way in which proposed domains were organized. The RA core set selected one set of domains that represent the minimal to be included in every clinical study, with one additional domain (radiology) that should be included in studies lasting one year or more. The PsA core set distinguished a minimal number of domains to be included in every clinical study, a number of optional domains that can be of further interest, and a number of domains that need further study before including them. The core set for LOS in rheumatology first selected subdomains that were then grouped into broader domains, some of which are optional. Finally, the AS core set distinguished 3 settings symptom-modifying antirheumatic drugs (or SMARD) or physiotherapy, disease-modifying antirheumatic drugs, and clinical record-keeping — with increasing number of (obligatory) core domains.

Second, a difference in the number of domains or subdomains was seen. For the core set of PsA, 6 core domains
were selected, 8 optional domains, and 5 domains for the research agenda; for RA, 7 (8 for studies of more than 1 year); for LOS, 11 subdomains within 5 broader domains (2 of which are optional); and for AS a total of 9 domains for the 3 settings together.

The names of the domains were surprisingly difficult to compare and could not easily be classified under the proposed “five D” structure of domains. Specifically, the domains in the core set for LOS were difficult to compare with other core sets: the starting point of the core set for LOS considered not only the “five D’s” but also included the “impairment→disability→handicap model” or the International Classification of Impairments, Disabilities and Handicaps model of the World Health Organization (ICIDH), which preceded the ICF and aimed to make a distinction between process versus damage. This resulted in names and subdomains that differed substantially from other core sets. Further, the domain “health-related quality of life” (HRQOL) posed some challenges. In the LOS, HRQOL was synonymous with health state, and together with Disease process and Damage, part of the domain “health.” HRQOL (or health status) domains included symptoms, physical function, and psychosocial function. In the PsA core set, HRQOL was selected as a domain in addition to the domains pain, physical function, global assessments, and participation. When taking into account the specifications of the subdomains that were proposed for the (sub)-domains of the LOS core set, then pain, physical function, patient global assessment of disease activity, and joint assessment were domains common to all 4 core sets. Damage (including radiographic damage) and acute-phase reactants were presented in all core sets except for the PsA core sets, where it is part of the research agenda. Fatigue was a separate domain in the AS core set, was mentioned among the domain “symptoms” of the LOS core set, was not included in the RA core set, and was on the research agenda for the PsA Core Set. The core set for LOS was the only one to include psychosocial function as a specific domain within HRQOL. The core set for psoriatic arthritis was the only one to include the domain “skin.” When considering the “five D” model, it is clear the OMERACT core sets emphasize Discomfort, Disability, and Damage. The LOS was the only one to include Drug toxicity and Death. Costs were included only as an optional domain in the core set for LOS.

When trying to compare the specific concepts of “what to measure” contained in the domains of the different core sets, it was assumed that the content of domains is specified by the concepts included in the items of the selected instruments. However, only for the core set of AS were measures formally selected by consensus. Notwithstanding, the PsA core set and the core set for RA, LOS, respectively, considered or proposed measures for domains. For the purpose of the present study, these “considered or proposed” measures were linked to the ICF categories. A separate table that represents the full result of this linking exercise is available from the authors on request.

Linking to the ICF revealed that the majority of concepts

Table 1. Comparison of the domains selected for the OMERACT core sets for psoriatic arthritis, rheumatoid arthritis, longitudinal observational studies, and ankylosing spondylitis.

<table>
<thead>
<tr>
<th>Common Domains</th>
<th>Psoriatic Arthritis</th>
<th>RA Clinical Trials</th>
<th>Longitudinal Studies in Rheumatology</th>
<th>Ankylosing Spondylitis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pain</td>
<td>Pain</td>
<td>Pain</td>
<td>Health status: “symptoms”</td>
<td>Pain&lt;sup&gt;b,c&lt;/sup&gt;</td>
</tr>
<tr>
<td>Stiffness</td>
<td>Physical function</td>
<td>Physical disability</td>
<td>Health status: “Physical function”</td>
<td>Physical function&lt;sup&gt;a,b,c&lt;/sup&gt;</td>
</tr>
<tr>
<td>Fatigue</td>
<td>HRQOL</td>
<td>HRQOL</td>
<td>Health status: “HRQOL”</td>
<td></td>
</tr>
<tr>
<td>Physical function</td>
<td>Participation</td>
<td>Patient global assessment</td>
<td>Disease process: “globals”</td>
<td>Patient global activity&lt;sup&gt;a,b,c&lt;/sup&gt;</td>
</tr>
<tr>
<td>HRQOL</td>
<td>Patient global</td>
<td>Patient global assessment</td>
<td>Disease process: “globals”</td>
<td></td>
</tr>
<tr>
<td>Physician global Joints</td>
<td>Peripheral joint activity</td>
<td>Swollen joints</td>
<td>Disease process: “Joint tenderness/swelling”</td>
<td></td>
</tr>
<tr>
<td>Psychosocial function</td>
<td>Acute-phase reactant</td>
<td>Acute-phase reactants</td>
<td>Disease process: acute-phase reactant</td>
<td>Acute-phase reactant&lt;sup&gt;b,c&lt;/sup&gt;</td>
</tr>
<tr>
<td>Mobility</td>
<td>Radiograph</td>
<td>Radiographs of joints (in studies ≥ 1 year duration)</td>
<td>Damage: Radiographic damage</td>
<td>Spinal mobility&lt;sup&gt;a,b,c&lt;/sup&gt;</td>
</tr>
<tr>
<td>Deformity</td>
<td></td>
<td></td>
<td>Damage: deformity</td>
<td></td>
</tr>
<tr>
<td>Organ damage</td>
<td></td>
<td></td>
<td>Damage: “organ damage”</td>
<td></td>
</tr>
<tr>
<td>Surgery</td>
<td></td>
<td></td>
<td>Damage: surgery</td>
<td></td>
</tr>
<tr>
<td>Skin</td>
<td>Skin assessment</td>
<td></td>
<td>Mortality</td>
<td></td>
</tr>
<tr>
<td>Mortality</td>
<td></td>
<td></td>
<td>Toxicity/adverse reactions</td>
<td></td>
</tr>
<tr>
<td>Toxicity</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup> In core set for symptom-modifying antirheumatic drugs (SMARD) and physiotherapy trials; <sup>b</sup> in core set for disease clinical record-keeping; <sup>c</sup> in core set for disease controlling antirheumatic drugs. HRQOL: health-related quality of life.
addressed in the OMERACT domains were pertinent to pain functions, movement/mobility, self-care, and participation. In addition, categories referring to energy and drive and emotional functions were addressed. To illustrate the advantage of linking, Table 2 presents the comparisons of the constructs included in the physical component of the Medical Outcome Study Short-Form 36 (SF-36)14 (considered in the domain “physical function” of the OMERACT core set for PsA), the Health Assessment Questionnaire (HAQ)2 (considered for the domain “physical function” in the RA and LOS core set, and the Bath AS Functional Index (BASFI) (selected for the domain “function” in the OMERACT core sets for AS). Consistent with the development for RA, the HAQ included more activities of the hands and upper limbs, and the BASFI included ICF categories referring to “maintaining and changing body positions,” which is more specific to patients with back problems. The SF-36 also contained categories on mental function and participation and thus exceeds the “traditional” domain of “physical function.”

**DISCUSSION**

The comparison of the OMERACT domains and more specific concepts represented within the domains across 4 selected OMERACT core sets proved useful. We learned that “what to measure” is not consistently defined across core sets and lacks specifications. There are 2 findings that we believe are relevant for further discussion within OMERACT.

First, within OMERACT core sets, the domains are grouped and named following different approaches. Although OMERACT suggested the “concept of five dimensions” as the starting point, this is not consistently reflected in the OMERACT core sets. Further, in the core set for LOS, subdomains...
are grouped in core domains comprising “health,” “process,” “damage,” “toxicity,” and “death.” Specifically, the domains “process” and “damage” represent a different construct, since it is a human interpretation that goes beyond the experience at the level of the patient. On the other hand, in the core set for AS, domains were grouped according to the setting of application [SMARD studies, disease-controlling antirheumatic therapy (DC-ART) studies, or clinical record-keeping]. Finally, in the core set for PsA and LOS, a number of optional domains are selected in addition to the “core domains.” As a consequence, in the core set for PsA, acute-phase reactant and radiographic damage are not in the core, while they are core domains for RA (radiographs in studies longer than 1 year), for AS (in DC-ART studies and clinical record-keeping) and for LOS (radiographs are a subdomain in the domain “damage”). Specifically challenging was the domain “quality of life” (QOL). While QOL is part of the core set for PsA and LOS, the domain covers different content. In the core set for LOS, HRQOL is the umbrella for “health” and includes symptoms as well as physical and psychological functioning. In the core set for PsA, HRQOL is a domain separate from pain and physical function.

Second, to be able to compare the specific concepts that are represented in the domains using the ICF as external reference, it was aimed to use the constructs in the items of selected questionnaires that need to represent the OMERACT domains. Remarkably, only the core set for AS agreed upon core instruments. Therefore, it was decided to also include in the comparison the instruments that were considered or recommended in the development of the OMERACT core sets. Comparison of the concepts in the items using the ICF again proved to be useful, especially for the domain “physical function.” The ICF allowed us to gain insight into which aspects of functioning are “typical and relevant” across all diseases versus those for specific diseases.

Historically, the choice of OMERACT to define “what to measure” in broad domains was a logical choice. First, a detailed framework that represented all possible aspects of health was not available. Second, the OMERACT process wanted, for domain specification, to make use of existing instruments that had proven usefulness and had been extensively validated. The need addressed by OMERACT was primarily to agree on selected instruments, in order to improve harmonization (uniformity) in the world of outcome measurement.

The question therefore remains whether there is a need to revisit the question “what to measure.” In the new structure of OMERACT, the selection and further specification of domains are given careful attention15. With improved insight into the “concept of health” and the emergence of new approaches to development of instruments, it appears that reexamining the domains and domain specification makes sense. Despite their proven usefulness for progress in rheumatology, the existing measures also have some drawbacks. Item response theory has already been applied to the HAQ (included in the core set for RA and considered for PsA) and BASFI (included in the core set for AS) and showed that although instruments are reasonably unidimensional, neither of them (HAQ or BASFI) are true interval measures and include items that show differential functioning and non-ordered responses6,17,18,19. Moreover, they lack items that help to distinguish between patients with lower levels of disability (i.e., floor effect). The ICF offers a classification by the WHO that is necessary to describe functioning and health. The ICF framework can be used as the external standard and therefore the starting point of defining “what to measure” for either one of the rheumatological conditions or within a particular setting. Moreover, the ICF could facilitate the management of a pool of items that are clinically valid and could be used in the development of new instruments. For AS, effort is going towards the development of a pool of items covering the categories of the ICF core set for AS20,21. In the area of the development of item pools, the Patient Reported Outcomes Measurement System (PROMIS) is sophisticated and useful22,23,24, and it would be a beneficial initiative for OMERACT to look at the active interface between the ICF and PROMIS15.

Other fertile ground within the OMERACT community is to consider the application of the ICF by way of the so-called ICF core sets. Different from OMERACT core sets, ICF core sets refer to a short list of ICF categories that are relevant to a health condition or health-related event. These are selections of ICF categories that are necessary to classify functioning of an individual using extensive and multiple methodologies. They encompass the perspective of patients, existing research, and health professionals. Such core sets are available for several rheumatologic or musculoskeletal conditions — RA25, AS21, osteoporosis26, chronic low back pain27, chronic widespread pain28, and osteoarthritis29. These core sets could be the starting point to specify OMERACT domains.

In addition to the comprehensive core sets, there are current initiatives to develop a statistically derived or combined statistical experts-derived Brief ICF Core Set based on the comprehensive core sets. Although the ICF (Core Sets) can be useful as external reference, we need to recognize that the ICF has limitations, such as the lack of universally agreed operationalization of its categories, and the lack of classification of personal factors. However, these limitations do not outweigh the advantage of having a common conceptual language and classification to measure functioning and disability domains across diseases. The ICF initiative within OMERACT is a starting point to propose for each “domain” the best (set of) concepts (ICF categories) to specify the domain.

In conclusion, we learned, by doing a comparison of some OMERACT core sets, that domain conceptualization and definition vary across diseases and domain specification is insufficient, if not lacking. A common framework to assess outcome and specification of domains could further improve outcome measurement as already addressed in the OMERACT report on domain selection15. The ICF proved useful in spec-
ifying the domains referring to physical function in the OMERACT core sets, but can also provide specification for domains that are not part of OMERACT core domains but are relevant to outcomes assessment.

REFERENCES