THE OMERACT HANDBOOK

FOR ESTABLISHING AND IMPLEMENTING CORE OUTCOMES IN CLINICAL TRIALS ACROSS THE SPECTRUM OF RHEUMATOLOGIC CONDITIONS

VERSION 2.1 | June 21st, 2021

Striving to improve endpoint outcome measurement through a data driven, iterative consensus process involving relevant stakeholder groups.

Contributors: Dorcas Beaton, Lara Maxwell, Shawna Grosskleg, Bev Shea, Peter Tugwell, Clifton O. Bingham III, Philip G. Conaghan, Maria-Antonietta D'Agostino, Catherine Hofstetter, Lyn March, Lee S. Simon, Jasvinder A Singh, Vibeke Strand, George Wells. With thanks to the many OMERACT participants, patients and other colleagues who contributed to the content of this Handbooks

Contact: OMERACT Secretariat 43 Bruyère St. Annex E Ottawa, ON K1N 5C7 Canada ▪ admin@omeract.org

Table of Contents

CHAPTER 4: DEVELOPING CORE DOMAIN SETS

INTRODUCTION

An OMERACT Core Outcome Set is made up of two important and sequential components: decisions about ***what*** to measure (Core Domain Set) and then decisions about ***how*** to measure each of the chosen domains (Core Outcome Measurement Set). The word *Domain* is closely linked to, or even equates with, the words 'concept', 'attribute' and 'construct' used in the literature (for example 'domain' or 'concept' used in the Cochrane Handbook). At OMERACT we use the term Domain.

Working Groups will use the procedures described here in Chapter 4 to gain endorsement on a Core Domain Set (what to measure) and then further work must be done to ensure that each Core Domain is addressed by at least one applicable instrument. The procedures for instrument selection are described in Chapter 5.

How to choose domains the OMERACT way

As with all our activities, it is based on evidence and consensus and is guided by the Spirit of OMERACT in its conduct (see Chapter 1). The steps of the Filter 2.2 Domain Selection process are summarized in the How to choose domains the OMERACT Way flowchart to select a core domain set (Figure 4.1).

This process is the basis for the OMERACT Master Checklist for Developing Core Domain Sets (table 1) and accompanying Workbook (Appendix A) which OMERACT Working Groups will follow as they move through the process of identifying their domains and creating the OMERACT Onion.

In this process, domains will be selected and defined to reflect the domains that are important to OMERACT stakeholders (including the 7 P's). They will then be prioritized for their level of importance for clinical trials, and this is where the working group will be placing the domains into an OMERACT Onion (Figure 4.4) for endorsement by the OMERACT community.

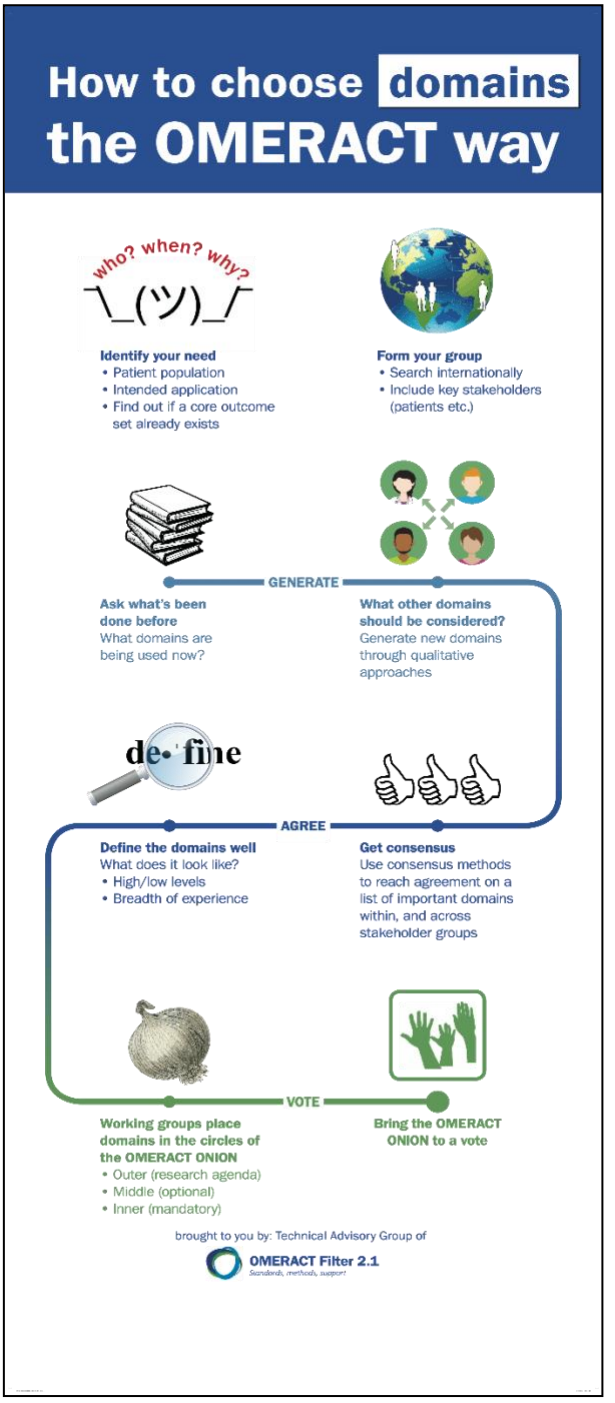


Figure 4. 1 The OMERACT Way flowchart to select a Core Do-main Set

Generating, Agreeing and Voting

Once the need has been identified and the Working Group has been formed (Chapter 2: OMERACT Working Groups) the process could be described in three phases – generating and defining domains, agreeing on the domains and their prioritization, and finally assembling and voting on the OMERACT Onion (Figure 4.2).

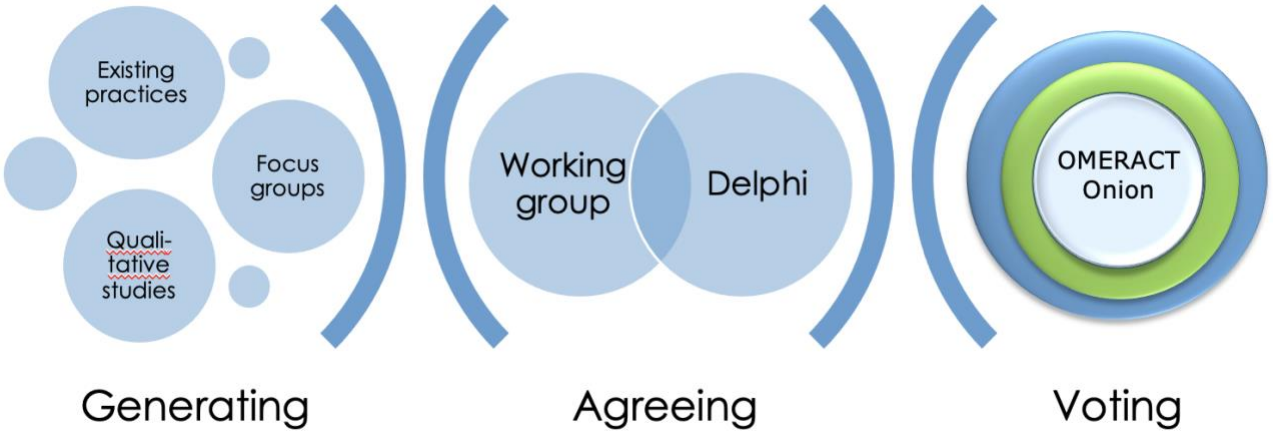


Figure 4. 2 The three phases of OMERACT's domain selection process: (i) generating candidate domains, (ii) agreeing on the domains and their prioritization using consensus methods, and (iii) voting on the OMERACT Onion by the OMERACT

When to begin thinking about Domains and their relevant definitions

Domain selection begins with a review of the key areas of health that need to be represented in the Core Domain Set. Boers et al. first described these as "s" and used a conceptual framework, OMERACT Filter 2.0, based on the International Classification of Functioning to define areas we are asking groups to consider when selecting domains (1). This framework was revised to OMERACT Filter 2.1 to address challenges in the framework application caused by unclear or ambiguous wording and terms, and incompletely developed concepts (2). This current iteration (Filter 2.2) corresponds to adjustments made to defining the domains and the addition of the new domain definition template as a tool to create those definitions. The framework is designed to help the development of a core set by encouraging us to think about what is considered a full breadth of domains when developing a core domain set. It is briefly described below, and readers are

Review of the meaning of the Core Areas in the OMERACT Framework.

Manifestations/abnormalities: Essential to assess whether the effect of the intervention specifically targets the pathophysiology of the health condition.

Manifestations corresponds with ICF body functions and structures; it includes psychosocial manifestations. Example Domains include organ function (e.g., renal function), reversible manifestations (including modifiable risk factors and actual manifestations of ill health), and irreversible manifestations (including unmodifiable risk factors and damage). This Area also encompasses all biomarkers and surrogate outcomes. In trials primarily focused on Impact, the core set will describe the minimum to be measured under Manifestations.

Resource Use: describes the economic impact of health conditions both on society and on the individual. Both the presence of a health condition and its treatment incur resource use.

Life Impact: Under Life Impact, OMERACT strongly suggests that core set developers consider both the domains of the International Classification of Functioning, Disability and Health (ICF) and domains within the concept of health-related quality of life, for example, as elaborated by Wilson and Cleary's model. In trials primarily focused on understanding a mechanism of action or proof of concept, the core set will describe the minimum to be measured under Life Impact

Death: Possible specifications include generic and disease-specific, that is, all cause vs. disease-specific mortality; and intervention-specific (e.g., death owing to surgery or transplantation). In conditions where death rarely occurs during a trial, this area could be covered in the core set by requiring a simple report of any deaths (or their lack), which is already a standard requirement in current guidelines.

encouraged to read the details in the framework elaboration paper (2).

The relationship between these areas is shown in Figure 4.3 & 4.4. The framework also includes variables known as 'contextual factors', which need to be measured to fully understand study results. Working Groups will use the procedures described here in Chapter 4 to identify the Core Domain Set and then further work must be done to ensure that each Core Domain is addressed by at least one applicable instrument. The procedures for instrument selection are described in Chapter 5.

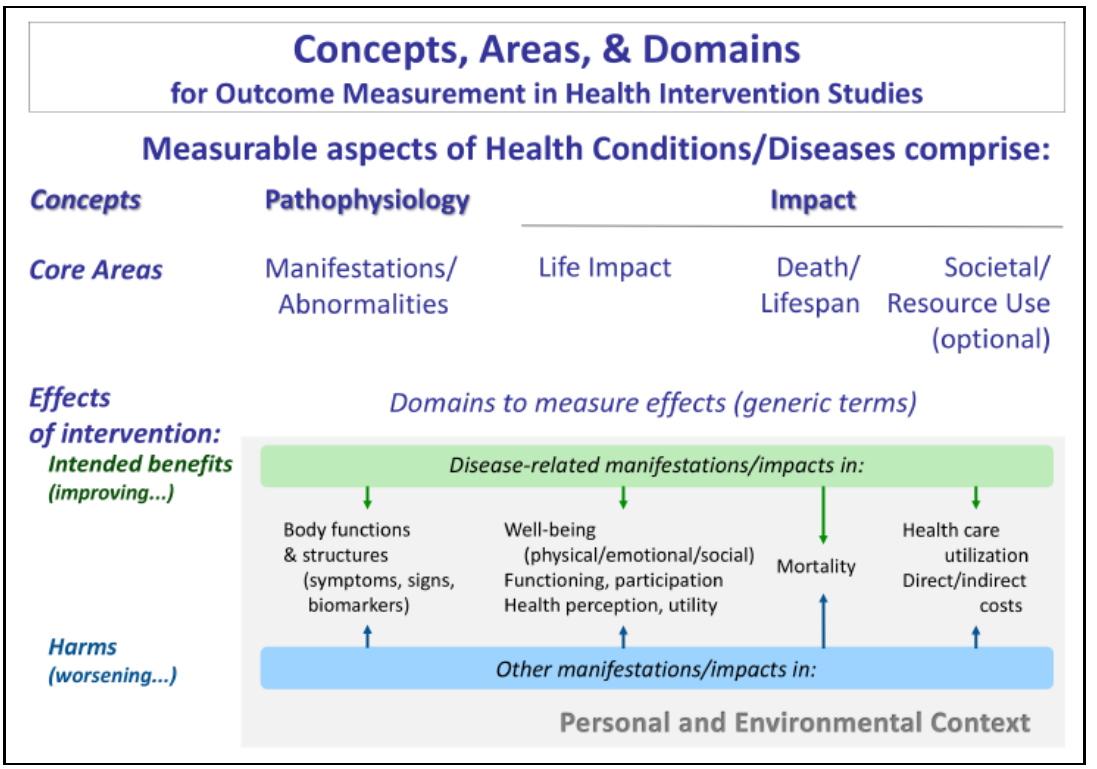


Figure 4. 2 Revised OMERACT Filter 2.1 Framework. Emphasis is on clarification of the Core Areas that need to be considered for representation in a Core Outcome Set, and the consideration of both intended effects and harms or costs of an intervention.

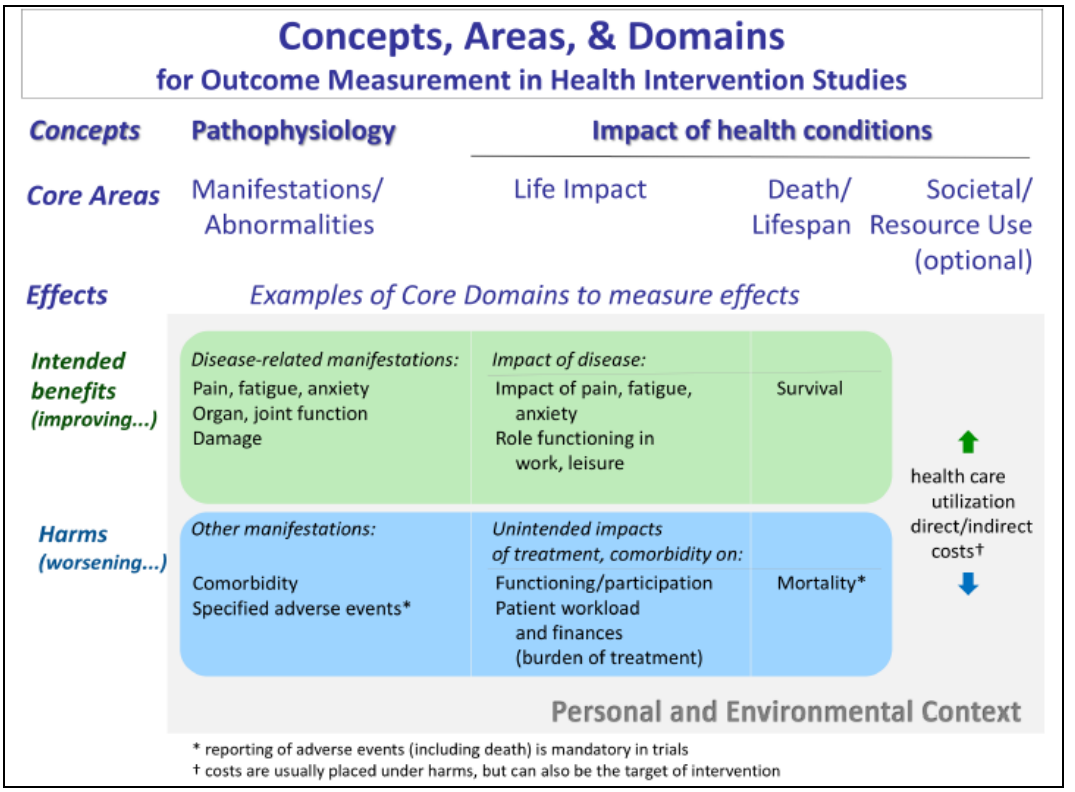


Figure 4. 3 Revised OMERACT Filter 2.1 Framework with examples. Societal and resources use are outside of the shaded boxes as the same indicator such as cost would represent both intended effects and harms.

Throughout this chapter you will see a lot of discussion about "consensus", one of the main tenets of OMERACT's work. Consensus is needed for your working group to agree on the domains, and where each domain will be placed in the Onion. Consensus is needed when you seek the agreement of the whole OMERACT community. Consensus is not a majority vote; it is getting to a decision that everyone can agree to or live with. It means thoughtfully engaging people in your decision-making processes and the content of your work so that they feel they can agree with or live with your recommendations. Often, we want to make sure that consensus has been achieved across patient and other OMERACT participants, not to highlight their differences, but to ensure our patient research partners (who make up a smaller proportion of our membership) are heard.

The OMERACT Onion

The OMERACT Onion is a visual representation of the domains that are considered mandatory, optional, or worthy of further research. These are transformed into the layers of an onion that has at its core the inner circle of mandatory domains. This is split into two layers: domains that are mandatory in **all trials** and domains that are mandatory in **specific circumstances** (for example, if there is eye involvement, include these domains). These mandatory domains should be kept to only critically important and essential domains to not overburden clinical trials. In some cases, a domain might be important for certain types of trials, but not necessarily all of them. These would go into the middle circle of the onion listing domains that are important but optional. In other cases, the domain might need a bit more unpacking and research before your working group is ready to recommend a specific location for it, so it would be parked in the outer circle, the research agenda domains, where more information is needed. We will cover the placement of domains in the OMERACT Onion in further details in section 7 below

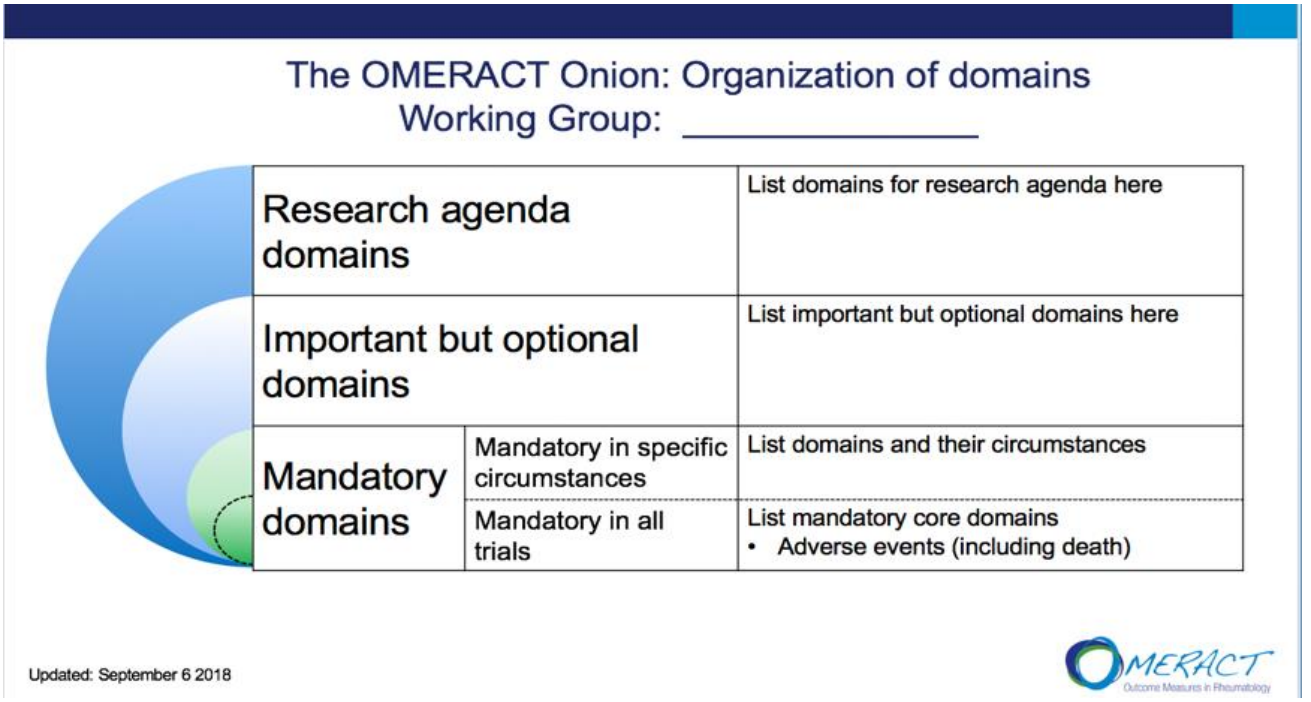


Figure 4. 4 the OMERACT onion. The onion shows the results of domain generation, agreement, and voting. The inner circle of mandatory domains is the "core domain set" fielded in each trial

Remember: Working Groups **must** have an OMERACT-endorsed Core Domain Set as described in this chapter BEFORE they can move to selecting instrument(s)

OMERACT MASTER CHECKLIST FOR DEVELOPING OR UPDATING CORE DOMAIN SETS

The OMERACT Master Checklist for Developing Core Domain Sets is a tool for OMERACT Working Groups to use as they move through the process of identifying their domains and creating the OMERACT Onion.

#	OMERACT Core Domain Set Checklist Item	Mark when complete
Core Domain Set selection		
Assembly of Working Group		
1	Assemble working group	o
Develop Methods Protocol		
2	Describe PICOC (Population, Intervention, Control, Outcome, Context)	o
3	Protocol development	o
4	**Deliverable:** Submission of protocol to Technical Advisory Group based on OMERACT Core Domain Workbook	o
5	Review and approval of protocol for Core Domain selection by Technical Advisory Group	o
Generating		
6	Generate candidate domains covering each Core Area	o
Agreeing		
7	Prioritization of candidate domains through DelphiManager modified for OMERACT	o
8	Formulation of Draft Core Domains and Definitions	o
9	Formulation of Core Contextual Factors	o
10	Working Group agrees on, finalizes, & submits Draft Core Domain Set to Technical Advisory Group	o
Voting		
11	Result of final vote by full OMERACT community on Core Domain Set	o

A Domain Selection Workbook is available in the Appendix at the end of this chapter. This workbook has been developed to help Working Groups keep track of their progress as they move through the Filter 2.2 process. Using the workbook will make it easier for Working Groups to report on their progress to the OMERACT Executive and

OMERACT Technical Advisory Group and ensure that all necessary steps are fulfilled. As well, it will ensure full and transparent reporting according to the Core Outcome Set-Standards for Reporting (COS-STAR) statement (4).

ASSEMBLY OF WORKING GROUP

1. Assemble working group

OMERACT has established a philosophy around the communication and engagement of members entitled the 'Spirit of OMERACT'. This is outlined in detail in Chapter 1 of this Handbook. Working Groups are expected to foster the Spirit of OMERACT (e.g., collaboration, consensus) in all their work.

Following the guidelines for establishing a working group found in Chapter 2: OMERACT Working Groups, ensure your working group has all the required elements:

1. **International Representation – co-chairs from a minimum of three continents.**
2. **Stakeholder Engagement – particularly ensuring you have patient research partners (PRP), fellows, and other key stakeholders.**
3. **Topic Redundancy – a group that can represent all aspects your topic/disease area well.**

Working group will be asked to complete an online form in the early stages of formation more details on assembly on a Working Group can be found in Chapter 2.

DEVELOP METHODS PROTOCOL

2. Describe PICOC (Population, Intervention, Control, Outcome, Context)

The first step in developing a protocol, or work plan, is to formulate a detailed description of the setting (scope) of the core outcome set. Central to this activity is defining the "PICOC statement" to which the Core Domain Set will apply; that is, the Patients/Population, Intervention, Comparator/Control, Outcome, and Context, with the understanding that the 'O-Outcome' is what will be defined during the project.

A comprehensive explanation on what the core domain set will cover is generated from Working Group discussions and may be modified based on discussions at the Special Interest Group activity. Items that need to be agreed upon include the: health condition(s) (disease or disease group) or population to which the intervention can be applied; type of interventions being compared (for example same or different class of treatments, drugs/biologics, non-pharmacologic, surgery and other interventions); etc. The Working Group also needs to decide whether the core set will apply only to randomized trials (or a subset of trials, e.g., effectiveness trials), or to longitudinal observational studies as well. A Core Outcome Set is only useful if the information it provides is adequate to enable treatment decision making. This means that the context in which the decision-making is expected to happen is also part of the setting (scope) of the Core Outcome Set (for example, information for the public, the patient, the physician, guideline developers, payers, etc.). Note that the word context has a very broad meaning, in contrast to the concept 'contextual factor' that is used in a very specific way within OMERACT.

Component of PICOC	*Description of your criteria for each component*
Population	
Intervention	
Control	
Outcome	*Core Domain Set under development*
Context (Setting)	

Please find an example of a PICOC from the OMERACT Hip & Knee Osteoarthritis Working Group (26)

Component of PICOC	*Description of your criteria for each component*
Population	People with hip and/or knee osteoarthritis
Intervention	Non-surgical interventions (Pharmacological and Non-Pharmacological Interventions)
Control	Non-surgical interventions (Pharmacological and Non-Pharmacological Interventions)
Outcome	*COS under development*
Context (Setting)	Phased 3 or 4 randomised controlled trials and non-randomised controlled trials

3. Protocol Development

Once the PICOC has been developed, the Working Group should check that there is no other core outcome set already in existence in the literature, for example, from professional clinical associations such as ACR, OARSI, EULAR, etc., and check with content experts. Note that the "Core Outcome Measures in Effectiveness Trials" (COMET) Initiative houses a database on their website (www.comet-initiative.org) of completed or in progress core outcome sets. If no overlap is identified, or there is strong justification for developing a new core outcome set, although there is an existing one (e.g., an existing core outcome set was lacking patient participation in its development), then the Working Group can proceed. Other gaps such as an important potential domain(s) or lack of accurate methods in the development of existing core outcome sets may also be identified during Working Group discussions. The OMERACT Core Domain Set Selection Workbook provides templates and guidance for describing your work plan (or 'protocol') for developing the Core Domain Set. It consists of your plans for both generating candidate domains and prioritizing those do-mains completing the protocol for submission to the Technical Advisory Group.

We recommend reviewing the COS-STAD (Core Outcome Set-STAndards for Development) and COS-STAP (Core Outcome Set-STAndardised Protocol Items) guidance documents (3,4) and making your protocol publicly available by entering it into the COMET database or publishing in a journal (e.g., Trials, BMC Med Res Methodol).

4. Deliverable: Submission of protocol to Technical Advisory Group based on OMERACT Core Domain Workbook

Once the Working Group has developed the protocol (checklist items 1, 2 & 3) they are ready to submit this for review by the Technical Advisory Group. This group will review the proposed methods, particularly if they vary from the OMERACT developed methods. The protocol must adequately describe how they will conduct literature searches and qualitative work to generate candidate domains, how they will track domain definitions, how they will agree on the candidate domains, and reach consensus for deciding on the Draft Core Domain Set. The role of the Technical Advisory Group is to critically appraise submitted documentation from OMERACT Working Groups for adherence to Filter 2.2 checklist requirements. Specifically, their role is to verify that the recommended steps and methods are being followed and to identify any potential challenges (i.e., not enough continents involved).

The Generating and Agreeing sections below provide details on the methods needed to help Working Groups draft their methods protocol.

5. Review and approval of protocol for Core Domain Set work by Technical Advisory Group

The Technical Advisory Group (TAG) will review and provide written comments on the methods described in the protocol to the co-chairs of the Working Group. At this stage the TAG will be looking at whether you have engaged a broad enough representation in your group, whether your planned methods on generating and agreeing on domains are appropriate, and whether you have identified ways of sharing your work that will facilitate communication. Once the Technical Advisory Group has approved the protocol in writing, the Working Group can start work on the steps below to develop their Core Domain Set. This may take a couple of iterations back and forth between TAG and the working group. The result will be a strong protocol. At this stage we recommend registering the protocol on the COMET website of core outcome sets (https://www.comet-initiative.org/About/SubmitNewStudy). Once the core set work is complete the COMET database should also be updated.

GENERATING

6. Generate candidate domains covering each Core Area

The first phase of the Core Domain Set development is called generating, where a wide comprehensive set of potential domains is generated. This phase is very inductive, creative, and generative. The wider the points of view, the more likely you are to capture the set of important domains.

There are often two initiatives in this phase:

- A scoping or literature review can be conducted to identify existing domains or previously published qualitative work and
- Qualitative work, including focus groups and/or interviews of appropriate stakeholder groups to identify additional domains. The methods should be followed as outlined in the protocol and in cases where changes were made to the methods, these can be described in the Domain Selection Workbook.

There is a good reason why both literature review and qualitative work is needed. A literature review alone will reinforce what is already being done, even in the absence of a core outcome set. A good example of the importance of both is shown in the work done to identify candidate domains for low back pain. Figure 4.5 shows the domains identified during the literature search in the left-hand column, and those identified from surveys and

focus groups of key stakeholder groups. The qualitative work elicited almost all the domains that were previously used in the literature, save two, however it also elicited many more domains of interest for consideration. These additional domains would have been missed. Those domains in the overlapping area were identified through both means. Qualitative input is important to broaden the scope of domains that would have been otherwise missed if only looking at what is used in the current and past literature

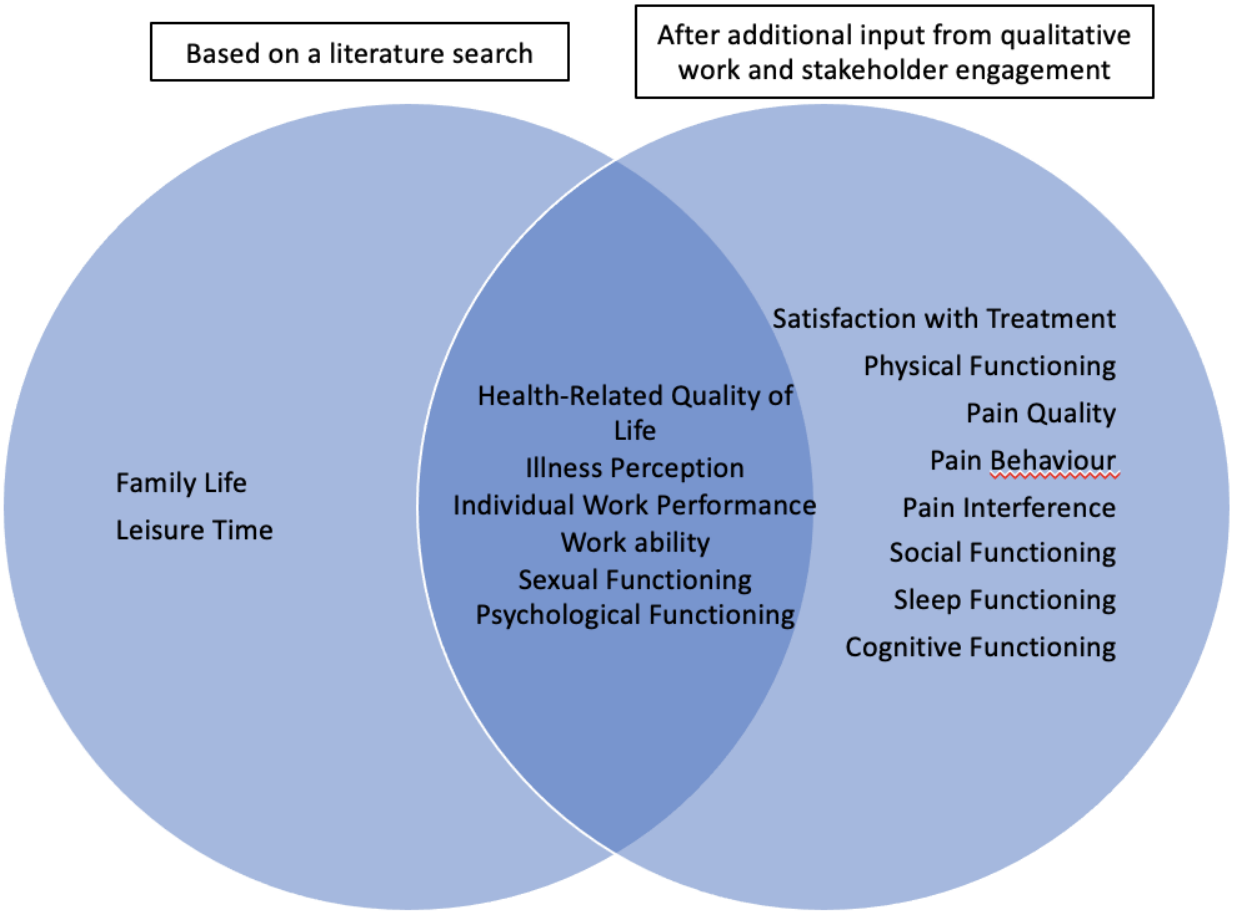


Figure 4. 5 Example of the difference in domains identified from literature search, and from more inductive approaches to generation.

We will now go into more detail of these two phases of generation.

6.1. Scoping review or literature review of domains: what has been measured?

TIP: Make use of a librarian or information specialist when developing your literature search strategy

A review of the literature to see what domains have been used in the past is often a good first step in the generation process. It is up to the working group to decide who will lead this review, but many working groups have found this is a good project for an OMERACT Fellow and it has often led to a publication.

A librarian or information specialist should be consulted when developing the search strategy. Working Groups should note that there are often new domains that are identified and operationalized after focus groups and/or surveys with key stakeholders have been held.

Often groups decide to search all randomized controlled trials for the domains used in trials. Others have extended this to cohort studies as well to have a broader sense of potential domains. Some groups like the Shoulder Working Group found enough published qualitative literature to allow them to do a review of the qualitative literature. This yielded a good range of domains including emotional distress and cognitive dysfunction (6) that may not have been found otherwise, and a lot of descriptive information to help with future definitions. They did a full systematic review.

OMERACT recommends that this review is at the level of a scoping or systematic review rather than a narrative or descriptive literature review. Here is an example of a scoping review for outcomes in another field (7). Either of these would be enough to give a robust sense and answer the broader question of what domains are being captured in the outcome measurement in the field. In general, we feel that a scoping review would be more than adequate for this need. As described below a scoping review is characterized by a comprehensive literature review with explicit methods of how articles were selected. Critical appraisal, which is a pillar of a "systematic" review, is not needed for a scoping review. As the name denotes, a scoping review is aiming to get a broad sweep of a field, rather than a very specific meta-analysis that could be the goal of a systematic review.

	Literature Reviews	Scoping Reviews	Systematic review
QUESTION	Broad	Focused	Focused
SOURCES	Usually, unspecified. possibly biased/	Comprehensive. explicit	Comprehensive, explicit
SELECTION	Unspecified. possibly biased/ All study types/Developed post hoc at study selection stage	Criterion-based. Uniformly applied	Criterion-based; uniformly applied
APPRAISAL	Not needed	Not needed	Quality appraisal of the methods used in the study is conducted
SYNTHESIS	Usually qualitative/ "Charts" data according to key issues, themes, etc.	Generally, not a quantitative meta-analysis conducted	Often is a formal synthesis such as a meta-analysis.

INFERFENCE	Generally, not evidence based	Usually, evidence based	Usually evidence-based

Table 4. 1 Comparison of Literature, Scoping and Systematic Review

During this review, both the domain and its working definition should be extracted from the study where it was found. Enough detail should be there to really communicate *what* they were trying to measure. It is at this time, when your head is in the literature, that it would be a good to begin working on your definitions. It is hard to remember all the details that you will have access to in the articles and it is very difficult to go back to an article to look at it again. We suggest you consider the parts of a definition now as you discover and uncover the domains, and then track and modify them as your process goes on.

Thanks to our work at OMERACT 2020 we have now introduced a detailed definitional template. It will be described in further detail later in the chapter section 6.3; however, it should be started at this step in the process and continue throughout the generation phase.

This template can be started, completed, and updated as you move through this chapter. One thing we have realized is that working groups often stopped their definition at the broad domain level of, for example, “Pain”. This was quickly problematic when anyone tried to operationalize it or begin to select an instrument at some future point in time. Conceptual clarity is very important when talking about domains (8). No one could remember if it was pain intensity, or the impact of pain on daily living. The broad domain would be pain intensity, or it could be the impact of pain on daily activities.

It is important to make sure you are generating domains that cross all the core areas describe in the OMERACT framework (see above Figure 4.2 and 4.3). OMERACT is looking for at least one domain from each of the core areas.

In the example of pain described above, pain intensity would fit into the OMERACT Framework Core Area of ‘manifestations/abnormalities’ while pain impact on daily activities would fit into ‘life impact’. Three core areas are required; Manifestations/Abnormalities, Life Impact, and Death/Lifespan; and one strongly recommended: Societal/Resource Use. If Societal/Resource Use will not be included, there needs to be an adequate and agreed upon justification for its exclusion.

6.2 Conduct qualitative work to identify candidate domains

> TIP: Remember that qualitative work must have representation from each relevant stakeholder group and from at least 3 continents with a suggested minimum of 30 participants' **total**

The Working Group initiates stakeholder consultation to determine what domains each stakeholder group deems essential or desirable to consider for measurement across all Core Areas. Formal qualitative research is an excellent way of obtaining the experiences of patients, family, and health care providers with the goal to explore the nature and the spectrum of the domain (e.g., fatigue or pain) encountered within the disease. Similarly qualitative work can engage another set of stakeholders in an imaging domain. In this area, more physicists, technicians, and engineers might join clinicians and researchers in a field to get a fuller array of options for imaging something like bone density, or inflammation.

Rigorous qualitative methods must be used with the collaboration of a qualitative methodological expert. This is to ensure scientific rigor in study design (theoretical underpinning; patient selection; conduct, recording and transcribing of interviews; data analysis and interpretation). Having an informal discussion with a few patients can be a useful precursor but is not qualitative research. There are guidelines for well conducted inductive research technique to elicit new concepts. These include recent documents from the FDA (9) on qualitative research to identify key domains, and the ISPOR work on concept elicitation (10). Both emphasize the conduct of solid qualitative work, engaging the correct stakeholder points of view, and a thorough exploration of each concept or domain. In qualitative terms this is often called continuing your work until you hit "saturation" which means that no new domains are being identified after a series of interviews, and there is a clear understanding of the breadth and depth of the domain that was discussed. Qualitative research is an entire field of research, too broad for this handbook. However, these two guidelines provide a focused approach to qualitative work aiming specifically at domain elicitation and definition.

> At this point you will have a strong understanding of the domain you have just studied. No time like the present to add this to your definitional template.

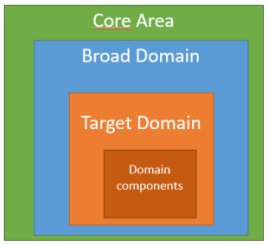
Disease experience and the way it is talked about can vary by geography and culture. Because we are an international organization and trials are often conducted or interpreted internationally, OMERACT recommends the qualitative work (individual interviews or focus groups) should aim to be as representative as possible of potential clinical trial participants with a minimum of 30 participants with a relevant stage and experience of disease, and with representation from at least 3 continents (11).

Publication of the results of qualitative work is strongly encouraged (6, 12, 13). Also having qualitative methodological expertise on the team will ensure this is to Consolidated criteria for reporting qualitative research (COREQ) standards (equivalent to CONSORT standards for RCTs) (http://www.equator-network.org/reporting-guidelines/coreq). Working Groups are encouraged to review this checklist for interviews and focus groups prior to starting qualitative work to ensure that all key items can be addressed in the reporting of their results. Any required ethics approvals and consent issues should be identified during the project plan discussions.

6.3 Domain definitions

While a domain label might be brief, the definition of the domain provides the rich description of what is included in that domain and what is not. If we use the metaphor of a domain being like a window in a house offering a certain view of an outcome, the right view will be when the window has the right breadth and depth. A rich definition accompanied by examples from qualitative or quantitative work helps us to describe that breadth and depth.

Earlier we introduced the idea of layers of definitions. Imaging outcomes have been using a layered approach to define their outcomes. They might start with a broad term like inflammation, but then they go to what they can measure, synovitis for example, and then they go even more specifically to blood flow and synovial tissue thickness for example (14). We used this layered approach to develop a template for all definitions, not just imaging, and it worked. This layered definition captures the core area, the broad domain, the target domain (what we are trying to measure) and the domains' components (things that should be included in the domain). If you want to think about this some more, we highly recommend reviewing the whiteboard lesson #1 for OMERACT from Imaging - Detailed Domain Definitions (https://youtu.be/omKD1z2MO78)



	Description	Example	Example	Example
Core Area	One of the Core Areas as defined in Boers, 2019.	Pathophysiological manifestation	Life Impact	Pathophysiological manifestations
Broad Domain	Often there is a general term for an area.	Pain	Pain	Bone fragility
Target Domain	This is the more focused view of the domain that would be placed in the onion.	Intensity of pain	Pain impact on daily activities	Bone density
Target Domain: more detail in definition	In this section, we are at the same level as Target Domain, however we are asking for fleshing this out. This might fit in as the definition used for this domain in the Delphi survey	The daily average of the intensity of the sensation of pain expressed on a range from no pain to worst pain imaginable.	A sense of the amount of day people are impacted by pain in terms of the accomplishment of daily activities and roles other than work.	The volume of mineral bone structure in a given volume of bone. It is capturing the density of the bone component of the space-bone matrix that makes up cancellous

				bone.
Elemental components of the domain	What areas are essential to capture to measure this target domain well?	24 window of pain, average day not special events or activities. Provide numeric scale, higher number = higher magnitude of pain sensation	Should consider selfcare, leisure, social roles at home (parenting). Not work. Generalized over whole day rather than in morning or evening.	Proportion of bone in g/cl at femoral neck and specific settings/machine

Table 4. 2 Some examples of the layers of a definition.

This part of our template plays another important role. When consensus activities like a Delphi process begin, they need to have a name for the domain and a definition to its meaning. We suggest using the target domain as the label and the definition of the target domain as the definition. This template will be imported into excel and become the domain names and their definitions that are offered to respondents during the survey. So, it can serve both roles.

Our experience has told us that when groups evolve, or time passes we lose the detailed understanding of a domain and why it was important or considered. We need that to be held for the working group and other OMERACT groups who might be considering a similar domain.

OMERACT requires a clear definition of each of the potential domains in the Core Domain Set. This definition should be true to the work that uncovered it, and the experiences of people who generated the idea if it was found qualitatively. It is also possible to look for a theoretical or conceptual framework that includes the domain and could provide more insight into its definition and its relationship to other domains. An existing conceptual framework like this could define a concept adequately for the Working Group (i.e., self-efficacy from social cognitive theory and self-management literature). If it is a clear match to the domain as described by your literature review or qualitative work than you can use that conceptual framework's definition and cite it as so. For example, if the concept is self-efficacy, the literature supported by Bandura's definition (15) might be used to define self-efficacy as "an individual's belief in his or her capability to produce given attainments". It might also be operationalized in Lorig's work in arthritis (16) as a mediator of care utilization.

Working groups have often found it important to take a more complex definition to an OMERACT meeting so that you can get more input into how others are seeing the domain, or what might be impacting it's meaning or measurement. For example, groups can reflect on the qualitative finding for a set of domains under 'Life Impact', discussing and documenting the breadth and depth of the experience in patients. In OMERACT, the work on the domain of fatigue (17) is an example of how breakout sessions during an OMERACT meeting paid considerable attention to the meaning of fatigue in persons with arthritis and led to a good discussion that vastly improved the domain definition for Fatigue. Several avenues can be used to improve the understanding of the domain.

Rheumatoid Arthritis (RA) Flare is another group where an emerging domain of patients' experience of flare from a disease perhaps in remission is the management target for new therapies. This domain was developed and defined through the OMERACT process (18). Early engagement and coordination with existing disease-focused Working

Groups during the development of a domain that is relevant across several health conditions or might be new to a disease group is essential. Our Budapest meeting had a pre-meeting to discuss the measurement of pain and whether this was common across different disease groups (chronic versus episodic pain for example). As our sense of definitions evolved, and we realized the importance of having these detailed definitions for OMERACT working groups. We developed a domain definition report based on the work by d'Agostino (14). The table of examples above follows this report, and it in turn is following the layers of definition described. The table above describes several domains and is the first step in preparation for the Delphi process described in the prioritizing section below. But there are also two additional sections in this report: qualitative findings, and sources of variability.

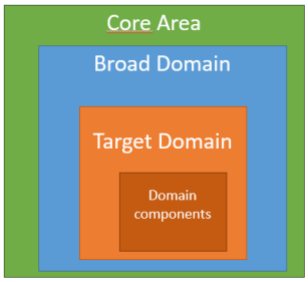
Qualitative findings. A quote from a patient, or a clinician can really provide a lot more meaning than a synopsis of that in a domain name or even its definition. It provides a rich description of the experience of that domain. We ask you to collect some of those salient quotes from the qualitative work or even from the literature so that you will always have that salience with you when you talk about domains. It is in the domain phase that these qualitative findings will be available in the mind of the person who did those interviews and analysis and that is precisely why we suggest using this time to document and write down some of those quotes.

Sources of variability. As you learn more about a domain that you are considering, you will hear a lot about sources of variability in its measurement or interpretation. For example, in worker productivity patient research partners started telling them "Well it depends" (a key sign that there is an important contextual factor here). It depended if they could pace their own work, or if they could take time off for important appointments or if they had switched to a less demanding job. These are all things that could make them respond to a question about their work ability or their overall productivity with a different numeric choice. In 2020 the contextual factors group led by fellow Sabrina Mai Nielsen described three main sources of variability in their operational definition of contextual factors, and their whiteboard video and background reading material is a great learning tool (https://youtu.be/WZStbgNNftc). From a domain definition perspective OMERACT feels that when you are in the literature about the domain definition you will also be finding out about contextual factors or sources of variability. For example, in imaging we learned that these definitions are discussed a lot and often with reliability tests to see if contextual factors, or sources of variability, are affecting a candidate outcome instrument score. This might be inter-rater differences, or a difference between brands of imaging machines. We suggest you document them as you go along in refining your definitions as you might forget them when you need them later!

Each domain nominated as part of your core domain set (inner layer of the onion) will require this domain definition sheet to be part of the report along with the OMERACT Onion and Workbook. We will hold a copy on the OMERACT website. Future groups would be well advised to search for existing definitions and see if they match an existing one. Such collaboration on definitions is exactly what we described around pain intensity.

Domain Definition Report

This is the one pager for each of your mandatory domains and any important but optional domains you feel you are ready to define. It will provide the domain definition in more detail than anywhere else and will be saved for future reference by OMERACT. In many situations this has become an invaluable resource when, perhaps years later, you are considering an instrument for your domain.



Working Group: _____

Target population _____

Intended use for this domain: _____ (e.g., RCT)

Intervention in trial: _____

Comparator in trial: _____

Core Area	*Life Impact*
Broad domain	*The general or broad domain, like "Pain Impact"* *Physical functioning*
Target Domain	*The name you are giving this more specific domain i.e., impact of pain on life activities in all realms of life- physical, social and role functioning. This is what we will be focusing on for measurement. Perceived physical functioning in daily activities*
Working definition of target domain	*Create a working definition in detail. Don't just repeat the domain name, flesh this out, this is what people will see in your paper. Sometimes this is a definition from another conceptual framework – for example the definition of pain impact should range from periodic interference over the course of one week to inability to do any activities due to this pain.* *Specify difficulty or frequency of experienced limitation.* *Consider the context – dependence on others because of limitations, difficulties doing activities, or more objective, capacity versus performance* *Appraisal – how is the assessment being done?*
Domain components	*Outline here the components of your domain that are important for a good instrument to capture.*

	e.g., pain impact on ADL's, pain impact on work life, pain impact on social activities. *Should not include pain impact on personality and personal relationships*
Qualitative or literature support	*Add in some quotes here from the qualitative work you have done. Consider examples of the breadth of the experience of this domain – high levels and low levels. Consider talking in more depth about what is included in this domain and what should NOT be considered part of this domain.* *This section is particularly important because it is easy to do as you work on your domains, and it will serve you well as a basis for your review of content validity when you start to look at candidate instruments. This definition sheet will be stored on OMERACT's website.*
Sources of variability in score	*Please think through sources of variability or contextual factors that might impact the results (scores) when you measure this domain. For example, are there sources of variability - contextual factors - that seem to influence the outcome scores (e.g. the type of job may explain differences in worker productivity), or factors that seem to result in random error or bias in the measurements (e.g. differences between different raters), or factors that generally seem to explain differences in treatment response (i.e. effect modifiers; e.g. disease duration, self-efficacy etc.).* *It may be difficult to distinguish the different types of contextual factors at this point, so it may be useful to put all sources of variability that the working groups encounter down on paper for future reference. Please see the paper by the contextual factors working group, Nielsen et al, Semin Arthritis Rheum, 2021 (*https://pubmed.ncbi.nlm.nih.gov/33875246/) and related material (https://omeract.org/working-groups/contextual-factors/)*, and work on lessons from imaging outcomes lesson #2 https://youtu.be/omKD1z2MO78*

AGREEING

7. Selection of candidate domains through consensus process

Once the candidate domains have been identified in step 6, the next step is to prioritize those domains through a consensus process. The purpose of this stage is to refine the initial list of generated domains to those that participants agree are critically important to a core domain set. These important domains are then placed into one of the three circles of the OMERACT Onion: inner circle of mandatory domains (which consists of two layers: (1) domains measured in all trials; (2) domains that are mandatory in specific circumstances) or middle circle of important but optional domains or the outer circle of domains requiring further research. This process is described more fully in section 8 below. The absolute essence of OMERACT is how formulating a Core Outcome Measurement Set is agreed upon through consensus. OMERACT recommends the use of Delphi Surveys to reach consensus and seek input across a large group of stakeholders. The Delphi is one part of the entire consensus process and is defined as a systematic means to measure and facilitate consensus (19, 20). The iterative process of ranking and then re-ranking the relative importance of different ideas and suggestions often generates some intellectual conflict as differing views are debated. While the result may not be everyone's preferred choice, the aim is to reach an agreement that all participants can accept as a working arrangement. Another important aspect is to refer to

evidence, wherever it exists, or to generate evidence that might inform the decisions of the participants, this is where the domain definition templates described above will be helpful. Consensus at one step feeds in to the next until there is international consensus on a core set of domains and outcome measures to use in all future rheumatologic clinical trials for a particular condition or disease.

From a list of candidate domains, participants select those domains they believe to be of critical importance for inclusion in a core domain set and this process is conducted iteratively over three, rounds until agreement is reached, often with an option in the first round to suggest additional domains. We have evidence that the number of items is predictive of the number of people who withdraw from Delphi surveys; a higher number of items results in a lower response rate (21); therefore, OMERACT suggests no more than 70 candidate domains in the initial round of the Delphi.

We have provided the following guidance to help Working Groups conduct their Delphi using DelphiManager modified for OMERACT. We recognize there are other survey methods that working groups might chose to use for prioritizing their complete list of domains generated from the search of the literature and qualitative work.

If a group decides to use DelphiManager modified for OMERACT, then they can simply complete the protocol template provided by OMERACT and submit it to their OMERACT Senior Methodologist for review. However, if the group decides to deviate from this method, we need to know ahead of time and the full Technical Advisory Group (TAG) will review these modifications.

The background information explaining the rationale for the Delphi should be prepared with consideration that it may be appropriate to prepare different information for the different stakeholders (e.g., patients may require more information to explain the study concepts). We recommend pilot testing the background information and initial Delphi questions with a small group prior to full implementation of the Delphi.

7.1 DelphiManager (DM) Modified for OMERACT

To further support OMERACT Working Groups OMERACT contracted COMET to modify their DelphiManager software to support the process of ranking then rating domains and sorting them into Core Areas. *DelphiManager modified for OMERACT*, is a web-based system designed to facilitate the building and management of Delphi surveys. This investment was made by OMERACT specifically to help groups follow the consensus process described above and to facilitate feedback between rounds by automatically collating and generating feedback.

At this stage of Domain agreement, we ask Working Groups to work closely with their OMERACT Senior Methodologist to develop the methods that they will use for their Delphi. OMERACT asks that groups follow the protocol for use of *DelphiManager modified for OMERACT* that can be found **HERE.**

This protocol would be suitable for submission for an ethics review for the Delphi portion of the domain selection. Some additional information might be required, such as a covering letter and consent form that is the responsibility of the investigators and consistent with the guidelines of their ethics review board.

The OMERACT approach has been reviewed against other standards for consensus (20) and Delphi techniques. The protocol as written addresses key aspects of these guidance documents. Any variation to the approach laid out in this protocol should be discussed with OMERACT and your Senior Methodologist to make sure you are still covering all aspects of best practices in using Delphi. Working groups are free to use the DelphiManager modified for

OMERACT with the cost covered by OMERACT. Any deviations from the protocol or to the software would be an independent process with DelphiManager, and costs would be borne by the working group.

7.2 Participant Selection & Recruitment

A foundational principle of OMERACT is the bringing together of multiple international stakeholders in collaborative research. Participants in the Delphi must represent a minimum of 3 continents and both patients and other stakeholders i.e., clinicians, researchers).
Examples of potential stakeholders include: 7 P's (22)

Category	**Description**
Patients and the public	Current and potential consumers of patient-centered health care and population-focused public health, their caregivers, families, and patient and consumer advocacy organizations
Providers	Individuals (e.g., nurses, physicians, mental health counselors, pharmacists, and other providers of care and support services) and organizations (e.g., hospitals, clinics, community health centers, community-based organizations, pharmacies, EMS agencies, skilled nursing facilities, schools) that provide care to patients and populations
Purchasers	Employers, the self-insured, government and other entities responsible for underwriting the costs of health care
Payers	Insurers, Medicare and Medicaid, state insurance exchanges, individuals with deductibles, and others responsible for reimbursement for interventions and episodes of care
Policy makers	The White House, Department of Health and Human Services, Congress, states, professional associations, intermediaries, and other policy-making entities
Product makers	Drug and device manufacturers
Principal investigators	Other researchers and their funders

OMERACT recommends starting with at least 100 participants per stakeholder group as we require stratification of the Delphi results by patients versus other stakeholder participants to see if there is a difference in which domains are considered important. Attrition always happens but starting with 100 will mean that working groups are likely to have a minimum of 30 to 50 participants in each of the 'patient' and 'other stakeholders' groups at end of the final round of the Delphi to be able to stratify and compare results.

The Working Group should consider which participant recruitment strategies they will use, e.g., direct, personalized contact, or indirect contact via websites, mailing lists, etc. Response rates may be improved by sending an initial email to potential participants outlining the purpose of the Delphi and the number of rounds planned. Consider multiple strategies to send automated reminders to participants as this may increase response rates and retention across Delphi rounds. Strategies to manage non-responders or partial responders should be identified at the protocol stage to reduce bias. Finally, all methods should be viewed through the lens of the Core Outcome Set-STAndardised Protocol Items (COS-STAD) (23) which provides a list of the critical elements in core outcome set development.

7.3 Domain Definitions

Following our domain definition work described in section 6 above. Working Groups are asked to complete a domain definition template as part of the protocol for *DelphiManager modified for OMERACT*. This template will feed into DelphiManager modified for OMERACT software providing the required background and details respondents will need to complete their survey.

The domains, the core area they represent, and their definitions are shown in table 4.2. In addition, under Pathophysiological manifestations/abnormalities, we have asked the working group to identify if this is a Symptom, a Sign, or a Biomarker (imaging, blood work).

In total there is a limit of a maximum of 70 domains allowed in the *DelphiManager modified for OMERACT* across all the core areas. We highly recommend the groups spend time refining their list to reduce redundancies and trim their list to ideally far less than 70, but not more.

Which Core Area does this belong to (Pathophysiological manifestations, Life Impact, Death/longevity, Resource Use)	Sub-category for Pathophysiological manifestations/abnormalities (symptoms, signs, biomarkers, other)	Broad Domain	Target Domain	Working Definition of target domain (will appear to help respondents understand domain)	Qualitative or literature support	Sources of variability in score
required	required	required	required	required	optional	optional
Select Core Area: Pathophysiological manifestations, Life Impact, Death/longevity, Resource Use	Sub-category for Pathophysiological manifestations/abnormalities (symptoms, signs, biomarkers, other)	The general or broad domain	The name you are giving this more specific domain i.e, impact of pain on life activities in all realms of life- physical, social and role functioning. This is what we will be focus-ing on for measurement. Perceived physical functioning in daily activities	Create a working definition in detail. Don't just repeat the domain name, flesh this out, this is what people will see in your paper. Sometimes this is a definition from another conceptual framework – for example the definition of pain impact should range from periodic interference over the course of one week to inability to do any activities due to this pain.	Add in some quotes here from the qualitative work you have done. Consider examples of the breadth of the experience of this domain – high levels and low levels. Consider talking in more depth about what is included in this domain and what should NOT be considered part of this domain. This section is particularly important because it is easy to do as you work on your do-mains, and it will serve you well as a basis for your review of content validity when you start to look at candidate instruments. This definition sheet will be stored on OMERACT's website.	Please think through sources of variability or contextual factors that might impact the results (scores) when you measure this domain. For example, is there a large differ-ence seen between people gathering the data? Is there a large difference between cul-tures or continents? Please see paper by contextual factors group Sabrina Nielsen et al, Ann

Table 4. 3 DelphiManager modified for OMERACT Domain Definitions Template

These domains along with their definitions will be submitted along with the *DelphiManager modified for OMERACT protocol* and any deviations from the protocol that the working group has chosen to make be reviewed for completeness by the OMERACT Technical Advisory Group. Please note the *DelphiManager modified for OMERACT* software was developed in collaboration with the Delphi Manager group to align it with OMERACT approaches to consensus. If you use this software the cost is absorbed by OMERACT. If you need to modify the software, that cost will need to be absorbed by the working group, and the modification included as a revision to the *DelphiManager modified for OMERACT protocol* provided by OMERACT.

Please note the *DelphiManager modified for OMERACT* software was developed in collaboration with the Delphi Manager group to align it with OMERACT approaches to consensus. If you use this software the cost is absorbed by OMERACT. If you need to modify the software, that cost will need to be absorbed by the working group, and the modification included as a revision to the *DelphiManager modified for OMERACT protocol* provided by OMERACT.

7.4 Defining Consensus

OMERACT applies a threshold of ≥70% participant agreement that a domain is of sufficient importance that it should be included in a draft core domain set to achieve consensus. We look at the votes stratified by stakeholder groups: A] patients and B] all other stakeholders. The following definitions are used for consensus for DelphiManager modified for OMERACT:

1. Consensus that a domain is important for a core domain set: ≥ 70% of participants in both groups (patients and others) scored the item as "critically important domains to include in a core set" (score 7 to 9); these domains are acknowledged in subsequent rounds as having met criteria for importance to a core domain set and held for final rating.

2. Consensus that a domain will NOT be included: ≥ 70% of participants in both groups (patients and others) scored the item as of " not important domains for a core set in this disease group " (score 1 to 3); these domains are dropped from Delphi and will not be part of core domain set. Document those domains dropped.

3. Dissensus but important to one group: 70%+ participants in one stakeholder group (patients or others) score items as critically important for a core set (score 7 to 9); domain continues to next round as having no consensus yet; if domain does not reach consensus level at end of Delphi, but still important to one group, it will go to middle circle of OMERACT Onion (Important but optional). More on this in section 7 below.

4. No consensus: All other results; domain continues to next round as having no consensus yet. If domain does not achieve consensus by last round, and no groups have supported it ≥ 70%, then domain is not endorsed for core domain set. Document those domains in report.

Rating scale between 1 and 9 is used with the following meanings:

≥ 7 critically important domains
> 3 and < 7 important but not critical domains
≤ 3 not important domains for a core set in this disease group

Rating Scale

1	2	3	4	5	6	7	8	9
not important domains for a core set in this disease group			important but not critical domains			critically important domains		

Table 4. 4 DelphiManager modified for OMERACT Domain Rating scale

7.5 Delphi Rounds

Three rounds and a rating round are offered in *DelphiManager modified for OMERACT*. The first three rounds are a typical Delphi Survey for selecting the importance of the domains for inclusion in the Core Domain Set in this area. Core domains are clearly defined as those minimal set of domains that (usually numbering 7 or less) that are fielded in all clinical trials or research in the defined field. Respondents will be asked to rate the importance of each domain (table 4.4) above for this Core Domain Set on a scale of 1-9 where 1-3 represents not important domains for a core set in this disease group, 4-6 important but not critical domains and 7-9 represents critically important domains to include. Comment fields are available after each domain to allow people to add additional notes about their decision. At the end of the list of domains there will be a chance for the respondents to nominate additional domains that they felt were missed.

A final (4th) round is offered in DelphiManager modified for OMERACT to gather all domains that have reached consensus to include across the Delphi process, or those domains where there is dissensus (at least one group endorsing as important but not both groups) for final comment. In this final round, the respondents will be able to select up to 10 domains as their top or most important ten, and upon selection of their choice, they will be asked to offer comments on saying why this domain should, in their opinion, be included or perhaps excluded. They will also be offered the opportunity to say the level at which it should be fielded (Mandatory in all trials, Mandatory in specific circumstances such as certain manifestations of disease or certain types of interventions, Important but optional, more research needed). Domains where consensus was to exclude, or those where there is no consensus (no group voting > 70% to include or to exclude) will not be offered in the final round.

This feedback along with the results of all three rounds of the Delphi process will be provided to the Working Group in an anonymous format who will use the results of the voting rounds, the selection of the most important domains, and the comments to make their final evidence-based decisions on the content of each layer of the OMERACT Onion or OMERACT Core Domain Set.

8. Formulation of Draft Core Domains and their detailed definitions

The work of creating a draft core domain set is not finished with the Delphi, though the Delphi results combined with the qualitative findings from the generation phase provides the bulk of the information needed for this next phase of identifying the core domains and placing them within the OMERACT Onion.

The initial list of generated domains is pared down to those where both patients and other stakeholder participants agree are critically important to a core domain set.

However, this could still be too many domains to make mandatory for every trial. The working group must then look at the round 4 rating descriptions to get more insight into which domains to keep as core, that is mandatory in all trials. The final set is placed within the Mandatory inner layer of the onion. This could be mandatory in all trials, or in circumstances where some preconditions may make the domain relevant, it would be placed at Mandatory in specific circumstances, and those specific circumstances must be described. For example, in a trial if a person has uveitis, eye involvement, then that might trigger adding the domain of visual acuity to the core set. If another person does not have uveitis, that domain is not needed. Similarly, if a person is employed for pay it might trigger a productivity outcome, whereas if they are not, then they do not need to complete that outcome.

The middle layer of the onion is the place to put domains considered to be important to at least one of the stakeholder groups. These are not mandatory but are optional. The outer ring of the onion is called a research agenda and is reserved for those domains where there was interest in the domain, but further work is needed to expand the understanding of the domain. For example, participation in non-work activities might be of interest but require a review of the concept in more detail or in leisure research before it can be accepted as a core domain. Once this research is complete, this type of domain could be considered for the next round of revising the Onion (Figure 4.4).

8.1 Facilitation techniques to help groups prioritize for placement of domains in the layers of the Onion:

Different groups have used different techniques to decide on the placement of domains in the Onion. Many have chosen to use small group techniques to engage stakeholders in the placement process and the important decision of what will be in the inner core, the core domain set. These are some techniques that have been tried.

- Vignettes with small, videotaped testimonials about the domains that were voted on as critical could be placed online, and people could review at their leisure with the additional comments offered for each in Round 4 (rating).

- Card sorting exercises, as described in The Workshop Book (24), in which participants use file cards and sort them on a wall to prioritize their choices for important domains, may be used by individuals and then brought together for a group decision. This can be replicated online using something like Google Jamboard to organize domains according to their core area and their likely role in capturing benefit or harm. (See figure 4.6).

- Dot votes have been used at OMERACT meetings in which a fixed set of coloured dots are allocated to each participant to use as "votes" endorsing candidate domains they consider to be mandatory. “Speed dating" circles have participants move to stations across a room where Working Group members explain and help to champion a specific domain.

These and other techniques are only suggestions. Working groups should decide upon how they will make this final decision making on their recommended core domains, and how to engage the OMERACT community to better understand the importance of each of the domains. It has been shown that engaging people in the material they are about to decide on is critical to a good consensus process.

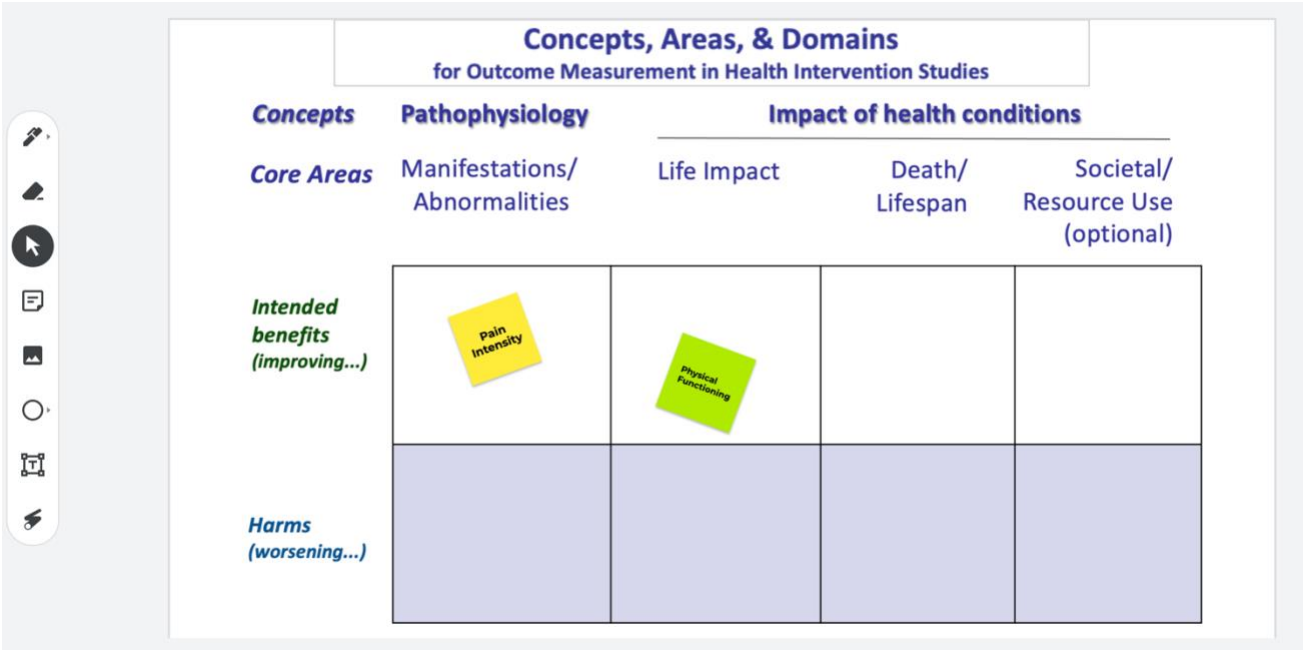


Figure 4. 6 Example of post-it notes on an online" Jamboard" for deciding on core domains.

At least one domain must be specified for each of the four Core Areas. At this point resource utilization (e.g., costs) is not mandatory. Therefore, Core Domain Sets will always contain a minimum of 3 Domains (corresponding to the three mandatory Core Areas of Manifestations/Abnormalities, Life Impact and Death/Lifespan). If the Core Area of Societal/Resource Use is not addressed, the Working Group must provide a statement explaining their decision.

We stress that a Core Domain Set aims to capture the *minimum* number of domains necessary to adequately capture what we want to know and communicate across clinical trials. The usefulness and uptake of a Core Domain Set strongly depends on this parsimony so that it can be included even if it does not include the primary endpoint for the trial. How many domains should be considered in the core? There is some evidence that 7 ± 2 individual items are the maximum the human brain can simultaneously consider (25). For example, the OMERACT core outcome set for rheumatoid arthritis that led to the development of the ACR 20/50/70 response criteria contains seven domains. Also, the Cochrane Summary of Findings tables allow up to 7 outcomes (domains). Therefore, OMERACT suggests Working Groups should strive for **approximately 5-7** domains in the Mandatory (Inner Circle) Core of a Core Domain Set in their onions

Groups at OMERACT are sometimes working on just one domain (e.g., Worker Productivity, Imaging). Sometimes these domains were in the outer circle of a Working Group's full Core Domain Set and a new Working Group was formed to address the research agenda to explore in which situations their domain should be considered either: inner circle (core domains; mandatory); middle circle (important but optional); or outer circle (a research agenda topic) for disease core sets. Often these domains are complex or require a lot of attention to detail or technique. The work for these Working Groups is similar – conceptual work, defining the domain(s) well and recommending placement in the Onion for the different working groups. The work is then integrated into the disease Working Groups' Core Domain Sets.

The figure below is an example of the core domains of the hip and knee osteoarthritis group (26). Five domains are considered mandatory in all trials, with one that is mandatory in certain circumstances (joint structure) and the defining circumstances is when trials have more than 2 years follow up.

2: Endorsed OMERACT-OARSI core domain set for trials of people with hip and knee osteoarthritis.

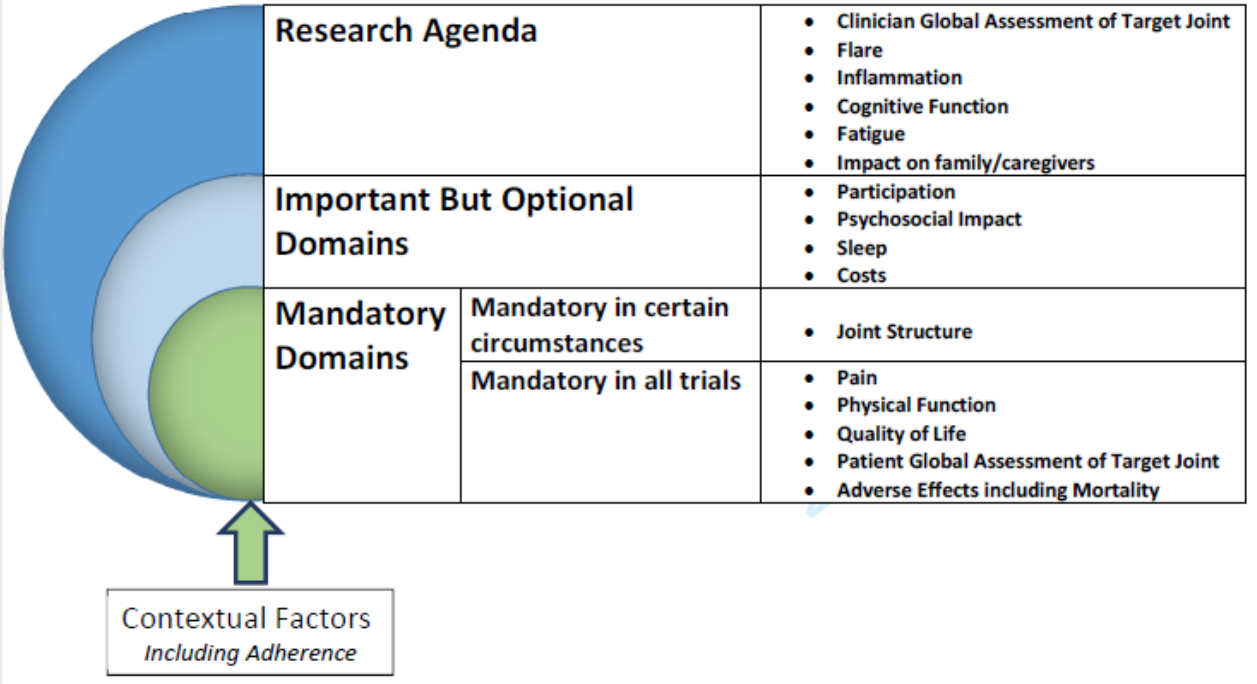


Figure 4. 7 'Inner Circle', 'Middle Circle' and 'Outer Circle' for domains in hip and knee osteoarthritis

8.2 Defining the Core Domains

Working groups will then be asked to provide the definitions for each of their core domains. Some of this work was done up in sections 6 and 7 using the domain definition template. At this point in the process, we are asking for this to be finalized on the domain definition sheets for all the domains that were placed into the mandatory categories of the Onion. The investment of time at this phase will serve the group well when they begin to move into the instrument phase. The domain definition template serves as the foundation of what the content of an instrument should be and help with the steps of concept match and content validity in the first steps of instrument selection process. To do this well, working groups should revisit each of the definitions for the mandatory domains and ensure there is enough detail. The more the better because this will be difficult information to remember when it is asked for at the instrument selection stage. We have had groups stumble later in the process because they did not have this detail. Working groups are encouraged to pull in qualitative quotes or more descriptions from the theoretical literature they used in the earlier stages of domain generation and selection for information to feed into the definitions. Remember this only needs to be done for the mandatory domains. A separate domain template needs to be created for each domain.

As we begin accumulating these domain definition templates on the OMERACT Website, we will be asking working groups to check and see if one of the other domain definition templates would be a match for their work. They might for example have pain intensity and find that another group has a domain definition template that matches their need exactly. This also means that these two disease groups are sharing a domain that is measured in all the clinical research.

9. Formulation of Core Contextual Factors

Initially, OMERACT defined a contextual factor as a "*variable that is not an outcome of the study but needs to be recognized (and measured) to understand the study results. This includes potential confounders and effect modifiers*" (1), and core outcome set developers were tasked to consider if there are any contextual factors that should be measured in all trials together with the core outcome set. However, the research that was presented at the OMERACT meeting in 2014 revealed that the working groups understood, approached and identified contextual factors in very different ways. To address this, the Contextual Factors Working Group was established (27). The working group has spent a tremendous amount of effort trying to sort through the various types of contextual factors by the way they impact the results of a trial. The group defined three types of contextual factors, describing different ways that factors can influence the results of a trial (see **Figure 4.8**) (28). These three types have been termed Effect Modifying - (EM-CFs), Measurement Affecting - (MA-CFs) and Outcome Influencing Contextual Factors (OI-CFs).

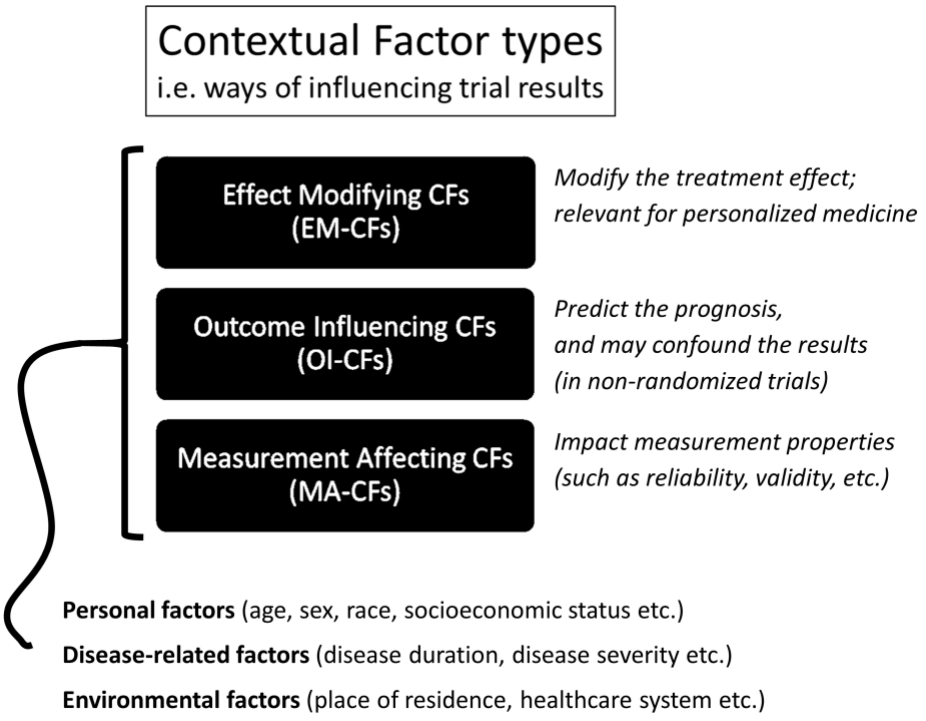


Figure 4. 8 Overview of the operational definition of contextual factors. The three contextual factor types describe different ways that contextual factors can influence the results of a trial. The three contextual factor types describe different ways that contextual factors can influence the results of a trial. To guide which possible factors could be considered within each of these types, specific factors must fit within one of the three classification categories, i.e. either personal-, disease-related, or environmental factors. Therefore, contextual factor types are not mutually exclusive, and some specific factors may both be an EM-CFs, OI-CFs and MA-CF. CFs, Contextual Factors.

In short:

- EM-CFs modify the treatment effect, i.e. some patient subgroups experience greater or less effect (net benefit) from a treatment compared to other subgroups.
  E.g. if women experience less effect (net benefit) from a treatment compared to men in a RCT, then sex is said to modify the treatment and, hence, is an EM-CF.

- OI-CFs are prognostic factors (sometimes called risk factors), i.e. factors predicting the course of a patient's condition and may confound the results of cohorts that are not randomized.
  E.g. if high body mass index (BMI) are associated with more disability in the future, then BMI is said to be prognostic (as it is associated with the outcome over time) for future disability and, hence, is an OI-CF.

- MA-CFs influence the performance of outcome measurement instruments, such as outcome measurement properties including reliability, validity, responsiveness, etc.
  E.g. if the test-retest reliability of the measurements of pain with the Visual Analog Scale (VAS) are lower for patients that are illiterate, then literacy influences the measurement properties and, hence, is a MA-CF.

EM-CFs are mainly relevant for core domain sets for clinical trials. OI-CFs relates to non-randomized trials and/or longitudinal observational studies comparing exposed with unexposed individuals. MA-CFs are mainly relevant for core outcome measurement sets, since it is the instruments' measurement properties that are being tested. However, candidates for any of the contextual factor types may appear in the qualitative work, and since it may be difficult to distinguish them in the beginning, noting down all contextual factors that the group encounters for later references may be useful. To gently guide which possible factors could be considered contextual factors, specific factors must fit within one of the three classification categories, i.e. either personal-, disease-related, or environmental factors (see **Figure 4.8**). The contextual factor types are ***not*** mutually exclusive, so some specific factors (e.g. age, sex, race) may both be an EM-CF, OI-CF and MA-CF.

Please note, since each of the contextual factor types are related to statistical concepts, high level of empirical evidence is needed to call a contextual factor '*core'*. Identified contextual factors should simply be termed '*important* contextual factors' until sufficient evidence is present sometimes in the future.

The topic of contextual factors is complex and under continuous development, and one may notice that there is some fortunate overlap with the equity extension for the OMERACT instrument selection process, as well as sources of variability for outcome measurement instruments by the Imaging group. A more in-depth explanation of each contextual factor type and relevant material are available from the website of the Contextual Factors Working Group https://omeract.org/working-groups/contextual-factors/

10. Working Group agrees on, finalizes & submits the Draft Core Domain Set to the Technical Advisory Group

A vote is taken by Working Group members to obtain agreement with the draft Core Domain Set as represented in the OMERACT Onion. The most important priority is to get a vote on the Core (Mandatory) part of the Onion. We recommend that this is the focus of the voting.

Your first vote would therefore be for the acceptance of the domains in the inner circle (mandatory domains).

TIP: The Domain Workbook should be complete by this stage and can be used to present the evidence to the entire Working Group

As with other OMERACT voting, at least 70% percent agreement is required in the Working Group to be considered endorsed. In addition to the overall vote from all Working Group members, the voting results by each subgroup of participant stakeholder should also be reviewed. These subgroups are the Patient Research Partners, and the other stakeholders (Researchers, Clinicians, Policy makers etc.). If there is a specific stakeholder subgroup (e.g., patients), who has not reached the 70% threshold or who has voiced a significant dissenting voice even if the 70% threshold has been reached, then the Working Group needs to address any concerns until all stakeholder groups agree. Disagreement at this stage often points to an area requiring further clarification. This is all valuable to work through as a Working Group and improve your product before you move to the full OMERACT vote.

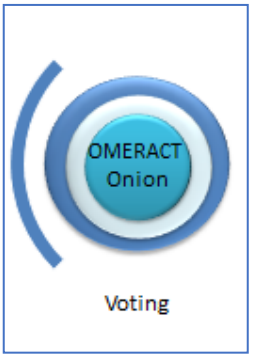
The Working Group then submits the Draft Core Domain Set reports to the OMERACT Secretariat (admin@omeract.org) for review and approval by the Technical Advisory Group.

The required reports include the following:

- Overview of the Draft Core Domain set in the format of the OMERACT Onion (approx. 1 page)
- Detailed definitions of each domain in the Core Set Domain (mandatory or mandatory in specific circumstances). This would mean a collection of one page of definitions and elaboration documents using the domain definition template.
- Completed Domain Workbook detailing specific methods (approx. 25 pages)
- Any other supplementary material such as ethics approvals or additional publications related to this work

VOTING

11. Result of final vote by full OMERACT membership on Core Domain Set



After a supporting nod from the Technical Advisory Group, the Working Group seeks OMERACT endorsement of the Core Domain Set through a vote from the entire OMERACT membership. In this process you will have time to engage the OMERACT community in your results through whiteboards, discussion boards, papers, and other creative means. Remember that the OMERACT community needs to understand that you have done a rigorous, thoughtful job of creating this core domain set to feel confident in providing you with an endorsing vote. Use your workbook to develop a flow chart of your activities.

OMERACT is a patient-centered organization, and we reflect this in our voting process. Votes will be stratified and shown separately for A] Patients and B] All other stakeholders [i.e., everyone else voting].

In the Domain Workbook, we have provided templates for the voting questions you plan to ask at the OMERACT membership endorsement vote. As with your Working Group vote, we recommend that there are 2 separate votes:

one on the draft Inner Circle Core Domain Set and the second on the completed OMERACT Onion as a whole. Your first vote would therefore be for the acceptance of the domains in the inner circle (mandatory domains).

OMERACT sets up the final voting on your domains set in the following manner:

1. Vote on the inner circle as a whole and the acceptance of that as the Core Domain Set.

    If 70 % (or greater) of both a] Patients and b) All other stakeholders approve, then these domains are approved for the Inner Circle of the Onion and are the Core Domain Set.
    If under 70 % in either or both groups, then the Inner Circle is NOT approved

2. If there is an under 70% endorsement vote on the inner circle, we recommend moving to a vote on each of the inner circle domains individually

    a. Your group can conclude the domains that do achieve a 70% vote from both a] Patients and b) All other stakeholders will be considered in the Core Domain Set that has received endorsement

    b. If an individual domain does not reach 70% from both a] Patients and b) All other stakeholders, the working group may move it to the middle circle. Importantly, sometimes one of the groups is just slightly under or over that threshold of 70%. In those circumstances we have specific ways to manage the vote:

        - If there is a discrepancy between Patient and Other Stakeholders votes where one group is above, and one is below 70% and the discrepancy is less than 10% samples are weighted equally using the following formula (i.e., pCombined = [pPts + pOther]/2)

        - If the discrepancy is greater than 10% and one is above and one is below 70% the domain, no average is taken. It is considered important to only one stakeholder group and is recommended for the middle layer of the onion (Important but optional domains).

    c. You can set a time to revisit the domains that were not endorsed. Often time, it is a matter of clarity about the domain definition or the work you had done.

3. The next vote is on the remainder of your “Onion”, that is the middle and outer circles.

The OMERACT secretariat will collect the votes and provide the results to the Working Group chairs to add to their Domain Workbook.

Once the Core Domain Set is endorsed, we recommend registering it in the COMET database (www.comet-initiative.org) as a start to your implementation activities. Implement your plan for other dissemination activities for your core domain set to your key stakeholders. OMERACT will post the endorsed core domain set on their website.

Shout it out loud! You are done, so how do you get your message out?

- Register core domain set at COMET – update phase of work as completed
- Submit definition templates to be added to OMERACT repository.

- Publish your core domain set creation in an OMERACT publication
- Create a KT plan for making clinical groups, pharma, regulators etc. aware of your core domain set. Uptake is just as important as all the work you have put into date.

Next Steps

Now that the Working Group has an OMERACT endorsed Core Domain Set, the next step is to identify at least one outcome measurement instrument per domain to develop the Core Outcome Measurement Set.

Chapter 5, 'Instrument selection for Core Outcome Measurement Sets' of the OMERACT Handbook and the accompanying Instrument Selection Workbook will guide you through this process. This process begins with the domain definition template, so bring that along into the instrument selection process.

OMERACT has a process, mentors, and supporting tools for OMERACT Instrument Selection. Have a look at Chapter 5 and get in touch with the OMERACT Secretariat to be directed to some of these resources. We also have some up on the website under Instrument Selection.

Congratulations!

REFERENCES

1. Boers M, Kirwan JR, Wells G, Beaton D, Gossec L, d'Agostino Maria-Antonietta, et al. Developing Core Outcome Measurement Sets for Clinical Trials: OMERACT Filter 2.0. Journal of Clinical Epidemiology 2014; 67:745-53

2. 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

3. Kirkham JJ, Davis K, Altman DG, Blazeby JM, Clarke M, Tunis S, Williamson PR. Core Outcome Set STAnDards for Development: the COS-STAD Recommendations. PLOS Med. 2017;14(11):e1002447

4. Kirkham JJ, Gorst S, Altman DG, Blazeby JM, Clarke M, Tunis S, Williamson PR and for the COS-STAP Group. Core Outcome Set-STAndardised Protocol Items: the COS-STAP Statement. Trials 2019 20:116

5. LOW BACK PAIN REF

6. Matthew J Page, Denise A O'Connor, Mary Malek, Romi Haas, Dorcas Beaton, Hsiaomin Huang, Sofia Ramiro, Pamela Richards, Marieke J H Voshaar, Beverley Shea, Arianne P Verhagen, Samuel L Whittle, Danielle A van der Windt, Joel J Gagnier, Rachelle Buchbinder, OMERACT Shoulder Core Set Working Group, Patients' experience of shoulder disorders: a systematic review of qualitative studies for the OMERACT Shoulder Core Domain Set, *Rheumatology*, Volume 58, Issue 8, August 2019, Pages 1410–1421, https://doi.org/10.1093/rheumatology/kez046

7. Tritschler T, Langlois N, Hutton B*, et al* Protocol for a scoping review of outcomes in clinical studies of interventions for venous thromboembolism in adults *BMJ Open* 2020;**10:** e040122. doi: 10.1136/bmjopen-2020-040122

8. A Call for Evidence-based Decision Making When Selecting Outcome Measurement Instruments for Summary of Findings Tables in Systematic Reviews: Results from an OMERACT Working Group Dorcas E. Beaton, Caroline B. Terwee, Jasvinder A. Singh, Gillian A. Hawker, Donald L. Patrick, Laurie B. Burke, Karine Toupin-April, and Peter S. Tugwell J Rheumatol October 2015 42(10):1954-1961; published online before print September 15, 2015, doi:10.3899/jrheum.141446

9. Patient-Focused Drug Development: Methods To Identify What Is Important to Patients; Draft Guidance for Industry, Food and Drug Administration Staff, and Other Stakeholders https://www.fda.gov/regulatory-information/search-fda-guidance-documents/patient-focused-drug-development-methods-identify-what-important-patients-guidance-industry-food-and

10. Patrick DL, Burke LB, Gwaltney CJ, Leidy NK, Martin ML, Molsen E, Ring L. Content validity--establishing and reporting the evidence in newly developed patient-reported outcomes (PRO) instruments for medical product evaluation: ISPOR PRO good research practices task force report: part 1--eliciting concepts for a new PRO instrument. Value Health. 2011 Dec;14(8):967-77. doi: 10.1016/j.jval.2011.06.014. Epub 2011 Oct 13. PMID: 22152165.

11. Chapter 2: Working Groups

12. Fuller A, Cai K, Diaz-Torne C, Filippou G, Pascart T, Hensey O, Grossberg D, Christensen R, Shea B, Singh JA, Tedeschi SK. Outcome domains reported by patients, caregivers, healthcare professionals and stakeholders for calcium pyrophosphate deposition (CPPD): A content analysis based on semi-structured qualitative interviews from the OMERACT CPPD working group. In Seminars in Arthritis and Rheumatism 2021 Jan 6. WB Saunders.

13. Hawker GA, Stewart L, French MR, Cibere J, Jordan JM, March L, Suarez-Almazor M, Gooberman-Hill R. Understanding the pain experience in hip and knee osteoarthritis–an OARSI/OMERACT initiative. Osteoarthritis and cartilage. 2008 Apr 1;16(4):415-22.

14. Imaging article
15. Bandura A. Self-efficacy: Towards a Unifying Theory of Behavioral Change. Psychological Review 1977. 84(2) 198-215
16. Lorig KR, Holman HR. Self-management education: History, Definitions, Outcomes, and Mechanisms. Ann Behav Med 2003. 26(1):1-7
17. Kirwan J, Hewlett S. Patient Perspective Workshop: Reasons and methods for measuring fatigue in rheumatoid arthritis. J Rheumatol 2007; 34: 1171-3.
18. Bartlett SJ, Hewlett S, Bingham CO, Woodworth TG, Alten R, Pohl C et al. Identifying core domains to assess flare in rheumatoid arthritis: an OMERACT international patient and provider combined Delphi consensus. Ann Rheum Dis. 2012 Nov;71(11):1855-60.

19. Sinha IP, Smyth RL, Williamson PR. Using the Delphi technique to determine which outcomes to measure in clinical trials: recommendations for the future based on a systematic review of existing studies. PLoS Med. 2011 Jan 25;8(1): e1000393.

20. Humphrey-Murto S, Crew R, Shea B, Bartlett SJ, March L, Tugwell P, Maxwell LJ, Beaton D, Grosskleg S, de Wit M. Consensus building in OMERACT: recommendations for use of the Delphi for core outcome set development. The Journal of rheumatology. 2019 Aug 1;46(8):1041-6.

21. Gargon E, Crew R, Burnside G, Williamson PR. Higher number of items associated with significantly lower response rates in COS Delphi surveys. J Clin Epidemiol. 2018 Dec 14. pii: S0895-4356(18)30776-5. doi: 10.1016/j.jclinepi.2018.12.010. [Epub ahead of print]

22. Concannon TW, Meissner P, Grunbaum JA, et al. A new taxonomy for stakeholder engagement in patient-centered outcomes research. J Gen Intern Med. 2012;27(8):985-991. doi:10.1007/s11606-012-2037-1
23. Kirkham JJ, Davis K, Altman DG, Blazeby JM, Clarke M, Tunis S, Williamson PR. Core Outcome Set STAnDards for Development: the COS-STAD Recommendations. PLOS Med. 2017;14(11): e1002447

24. Stanfield RB. The Workshop Book: From Individual Creativity to Group Action. Gabriola Island and Toronto: New Society Publishers and Canadian Institute of Cultural Affairs; 2002.

25. Miller, G. A. (1956). "The magical number seven, plus or minus two: Some limits on our capacity for processing information". Psychological Review. 63 (2): 81–97. PMID 13310704. doi:10.1037/h0043158.

26. Smith TO, Hawker GA, Hunter DJ, March LMM, Boers M, Shea BJ, et al. The OMERACT-OARSI Core Domain Set for Measurement in Clinical Trials of Hip and/or Knee Osteoarthritis. J Rheumatol January 2019, jrheum.181194; DOI: https://doi.org/10.3899/jrheum.181194

27. Finger ME, Boonen A, Woodworth TG, et al. An OMERACT Initiative Toward Consensus to Identify and Characterize Candidate Contextual Factors: Report from the Contextual Factors Working Group. The Journal of rheumatology 2017;44(11):1734-39. doi: 10.3899/jrheum.161200 [published Online First: 2017/05/04]

28. Nielsen, Sabrina Mai, et al. "OMERACT consensus-based operational definition of contextual factors in rheumatology clinical trials: A mixed methods study." *Seminars in arthritis and rheumatism*. WB Saunders, 2021.