

What Do the OMERACT Shoulder Core Set Candidate Instruments Measure? An Analysis Using the Refined International Classification of Functioning, Disability, and Health Linking Rules

Yngve Røe¹ , Rachele Buchbinder² , Margreth Grotle¹ , Samuel Whittle³ , Sofia Ramiro⁴ , Hsiaomin Huang⁵, Joel Gagnier⁵ , Arianne Verhagen⁶ , and Sigrid Østensjø¹ 

ABSTRACT. *Objective.* The objective of this paper is to assess the content and measurement constructs of the candidate instruments for the domains of “pain” and “physical function/activity” in the Outcome Measures in Rheumatology (OMERACT) shoulder core set. The results of this International Classification of Functioning, Disability, and Health (ICF)-based analysis may inform further decisions on which instruments should ultimately be included in the core set.

Methods. The materials for the analysis were the 13 candidate measurement instruments within pain and physical function/activity in the shoulder core domain set, which either passed or received amber ratings (meaning there were some issues with the instrument) in the OMERACT filtering process. The content of the candidate instruments was extracted and linked to the ICF using the refined linking rules. The linking rules enhance the comparability of instruments by providing a comprehensive overview of the content of the instruments, the context in which the measurements take place, the perspectives adopted, and the types of response options.

Results. The ICF content analysis showed a large variation in content and measurement constructs in the candidate instruments for the shoulder core outcome measurement set.

Conclusion. Two of 6 pain instruments include constructs other than pain. Within the physical function/activity domain, 2 candidate instruments matched the domain, 3 included additional content, and 2 included meaningful concepts in the response options, suggesting that they should be omitted as candidate instruments. The analyses show that the content in most existing instruments of shoulder pain and functioning extends across core set domains.

Key Indexing Terms: ICF, OMERACT, outcome assessment, shoulder pain

Shoulder pain is a common musculoskeletal disorder with an incidence of 10 per 1000 patients in primary care and point prevalence estimates of 7% to 26% in the general population¹. Shoulder disorders can be long lasting; in a Dutch study of patients presenting to their general practitioner with a new episode of shoulder pain, a considerable number (41%) showed persistent symptoms after 12 months². The associated disability and effect in terms of earnings, missed workdays, and disability payments are substantial^{3–7}.

The domains and measurement instruments reported in trials on shoulder disorders are widely diverse; therefore, the development of a core outcome set for use in clinical trials across shoulder disorders has been advocated⁸. Since 2016, there has been an ongoing effort to develop a shoulder core set within the Outcome Measures in Rheumatology (OMERACT)^{9–12}. At the OMERACT 2018 conference, a shoulder core domain set was approved by the delegates¹³. It consists of 4 mandatory domains for all trials of shoulder disorders: pain, physical function/activity, patient global – shoulder, and adverse events including death; and 4 important but optional domains: participation (recreation/work), sleep, emotional well-being, and condition-specific pathophysiological manifestations¹³. The next phase will be to recommend specific measurement instruments for a core outcome measurement set¹⁰.

Preliminary work has investigated instruments within 2 of the mandatory domains, pain and physical function/activity, identified from a systematic review of outcome domains and measurement instruments reported across randomized trials of any interventions for various shoulder disorders⁸. Pain was defined as “how much a person’s shoulder hurts, reflecting the

¹Y. Røe, Associate Professor, PhD, M. Grotle, Professor, PhD, S. Østensjø, Professor, PhD, Department of Physiotherapy, Oslo Metropolitan University, Oslo, Norway; ²R. Buchbinder, Professor, PhD, Department of Epidemiology and Preventive Medicine, Monash University, Sydney, Australia; ³S. Whittle, MD, PhD, The Queen Elizabeth Hospital, Woodville South, Australia; ⁴S. Ramiro, MD, PhD, Leiden University Medical Center – Rheumatology, Leiden, the Netherlands; ⁵H. Huang, Associate Professor, PhD, J. Gagnier, Associate Professor, PhD, Department of Epidemiology and Orthopaedic Surgery, University of Michigan, Ann Arbor, Michigan, USA; ⁶A. Verhagen, Associate Professor, PhD, University of Technology, Melbourne, Australia
Address correspondence to Y. Røe, OsloMet – Oslo Metropolitan University, Department of Physiotherapy, P.O. Box 4, St. Olavs Plass, N-0130 Oslo, Norway. Email: yngveroe@oslomet.no.

Accepted for publication January 27, 2020.

overall magnitude of the pain experience (i.e., at rest, during and after activity, at night)". Physical function/activity was defined as "a person's ability to carry out daily physical activities, ranging from self-care (e.g., bathing, combing hair) to more complex activities that require a combination of skills (e.g., driving a car)"¹³. Thirty-eight instruments within the pain domain and 45 within the physical function/activity domain were further investigated with the "Truth" Part 1 and "Feasibility" filters of OMERACT^{11,12}. Altogether, 6 instruments in the pain domain and 7 within the physical function/activity domain passed both filters and are candidates for further assessment^{14,15}. However, 5 in the pain domain and 3 in the physical function/activity domain received amber ratings for content validity, indicating potential limitations in their utility^{14,15}.

The International Classification of Functioning, Disability, and Health (ICF) is the World Health Organization framework for measuring health and disability¹⁶. Since its publication in 2001, the ICF has been used to describe and compare health information. To establish a standardized procedure to translate the content of measurement instruments into ICF concepts, a set of 10 linking rules were published in 2002 and updated in 2005^{17,18}. Since their introduction, a number of instruments have been linked to the ICF^{19–21}. To enhance the comparability of instruments, and ultimately to be able to aggregate information gathered with various instruments, it does not only require content comparability of items but also a reflection on the perspective they have adopted and the categorization of their response options.

In 2016, the linking rules were refined to account for these aspects, offering a more transparent tool to assess the content of measurement instruments and the context in which the measurements take place²². Thus, content linking of outcome measurement instruments based on the refined ICF linking rules provides information on important aspects of content validity. Content validity is considered to be the most important measurement property of an outcome measurement instrument, because if it is unclear what an instrument is actually measuring, the assessment of other measurement properties may be irrelevant²³.

The aim of the present study was to assess the content and measurement constructs of the candidate instruments for the domains of pain and physical function/activity in the OMERACT shoulder core set, using the refined ICF linking rules. The results of this ICF-based analysis may inform further decisions on which instruments should ultimately be included in the Core Set.

MATERIALS AND METHODS

The materials for the analysis were the 13 candidate measurement instruments within pain and physical function/activity in the shoulder core domain set, which either passed or received amber ratings (meaning there were some issues with the instrument) in the OMERACT filtering process^{14,15}. The 6 candidate instruments within the pain domain and the 7 within the physical function/activity domain are presented in Table 1^{24–34}.

These instruments are widely used in the clinical and epidemiological research of shoulder pain conditions⁸.

Analysis of content and measurement constructs. The ICF is based on an integrative model of health that classifies functioning within the components of body functions (b), body structures (s), activities and participation (d), environmental factors (e), and personal factors (not classified)¹⁶. The ICF provides 4 subclassifications (b, s, d, e), where categories of functioning and environmental factors are arranged hierarchically using an alphanumeric coding system. At the first level, the initial letter is followed by a numeric code (1-digit; e.g., d4 Mobility), 2 more digits for the second level (e.g., d445 Hand and arm use), and a total of 4 digits for third level categories (e.g., d4452 Reaching). A fourth level is also available when appropriate. An overview of the chapter structure of the components body functions and activities and participation is shown in Table 2.

The content from each item in the measurement instruments was linked to the ICF according to the 10 refined linking rules²². Linking rules 1 to 3 specify how to get familiar with the ICF and identify the purpose of an instrument and concepts to be linked to the ICF. Both the researchers who conducted the analyses (YR and SØ) had previously linked the content of shoulder pain instruments to the ICF²¹.

First, the actual meaning (main and additional concepts) of the information to be linked was identified, consistent with rules 2 and 3²². When identifying the concepts, both the item text and the text that sets the premises for the interpretation of the item content were taken into consideration. For most items, it was straightforward to identify the main and additional concepts. For example, for the item "How severe is your pain: pushing with the involved arm?", "pain" was identified as the main concept and "pushing with the arm" as an additional concept. In this item, the additional concept defines the context in which pain is assessed. Sometimes more than 1 activity was listed in the same item. When this was the case, all of the listed activities were recognized as main concepts. In a few cases, the item was framed in general terms, while specific activities were included in the response options, such as in the function subscale of the University of California at Los Angeles (UCLA) Shoulder Score function subscale, and the activities of daily living subscale of the Shoulder Function Assessment scale (SFA)^{32,33,34}. The naming of the item was then identified as the main concept and the specific activities as additional concepts.

The next step was to document the perspectives from which the information was collected (linking rule 4). The most common perspectives included in measurement instruments are the descriptive, appraisal, and needs or dependency perspectives²². The descriptive perspective refers to a person's function of the body, ability to perform a task in a standardized environment (capacity), or actual performance of certain tasks or activities in the natural environment. According to linking rule 5, the categorization of the response option in every measurement instrument was identified and documented.

Finally, all main and additional concepts identified during steps 2 and 3 were linked to the most precise ICF category (linking rules 6–10). For concepts not sufficiently specified to be linked, the "not definable" option was used. If a concept was not covered by any of the ICF classifications, the option "not covered" was used.

All instruments were independently assessed by 2 researchers (YR and SØ). In the case of differences in linking, this was solved by discussion. There were no cases of disagreement in the identification and documentation of perspectives and response options.

Agreement between the researchers in the linking of concepts at the second ICF category level was calculated with the Cohen k coefficient. The 95% CI for the k coefficient were calculated using the standard error (SE) of the kappa³⁵: $k - 1.96 \times SE_k$ to $k + 1.96 \times SE_k$. The calculated k coefficient of the linking of the main and additional concepts was 0.85 (95% CI 0.78–0.91) and considered as excellent (range, 0.61–1.00)³⁶. The study did not include data from patients or any other sensitive material, thus ethical approval was waived.

Table 1. Candidate instruments for the shoulder core domain set within pain and physical function/activity.

	Abbrev	Items	Response Scales (Range Score)
Pain instruments			
Visual analog scale	VAS	1	Visual analog scale (0–100)
Numerical rating scale	NRS	1	Ordinal scale (0–10)
Verbal rating scale	VRS	1	Categorical scale with optional response formats
Oxford Shoulder Score ^{24,25}	OSS	12	5-point ordinal scale (0–48)
Shoulder Pain And Disability Index, pain subscale ^{26,27}	SPADI	5	Visual analog pain scale or 10-point ordinal scale (0–100)
Shoulder Pain Score ²⁸	SPS	7	Items 1–5: four-point ordinal scale; item 6: visual analog pain scale (0–100); item 7: categorical scale, response formats indicating degree of pain radiation
Physical function/activity instruments			
Penn Shoulder Score, Function subscale ²⁹	Penn	20	4-point ordinal scale (0–60)
L'Insalata Shoulder Rating Questionnaire, daily activities subscale ³⁰	SRQ	6*	5-point ordinal scale
Simple Shoulder Test ³¹	SST	12	Categorical with yes/no (0–12)
American Shoulder and Elbow Surgeons Shoulder Outcome Score, activities of daily living subscale ³²	ASES	10	4-point ordinal scale (0–40)
University of California at Los Angeles Shoulder Score, function subscale ³³	UCLA	1	Intensity or categorical (1–10)
Shoulder Pain and Disability Index, disability subscale ^{26,27}	SPADI	8	Visual analog pain scale or 10-point ordinal scale (0–100)
Shoulder Function Assessment scale, activities of daily living subscale ³⁴	SFA	3	Intensity or categorical (0–20)

* Only items 6–11 were selected as candidates for the OMERACT core measurement instrument set. Abbrev: abbreviation; OMERACT: Outcome Measures in Rheumatology.

Table 2. Overview of the ICF domain (chapter) structure within the subclassifications of “body functions” and “activities and participation.”

Body Functions	Activities and Participation
Chapter b1: Mental functions	Chapter d1: Learning and applying knowledge
Chapter b2: Sensory functions and pain	Chapter d2: General tasks and demands
Chapter b3: Voice and speech functions	Chapter d3: Communication
Chapter b4: Functions of the cardiovascular, hematological, immunological, and respiratory systems	Chapter d4: Mobility
Chapter b5: Functions of the digestive, metabolic, and endocrine systems	Chapter d5: Self-care
Chapter b6: Genitourinary and reproductive functions	Chapter d6: Domestic life
Chapter b7: Neuromusculoskeletal and movement-related functions	Chapter d7: Interpersonal interactions and relationships
Chapter b8: Functions of the skin and related structures	Chapter d8: Major life areas
	Chapter d9: Community, social, and civic life

ICF: International Classification of Functioning, Disability, and Health.

RESULTS

Descriptive information about the 6 candidate instruments within the pain and physical function/activity domains of the shoulder core set is shown in Table 1.

“Pain” candidate instruments. The analysis of the perspectives showed that the “descriptive performance perspective” was adopted in all 6 instruments. The response options in the visual analog scale (VAS), the numerical rating scale (NRS), the Oxford Shoulder Score (OSS), and the Shoulder Pain and Disability Index (SPADI) pain subscale reflect intensity. In the verbal rating scale (VRS) the response options reflected “qualitative attributes”, and in the Shoulder Pain Scale (SPS), a combination of intensity and qualitative attributes.

All instruments had main concepts linked to sensory functions and pain categories in the ICF (Table 3). For the 3 overall pain scales (VAS, NRS, and VRS) and the SPADI pain subscale, all main concepts were linked to a pain ICF category. The overall pain scales only cover a single ICF pain category, while the SPADI includes 5 categories.

In addition to pain categories, the SPS included a main concept linked to a mobility category in the ICF. One instrument stood out from the others: in the OSS, 10 of 14 main concepts were linked to ICF categories other than pain, namely to activity and participation categories within the mobility, self-care, and domestic life chapters.

The additional concepts in the pain instruments provide

Table 3. Overview of linked ICF categories in the 6 candidate instruments of the pain domain.

		VAS	NRS	VRS	SPS	SPADI Pain	OSS
		M/A	M/A	M/A	M/A	M/A	M/A
Body Functions	b1 Mental functions						
	b1349 Sleep functions, unspecified				1		
	b2 Sensory functions and pain						
	b2801 Pain in body part				1		
	b28016 Pain in joints	1	1	1	4 1	5	4 2
	b2804 Radiating pain in a segment or region				1		
Activities and Participation	d4 Mobility						
	d4150 Maintaining a lying position				1	1	
	d4301 Carrying in the hands						1
	d4451 Pushing					1	
	d4452 Reaching					1	
	d4453 Turning or twisting the hands or arms						1
	d4458 Hand and arm use, other specified					1	
	d4701 Using private motorized transportation						1
	d4702 Using public motorized transportation						1
	d5 Self-care						
	d5100 Washing body parts						1
	d5102 Drying oneself						1
	d5202 Caring for hair						1
	d5409 Dressing, unspecified						1
	d6 Domestic life						
	d6200 Shopping						1
	d6408 Doing housework, other specified						1
	d6409 Doing housework, unspecified						1
	d8 Major life areas						
	d8509 Remunerative employment, unspecified						1
	Not definable				1		2
	Not covered by the ICF				2	1	
	Sum of linked ICF categories	1	1	1	12	10	20

ICF: International Classification of Functioning, Disability, and Health; M/A: main and additional concepts; NRS: numerical rating scale; OSS: Oxford Shoulder Score; SPADI: Shoulder Pain and Disability Scale, Pain subscale; SPS: Shoulder Pain Scale; VAS: visual analog scale; VRS: verbal rating scale.

information about the context in which the pain is assessed. In the 3 overall pain scales, no additional concepts were identified. In the SPS, 3 additional concepts were not sufficiently specified to be classified in the ICF (at rest, in motion, and nightly), whereas in the SPADI pain subscale, pain was measured in the context of 4 different mobility activities. Of the 4 main concepts in the OSS that assessed pain, 2 were provided without any additional concepts, one was linked to a mobility category, and another was assigned to “not definable.”

“Physical function/activity” candidate instruments. The analysis of perspectives in the candidate instruments showed that a descriptive performance perspective was adopted in all 7 instruments. With respect to the response options, 4 instruments including the Penn Shoulder Score, Function Subscale (Penn), the L’Insalata Shoulder Rating Questionnaire (SRQ), the American Shoulder and Elbow Surgeons Shoulder Outcome Score (ASES), and the SPADI disability subscale, assessed “Intensity”; the Simple Shoulder Test (SST) and the University of California at Los Angeles Shoulder Score (UCLA) “Confirmation/agreement,” and the SFA scale “Qualitative attributes.”

The instruments varied with respect to the depth and breadth

of information (Table 4). The additional concepts in the physical function/activity measures were often used for specifying the content, and thus should be interpreted differently than in the pain candidate instruments. All physical function/activity candidate instruments included concepts linked to self-care ICF categories, and all except one, the SFA, included concepts linked to both self-care and mobility.

The Penn was the most wide-ranging instrument with concepts linked to categories in 5 chapters of the activities and participation component of the ICF. In particular, the Penn comprehensively covers mobility, self-care, and domestic life (23 of 27 main concepts). It is also worth noting that the Penn included 4 main concepts linked to a sleep category, which is classified as body functions in the ICF, and also linked to work and leisure activities in the activities and participation component of the ICF. Similarly, the ASES covered mobility and self-care comprehensively, but it also included concepts linked to sleep functions and to work and leisure activities. In the SST, 8 of 15 concepts were linked to mobility categories and the rest to work, sleep, and pain categories in the ICF.

Two instruments, the SRQ and the SPADI disability subscale,

Table 4. Overview of linked ICF categories in the 7 candidate instruments of the physical function/activity domain.

			Penn	SRQ	SST	ASES	UCLA	SPADI Dis	SFA
			M/A	M/A	M/A	Main and Additional Concepts	M/A	M/A	M/A
Body Functions	b1 Mental functions								
	b1348 Sleep functions, other specified			1	1				
	b1349 Sleep functions, unspecified	1		1					
	b2 Sensory functions and pain								
	b2801 Pain in body part				1				
	b28016 Pain in joints			1 2					
	b7 Neuromusculoskeletal and movement related functions								
	b7108 Mobility of joint functions, other specified								
	b7301 Power of muscles of 1 limb								
Activities and Participation	d4 Mobility								
	d4300 Lifting		1	2	1				
	d4301 Carrying in the hands	2	1	1					
	d4302 Carrying in the arms						1		
	d4452 Reaching		1		1				
	d4454 Throwing	1		2	1				
	d4458 Hand and arm use, other specified	8		2			2		
	d4459 Hand and arm use, unspecified			1		2			
	d4751 Driving motorized vehicles					1			
	d4759 Driving, unspecified		1						
	d5 Self-care								
	d5100 Washing body parts	2	1	1	1	1	2	1	
	d5109 Washing oneself, unspecified		1						
	d5202 Caring for hair	1	1		1				1
	d5308 Toileting, other specified				1				
	d5309 Toileting, unspecified	1							1
	d5400 Putting on clothes	1 1	1	1	1	1	3		
	d5401 Taking off clothes	1	1						
	d5409 Dressing, unspecified	1	1		1				1
	d599 Self-care, unspecified		1						
	d6 Domestic life								
	d6200 Shopping					1			
	d6309 Preparing meals, unspecified	1							
	d6400 Washing and drying clothes and garments	1							
	d6402 Cleaning living area	1							
	d6409 Doing housework, unspecified	1							
	d649 Household tasks, other specified and unspecified	1	1			2			
	d699 Domestic life, unspecified		1						
	d8 Major life areas								
	d8509 Remunerative employment, unspecified	1		1					
	d859 Work and employment, other specified and unspecified				1				
	d9 Community, social and civic life								
	d9201 Sports				1				
	d9204 Hobbies	2							
	Not definable					3			1
	Not covered by the ICF	10		2					
	Sum of linked ICF categories	38	13	18	12	11	8	5	

ASES: American Shoulder and Elbow Surgeons Shoulder Outcome Score activities of daily living subscale; ICF: International Classification of Functioning, Disability, and Health; M/A: Main and additional concepts; Penn: Penn Shoulder Score function subscale; SFA: Shoulder Function Assessment scale activities of daily living subscale; SPADI dis: Shoulder Pain and Disability Index disability subscale; SRQ: L'Insalata Shoulder Rating Questionnaire daily activities subscale; SST: Simple Shoulder Test; UCLA: University of California at Los Angeles Shoulder Score function subscale.

covered mobility and self-care comprehensively. In the SRQ, the content was linked to 3 mobility categories and 7 different self-care categories. In addition, 2 concepts were linked to domestic life activities. The SPADI disability subscale had concepts linked

to 4 mobility categories and 5 self-care categories (of these, only 2 are unique).

In the 2 last instruments, the SFA and the UCLA, the meaningful concepts were identified in the response options. For the

UCLA, these concepts were linked to mobility, self-care, and domestic life categories in the ICF, and for the SFA to self-care categories.

DISCUSSION

The ICF content analysis showed a large variation in content and measurement constructs in the candidate instruments for pain and physical function/activity for the shoulder core outcome measurement set.

Among the 6 pain candidate instruments, all included concepts linked to a pain category in the ICF. However, 2 of the instruments, the SPS and the OSS, also covered sleep functions, mobility, self-care, and domestic life activities. This was particularly prominent in the OSS, where more than two-thirds of the items covered concepts other than pain.

In pain assessments, it is important to take into account the context in which the pain is experienced. This is consistent with the definition of pain in the shoulder core set, relating pain experiences to a given context ("i.e., at rest, during, and after activity")¹³. The only candidate instrument where all main concepts cover pain and at the same time refer to a specific context was the SPADI Pain. It should, however, be noted that all except 1 SPADI item measure pain in the context of performing hand and arm mobility activities. In addition, a single item measures pain at its worst. Thus, one of its items measures pain in relation to self-care or domestic life activities, not pain at rest.

The overall pain candidate scales, the VAS, NRS, and VRS, measure the magnitude of the pain regardless of any contextual information. Because of the vagueness in construct definition, it has been recommended that such scales can only complement and not replace genuine, validated pain scales³⁷.

Based on our ICF analysis, no single candidate instrument completely matches the magnitude of the pain experience, as defined in the shoulder core outcome set¹³. However, the use of the SPADI Pain in combination with an overall pain scale (VAS, NRS, or VRS), might provide an acceptable coverage of the pain domain. Moreover, the documented inconsistencies in the content of the SPS and OSS should be considered in the further discussions regarding which pain instruments to be included in the core set.

Seven candidate instruments in the physical function/activity domain were included in the ICF content analysis. As defined in the core set, this domain covers functions ranging from self-care (e.g., bathing, combing hair) to more complex activities (e.g., driving a car)¹³. Our analysis showed that a majority of the candidate instruments cover mobility and self-care activities, which matches the domain definition of the core set^{13,16}. Nevertheless, a majority of the candidate instruments also cover content that falls outside the domain definition. In particular, 1 instrument, the Penn, included content from 5 of 9 chapters within the activities and participation component and content that was linked to the body functions component of the ICF. A similar content coverage was found in the SST and the ASES. This wider content coverage, as provided by the Penn, SST, and

ASES, is supported by empiric evidence showing that patient-reported problems are frequently reported within a range of body functions and activities and participation chapters³⁸.

The candidate instruments that provided the best match with physical function/activity were the SRQ and the SPADI Disability. Both instruments covered mobility and self-care activities, and included little additional content. Although both instruments had a similar content profile, an important difference was discovered: while the SRQ covers a range of self-care activities, the SPADI Disability only included 2 such activities. It should also be noted that only 6 of the 15 items in the full version of the SRQ were selected as candidates for the shoulder outcome measurement set. From our previous content analyses of shoulder pain instruments, we learned that the full version of the SRQ covers similar ICF domains as the most wide-ranging candidate instrument, the Penn²¹.

The 2 last candidate instruments, the UCLA Shoulder Score and the SFA have little or no content that address mobility activities of the hand and arm. In addition, they have a structure that implies that the meaningful concepts are included in the response options and not in the item itself. This limitation needs to be considered in the ongoing selection process.

Our ICF analysis showed that a majority of the physical function/activity candidate instruments had content that did not perfectly match the OMERACT domain definition¹³. In addition to mobility and self-care activities, most of the measures covered content belonging to pain and 2 optional core set domains, participation (recreation and work) and sleep¹³. There were also examples of domestic life activities (e.g., household tasks) in the instruments that are not included in any of the recommended core set domains¹³.

We suggest that the lack of alignment between the definition of physical function/activity in the shoulder core outcome set, and the content of the candidate measures needs further consideration by the OMERACT Shoulder working group. The group could consider either adjusting the domain definition or not including instruments that do not comply with the current definition. In this undertaking, the consensus-based guidelines for selection of outcome measurement instruments, developed as a joint initiative between the Core Outcome Measures in Effectiveness Trials (COMET) initiative and the COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) initiative, will be useful^{23,39}.

A limitation of our study was that some of the measures included content that neither could be defined nor is covered by the ICF. Because of this, the results do not provide a complete overview of the content in the measures.

The ICF-based analysis of the candidate instruments within the mandatory pain and physical function/activity domains of the OMERACT shoulder core outcome set showed large variations in the content and measurement constructs covered. Two of 6 pain instruments include constructs other than pain. Within physical function/activity, 2 candidate instruments matched the domain, 3 included additional content, and the

last 2 instruments included meaningful concepts in the response options, suggesting that they should be omitted as candidate instruments. The analyses show that the content in most existing instruments of shoulder pain and functioning extends across core set domains.

REFERENCES

1. Luime JJ, Koes BW, Hendriksen IJ, Burdorf A, Verhagen AP, Miedema HS, et al. Prevalence and incidence of shoulder pain in the general population; a systematic review. *Scand J Rheumatol* 2004;33:73-81.
2. van der Windt DA, Koes BW, Boeke AJ, Deville W, De Jong BA, Bouter LM. Shoulder disorders in general practice: Prognostic indicators of outcome. *Br J Gen Pract* 1996;46:519-23.
3. Largacha M, Parsons IM, Campbell B, Titelman RM, Smith KL, Matsen F 3rd. Deficits in shoulder function and general health associated with sixteen common shoulder diagnoses: A study of 2674 patients. *J Shoulder Elbow Surg* 2006;15:30-9.
4. Mather RC 3rd, Koenig L, Acevedo D, Dall TM, Gallo P, Romeo A, et al. The societal and economic value of rotator cuff repair. *J Bone Joint Surg Am* 2013;95:1993-2000.
5. Page MJ, Voshaar MJH, Shea B, Verhagen AP, Whittle SL, van der Windt DA, et al. Patients' experience of shoulder disorders: A systematic review of qualitative studies for the OMERACT shoulder core domain set. *Rheumatology* 2019;58:1410-21.
6. van der Windt DA, Koes BW, de Jong BA, Bouter LM. Shoulder disorders in general practice: Incidence, patient characteristics, and management. *Ann Rheum Dis* 1995;54:959-64.
7. Virta L, Joranger P, Brox JI, Eriksson R. Costs of shoulder pain and resource use in primary health care: A cost-of-illness study in Sweden. *BMC Musculoskelet Disord* 2012;13:17.
8. Page MJ, Huang H, Verhagen AP, Gagnier JJ, Buchbinder R. Outcome reporting in randomized trials for shoulder disorders: Literature review to inform the development of a core outcome set. *Arthritis Care Res* 2018;70:252-9.
9. Buchbinder R, Page MJ, Huang H, Verhagen AP, Beaton D, Kopkow C, et al. A preliminary core domain set for clinical trials of shoulder disorders: A report from the OMERACT 2016 shoulder core outcome set special interest group. *J Rheumatol* 2017;44:1880-3.
10. Gagnier JJ, Page MJ, Huang H, Verhagen AP, Buchbinder R. Creation of a core outcome set for clinical trials of people with shoulder pain: A study protocol. *Trials* 2017;18:336-45.
11. Boers M, Beaton DE, Shea BJ, Maxwell LJ, Bartlett SJ, Bingham CO, et al. OMERACT filter 2.1: Elaboration of the conceptual framework for outcome measurement in health intervention studies. *J Rheumatol* 2019;46:1021-7.
12. Boers M, Kirwan JR, Wells G, Beaton D, Gossec L, d'Agostino MA, et al. Developing core outcome measurement sets for clinical trials: OMERACT filter 2.0. *J Clin Epidemiol* 2014;67:745-53.
13. Ramiro S, Page MJ, Whittle SL, Huang H, Verhagen AP, Beaton D, et al. The OMERACT core domain set for clinical trials of shoulder disorders. *J Rheumatol* 2019;46:969-75.
14. Huang H, editor. OMERACT filter 2.1 – item 1 and 2: Truth and feasibility test. Domain of interest – physical function/activity. OMERACT - Outcome Measures in Rheumatology Conference; 2018; New South Wales, Australia.
15. Whittle S, editor. Assessment of pain in clinical trials of shoulder disorders. OMERACT - Outcome Measures in Rheumatology Conference; 2018; New South Wales, Australia.
16. World Health Organization (WHO). International classification of functioning, disability and health (ICF). WHO, Geneva (Switzerland); 2001.
17. Cieza A, Brockow T, Ewert T, Amman E, Kollerits B, Chatterji S, et al. Linking health-status measurements to the international classification of functioning, disability and health. *J Rehabil Med* 2002;34:205-10.
18. Cieza A, Geyh S, Chatterji S, Kostanjsek N, Ustun B, Stucki G. ICF linking rules: An update based on lessons learned. *J Rehabil Med* 2005;37:212-8.
19. Brockow T, Cieza A, Kuhlow H, Sigl T, Franke T, Harder M, et al. Identifying the concepts contained in outcome measures of clinical trials on musculoskeletal disorders and chronic widespread pain using the international classification of functioning, disability and health as a reference. *J Rehabil Med* 2004;44 Suppl:30-6.
20. Grotle M, Brox JI, Vollestad NK. Functional status and disability questionnaires: What do they assess? A systematic review of back-specific outcome questionnaires. *Spine* 2005;30:130-40.
21. Roe Y, Østensjø S. Conceptualization and assessment of disability in shoulder-specific measures with reference to the international classification of functioning, disability and health. *J Rehabil Med* 2016;48:325-32.
22. Cieza A, Fayed N, Bickenbach J, Prodinger B. Refinements of the ICF linking rules to strengthen their potential for establishing comparability of health information. *Disabil Rehabil* 2019;41:574-83.
23. Prinsen CA, Vohra S, Rose MR, Boers M, Tugwell P, Clarke M, et al. How to select outcome measurement instruments for outcomes included in a "core outcome set" – a practical guideline. *Trials* 2016;17:449.
24. Dawson J, Fitzpatrick R, Carr A. Questionnaire on the perceptions of patients about shoulder surgery. *J Bone Joint Surg Br* 1996;78:593-600.
25. Dawson J, Rogers K, Fitzpatrick R, Carr A. The Oxford shoulder score revisited. *Arch Orthop Trauma Surg* 2009;129:119-23.
26. Roach KE, Budiman-Mak E, Songsiridej N, Lertratanakul Y. Development of a shoulder pain and disability index. *Arthritis Care Res* 1991;4:143-9.
27. Williams JW Jr, Holleman DR Jr, Simel DL. Measuring shoulder function with the shoulder pain and disability index. *J Rheumatol* 1995;22:727-32.
28. Winters JC, Sobel JS, Groenier KH, Arendzen JH, Meyboom-De Jong B. A shoulder pain score: A comprehensive questionnaire for assessing pain in patients with shoulder complaints. *Scand J Rehabil Med* 1996;28:163-7.
29. Leggin BG, Michener LA, Shaffer MA, Brenneman SK, Iannotti JP, Williams GR Jr. The Penn shoulder score: Reliability and validity. *J Orthop Sports Phys Ther* 2006;36:138-51.
30. L'Insalata JC, Warren RF, Cohen SB, Altchek DW, Peterson MG. A self-administered questionnaire for assessment of symptoms and function of the shoulder. *J Bone Joint Surg Am* 1997;79:738-48.
31. Lippitt SB, Harryman DT II. A practical tool for evaluation of function: The simple shoulder test. In: Matsen FI, Fu F, Hawkins R, editors. *The shoulder: A balance of mobility and stability*. Rosemont: The American Academy of Orthopedic Surgeons; 1993:501-18.
32. Richards RR, An KN, Bigliani LU, Friedman RJ, Gartsman GM, Gristina AG, et al. A standardized method for the assessment of shoulder function. *J Shoulder Elbow Surg* 1994;3:347-52.
33. Ellman H, Hanks G, Bayer M. Repair of the rotator cuff. End-result study of factors influencing reconstruction. *J Bone Joint Surg Am* 1986;68:1136-44.
34. van Den Ende CH, Rozing PM, Dijkman BA, Verhoef JA, Voogt-van der Harst EM, Hazes JM. Assessment of shoulder

- function in rheumatoid arthritis. *J Rheumatol* 1996;23:2043-8.
35. McHugh ML. Interrater reliability: The kappa statistic. *Biochem Med* 2012;22:276-82.
36. Cicchetti DV, Sparrow SA. Developing criteria for establishing interrater reliability of specific items: Applications to assessment of adaptive behavior. *Am J Ment Defic* 1981;86:127-37.
37. Franchignoni F, Salaffi F, Tesio L. How should we use the visual analogue scale (VAS) in rehabilitation outcomes? I: How much of what? The seductive VAS numbers are not true measures. *J Rehabil Med* 2012;44:798-9.
38. Roe Y, Bautz-Holter E, Juel NG, Soberg HL. Identification of relevant international classification of functioning, disability and health categories in patients with shoulder pain: A cross-sectional study. *J Rehabil Med* 2013;45:662-9.
39. Williamson PR, Altman DG, Bagley H, Barnes KL, Blazeby JM, Brookes ST, et al. The COMET handbook: Version 1.0. *Trials* 2017;18:280.