

Full title of manuscript:

What do the OMERACT Shoulder Core Set candidate instruments measure? An analysis using the refined ICF linking rules

Complete given names and surnames of all authors with ORCID ID if any:

Yngve Røe

<https://orcid.org/0000-0003-2428-571X>

Rachelle Buchbinder

<https://orcid.org/0000-0002-0597-0933>

Margreth Grotle

<https://orcid.org/0000-0001-8243-1143>

Samuel Whittle

<https://orcid.org/0000-0001-7417-7691>

Sofia Ramiro

<https://orcid.org/0000-0002-8899-9087>

Hsiaomin Huang

Scopus Author Identifier: 56883268800

Joel Gagnier

<https://orcid.org/0000-0002-3162-3935>

Arianne Verhagen

<https://orcid.org/0000-0002-6195-0128>

Sigrid Østensjø

<https://orcid.org/0000-0002-7418-2973>

Key Indexing Terms:

OMERACT, Outcome assessment, ICF, Shoulder pain

Name of department(s) and institution(s) to which the work should be attributed:

Department of Physiotherapy, Oslo Metropolitan University, Oslo, Norway

Department of Epidemiology and Preventive Medicine, Monash University, Sydney, Australia

The Queen Elizabeth Hospital, Woodville South, Australia

Department of Epidemiology and Orthopaedic Surgery, University of Michigan, USA

University of Technology, Melbourne, Australia

The source(s) of support in the form of grants or industrial support:

None

Conflict of interest:

None

Initials, surnames, appointments, and highest academic degrees of all authors:

Y. Røe, Associated Professor, PhD

R. Buchbinder, Professor, PhD

M. Grotle, Professor, PhD

S. Whittle, MD, PhD

S. Ramiro, MD, PhD

H. Huang, Associated Professor, PhD

J. Gagnier, Associated Professor, PhD

A. Verhagen, Associated Professor, PhD

S. Østensjø, Professor, PhD

Name, address, and e-mail of author responsible for correspondence:

Yngve Røe

OsloMet – Oslo Metropolitan University

Department of Physiotherapy

P.O. Box 4, St. Olavs plass

N-0130 Oslo

Norway

yngveroe@oslomet.no

Short running head:

Shoulder candidate instruments

Abstract

Objective: 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. The results of this ICF based analysis may inform further decisions on which instruments finally to 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 International Classification of Functioning, Disability and Health (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 large variation in content and measurement constructs in the candidate instruments for the Shoulder Core Outcome Measurement Set.

Conclusion: Two of six ‘Pain’ instruments include other constructs than pain. Within ‘Physical function/activity’ two candidate instruments matched the domain, three included additional content and the two last 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 extend across Core Set domains.

Introduction

Shoulder pain is a common musculoskeletal disorder with an incidence of 10 per 1000 in primary care and point prevalence estimates of 7-26% in the general population (1). Shoulder disorders can be long lasting; in a Dutch study in those patients presenting a new episode of shoulder pain for their general practitioner, a considerable number (41%) showed persistent symptoms after 12 months (2). The associated disability and impact in terms of earnings, missed workdays, and disability payments, is 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, have been advocated (8). 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 delegates (13). It consisted of four mandatory domains for all trials of shoulder disorders: 'Pain', 'Physical function/activity', 'Patient global – shoulder', and 'Adverse events including death'; and four important but optional domains: 'Participation (recreation/work)', 'Sleep', 'Emotional wellbeing' and 'Condition-specific pathophysiological manifestations' (13). The next phase will be to recommend specific measurement instruments for a Core Outcome Measurement Set (10).

Preliminary work has investigated instruments within two 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 disorders (8). Pain was defined as "How much a person's shoulder hurts, reflecting the overall magnitude of the pain experience (i.e., at rest, during and after activity, at night)" and physical function/activity as "A person's ability to carry out daily physical activities,

Accepted Article

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 the Outcome Measures in Rheumatology (OMERACT) (11, 12). Altogether, six instruments in the pain domain and seven within the physical function/activity domain, passed both filters and are candidates for further assessment (14, 15). However five in the ‘Pain’ domain and three 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 WHO framework for measuring health and disability (16). 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 ten linking rules were published in 2002, and updated in 2005 (17, 18). Since their introduction, a number of instruments has 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 place (22). Thus, content linking of outcome measure instruments based on the refined ICF linking rules provide information on important aspects of content validity. Content validity is considered to be the most important measurement property of an outcome measure instrument, because if it is unclear what an instrument is actually measuring, the assessment of other measurement properties may be irrelevant (23).

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 finally to 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 Outcome Measures in Rheumatology (OMERACT) filtering process (14, 15). The six candidate instruments within the pain domain and the seven within the physical function/activity domain are presented in Table 1. These instruments are widely used in clinical and epidemiological research of shoulder pain conditions (8).

Please insert Table 1, about here

Analysis of content and measurement constructs

The International Classification of Functioning, Disability and Health (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) and ‘Environmental factors’ (e) and ‘Personal Factors’ (not classified) (16). The ICF provides four sub-

classification (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 letters is followed by a numeric code (one-digit) (e.g., d4 Mobility), two 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.

Please insert Table 2, about here

The content from each item in the measurement instruments was linked to the ICF according to the 10 refined linking rules (22). Linking rules #1–3 specify how to get familiar with the ICF, identifying the purpose of an instrument and concepts to be linked to the ICF. Both the researchers who conducted the analyses (YR & SØ) had previously linked the content of shoulder pain instruments to the ICF (21).

First, the actual meaning (main and additional concepts) of the information to be linked was identified, consistent with rules #2 and 3 (22). When identifying the concepts, both the item text and the text that set premises for the interpretation of the item content were taken into consideration. For most items it was straightforward to identify main and additional concepts. Such as 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 one activity was listed in the same item. When this was the case, all 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 UCLA and the “Activities of daily living subscale” of the SFA (32, 34). Then the naming of the item was 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 the needs or dependency perspectives (22). 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 task or activities in the natural environment. According to linking rule #5, the categorisation 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 two researchers (YR and SØ). In 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 2nd ICF category level was calculated with the Cohen’s Kappa coefficient. The 95% confidence intervals for the Kappa coefficient were calculated using the standard error of the kappa:

$k-1.96 \times SE_k$ to $k+1.96 \times SE_k$ (35). The calculated Kappa coefficient of the linking of main and additional concepts was 0.85 (95% CI, 0.78 - 0.91), 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

Results

Descriptive information about the six 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 six instruments. The response options in the Visual Analogue 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 International Classification of Functioning, Disability and Health (ICF) (Table 3). For the three 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 five categories.

In addition to pain categories, the SPS included a main concept linked to a ‘Mobility’ category in the ICF. An 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 information about the context in which the pain is assessed. In the three overall pain scales, no additional concepts were identified. In the SPS, three 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 four different ‘Mobility’ activities. Of the four main concepts in the OSS that assessed pain, two were provided without any additional concepts, one was linked to a ‘Mobility’ category and another was assigned to ‘not definable’.

Please insert Table 3, about here

‘Physical function/activity’ candidate instruments

The analysis of perspectives in the candidate instruments showed that a ‘Descriptive-performance’ perspective was adopted in all seven instruments. With respect to the response options, four instruments, 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 Shoulder Pain And Disability Index (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 Shoulder Function Assessment scale (SFA) ‘Qualitative attributes’.

The instruments varied with respect to the depth and breadth of information (see 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 five 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 four main concepts linked to a sleep category which is classified as ‘Body Functions’ in the ICF, and also 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, covered ‘Mobility’ and ‘Self-care’ comprehensively. In the SRQ, the content was linked to three ‘Mobility’ categories and seven different ‘Self-care’ categories. In addition, two concepts were linked to ‘Domestic life’ activities. Another instrument, the SPADI disability subscale had concepts linked to three ‘Mobility’ categories and five ‘Self-care’ categories (of these only two unique).

In the two 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.

Please insert Table 4, about here

Discussion

The International Classification of Functioning, Disability and Health (ICF) content analysis showed 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 six pain candidate instruments, all included concepts linked to a pain category in the ICF. However, two of the instruments, the Shoulder Pain Score (SPS) and the Oxford Shoulder Score (OSS), also covered sleep functions, and mobility, self-care and domestic life activities. This was particularly prominent in the OSS, where more than two thirds of the items covered other concepts 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 Shoulder Pain And Disability Index (SPADI pain). It should, however, be noted that all except one SPADI items measure pain in the context of performing hand and arm ‘Mobility’ activities. In addition, a single item requests pain at its worst. Thus, one of its items measure pain in relation to self-care or domestic life activities, nor pain at rest.

The overall pain candidate scales, the Visual Analogue Scale (VAS), the Numerical Rating Scale (NRS) and the Verbal Rating Scale (VRS), measure the magnitude of the pain, regardless of any contextual information. Due to the vagueness in construct definition, it has

been recommended that such scales can only complement and not replace genuine, validated pain scales (37).

Based on our ICF analysis, no single candidate instrument completely matches the magnitude of the pain experience, as defined in the Shoulder Core Outcome Set (13). 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 content of the SPS and OSS should be considered in the further discussions on 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 one instrument, the Penn Shoulder Score (Penn), included content from five of nine 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 Simple Shoulder Test (SST) and the American Shoulder and Elbow Surgeons Shoulder Outcome Score (ASES). This wider content coverage, as provided by the Penn, SST and ASES, are supported by empirical evidence showing that patient-reported problems are frequently reported within a range of body functions and activities and participation chapters (38).

The candidate instruments that provided the best match with 'Physical function/activity', was the Shoulder Rating Questionnaire (SRQ and) the Shoulder Pain and Disability Index (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 only included two 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 Penn (21).

The two last candidate instruments, the University of California at Los Angeles Shoulder Score (UCLA) and the Shoulder Function Assessment scale (SFA) have little or no content that address ‘Mobility’ activities of the hand and arm. In addition, they have a structure that imply that the meaningful concepts are included in the response options, and not in the item itself. This limitation need 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 definition (13). In addition to mobility and self-care activities, most of the measures covered content belonging to ‘Pain’ and to two ‘optional’ Core Set domains, ‘Participation (recreation and work)’ and ‘Sleep’ (13). 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 domains (13).

We suggest that the lack of alignment between the definition of ‘Physical function/activity’ in 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 work, 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 not be defined, nor is covered by the ICF. Due to this, the results do not provide a complete overview of the content in the measures.

Conclusion

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 six ‘Pain’ instruments include other constructs than pain. Within ‘Physical function/activity’ two candidate instruments matched the domain, three included additional content and the two last 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 extend across Core Set domains.

Acknowledgement

Not applicable

Funding

None

Competing interests

The authors have no competing interests to declare

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 IMt, 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 (Oxford)* 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 JJ, 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;44:1880-83.
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;46:1021-27.
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;01:01.
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, Kuhlman 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;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* 2016;1-10.
23. Prinsen CAC, 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.

- Accepted Article
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 Of Rehabil Med* 1996;28:163-7.
 29. Leggin BG, Michener LA, Shaffer MA, Brennehan 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 DTII. 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, Illinois: The American Academy of Orthopedic Surgeons; 1993. p. 501-18.
 32. Richards RR, An K-N, 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, Hunker 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, Dijkmans BA, Verhoef JA, Voogt-van der Harst EM, Hazes JM. Assessment of shoulder function in rheumatoid arthritis. *Journal 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.

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 Analogue Scale	VAS	1	Visual analogue 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 analogue pain scale or 10-point ordinal scale (0-100)
Shoulder Pain Score (28)	SPS	7	Item 1-5, 4-point ordinal scale. Item 6: Visual analogue 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 analogue 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 item 6-11 were selected as candidates for the OMERACT core measurement instrument set

Table 2: Overview of the ICF domain (chapter) structure within the sub-classifications of ‘Body Functions’ and ‘Activities and Participation’

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

Table 3: Overview of linked ICF categories in the six candidate instruments of the ‘Pain’ domain

		VAS	NRS	VRS	SPS	SPADI pain	OSS
		Main and additional concepts (M/A)					
		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 (nd)				1		2
	Not covered by the ICF (nc)				2	1	
	Sum of linked ICF categories	1	1	1	12	10	20

Visual Analogue Scale (VAS), Numerical Rating Scale (NRS), Vocal Rating Scale (VRS), Shoulder Pain Scale (SPS), Shoulder Pain and Disability Scale pain subscale (SPADI pain), Oxford Shoulder Score (OSS)

Table 4: Overview of linked ICF categories in the seven candidate instruments of the ‘Physical function/activity’ domain

	Penn	SRQ	SST	ASES	UCLA	SPADI dis	SFA	
	Main and additional concepts (M/A)							
	M/A	M/A	M/A	M/A	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 movementrelated functions							
	b7108 Mobility of joint functions, other specified							
b7301 Power of muscles of one 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	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 (nd)					3		1
	Not covered by the ICF (nc)	10		2				
Sum of linked ICF categories	38	13	18	12	11	8	5	

Penn Shoulder Score function subscale (Penn), L'Insalata Shoulder Rating Questionnaire daily activities subscale (SRQ), Simple Shoulder Test (SST), American Shoulder and Elbow Surgeons Shoulder Outcome Score activities of daily living subscale (ASES), University of California at Los Angeles Shoulder Score function subscale (UCLA), Shoulder Pain and Disability Index disability subscale (SPADI dis), Shoulder Function Assessment scale activities of daily living subscale (SFA)