RMD Open
Rheumatic & Musculoskeletal Diseases

ORIGINAL ARTICLE

Identifying a core set of outcome domains to measure in clinical trials for shoulder disorders: a modified Delphi study

Matthew J Page,[1,2] Hsiaomin Huang,[3] Arianne P Verhagen,[4] Rachelle Buchbinder,[5,6] Joel J Gagnier[3,7]

To cite: Page MJ, Huang H, Verhagen AP, *et al*. Identifying a core set of outcome domains to measure in clinical trials for shoulder disorders: a modified Delphi study. *RMD Open* 2016;**2**: e000380. doi:10.1136/rmdopen-2016-000380

► Prepublication history and additional material is available. To view please visit the journal (http://dx.doi.org/10.1136/rmdopen-2016-000380).

RB and JJG Contributed equally as senior authors of this work.

Received 7 October 2016
Revised 17 November 2016
Accepted 28 November 2016

For numbered affiliations see end of article.

Correspondence to
Dr Joel Gagnier;
jgagnier@umich.edu

ABSTRACT

Objective: To achieve consensus on the most important outcome domains to measure across all clinical trials for shoulder disorders.

Methods: We performed an online modified Delphi study with an international, multidisciplinary and multistakeholder panel. A literature review and the OMERACT Filter 2.0 framework was used to generate a list of potential core domains, which were presented to patients, clinicians and researchers in two Delphi rounds. Participants were asked to judge the importance of each potential core domain and provide a rationale for their response. A core domain was defined a priori as a domain that at least 67% of participants considered core.

Results: In both rounds, 335 individuals were invited to participate (268 clinicians/researchers and 67 patients); response rates were 27% (n=91) and 29% (n=96), respectively. From a list of 41 potential core domains, four domains met our criteria for inclusion: 'pain', 'physical functioning', 'global assessment of treatment success' and 'health-related quality of life'. Two additional domains, 'sleep functioning' and 'psychological functioning', met the criteria for inclusion by some, but not all stakeholder groups. There was consensus that 'number of deaths' was not a core domain, but insufficient agreement on whether or not several other domains, including 'range of motion' and 'muscle strength', were core domains.

Conclusions: Based on international consensus from patients, clinicians and researchers, 'pain', 'physical functioning', 'global assessment of treatment success' and 'health-related quality of life' were considered core outcome domains for shoulder disorder trials. The value of several other domains needs further consideration.

Key messages

What is already known about this subject?

- There is a lack of uniformity in the outcomes measured in shoulder trials.

What does this study add?

- In a modified Delphi study, a majority of patients, clinicians and researchers identified 'pain', 'physical functioning', 'global assessment of treatment success' and 'health-related quality of life' as important domains to include in a core outcome set (COS) for shoulder disorders.
- The value of several other domains, including 'sleep functioning', 'psychological functioning', 'range of motion' and 'muscle strength', needs further consideration.

How might this impact on clinical practice?

- Uptake of a COS for shoulder disorders in future clinical trials should increase our ability to compare findings between studies and synthesise data in meta-analyses, which may provide more reliable evidence for decision makers.

INTRODUCTION

Shoulder pain is a common, activity-limiting symptom.[1 2] It has various causes, including rotator cuff disease (most commonly), adhesive capsulitis and other disorders of the glenohumeral joint (eg, osteoarthritis, humeral head fracture).[1 3 4] Regardless of the cause, people with different shoulder disorders tend to experience similar outcomes, including pain, problems with performing daily activities such as dressing, reduced participation in recreational activities and disrupted sleep patterns.[5 6] However, there is a lack of uniformity in the outcomes measured in shoulder trials,[7] which limits our ability to compare findings between studies and synthesise data in meta-analyses.

The Outcome Measures in Rheumatology (OMERACT)[8] and Core Outcome Measures in Effectiveness Trials (COMET)[9] initiatives have responded to the problem of outcome

diversity by developing guidance for the creation of core outcome sets (COSs). A COS is a collection of outcome domains (true states or endpoints of interest) recommended for measurement in *all* trials for a particular condition.[9 10] No COS for shoulder disorders currently exists, so the OMERACT Shoulder Core Outcome Set Special Interest Group was established to address this gap.

It is recommended that COS development starts by identifying, through literature review, a list of potential domains to measure, which is then refined via consensus processes.[9 10] We recently completed a review of outcome domains and measurement instruments reported in 409 randomised trials of interventions for shoulder disorders published between 1954 and 2015.[18] We identified 32 domains across these trials. The majority included a measure of pain (90%), range of motion (78%) and physical functioning (71%), while muscle strength was reported in 44% of trials, and imaging outcomes were reported in 21% of trials. Other patient-reported outcome measures (PROMs) such as health-related quality of life were each reported in 15% or fewer trials. Most domains were reported at a similar frequency across different shoulder disorders.[18]

Informed by our previous literature review, this paper presents the results of an international modified Delphi study performed to achieve consensus from patients, clinicians and researchers on which domains should be included in a COS for all shoulder disorders.

METHODS

Our prespecified methods are described elsewhere (Gagnier JJ, Page MJ, Huang H, *et al*. Creation of a core outcome set for clinical trials of patients with shoulder pain: A protocol. Trials 2016. Under review). We predetermined that the COS should be applicable to most common shoulder disorders, including: rotator cuff disease (an umbrella term to classify disorders of the rotator cuff muscles or tendons, including subacromial impingement syndrome, rotator cuff tendinopathy or tendinitis, partial or full rotator cuff tear, calcific tendinitis and subacromial bursitis[11]); adhesive capsulitis; shoulder instability; glenohumeral osteoarthritis; dislocation of the shoulder and proximal humeral or humeral head fractures. It should also be applicable to unspecified shoulder pain, to capture trials in primary care that have included people with shoulder pain without a specific diagnosis or cause.

We used an international online modified Delphi process since it permits involvement of geographically distant and informed individuals, anonymity of responses and lack of interaction among participants to prevent the views of prominent personalities from dominating the group.[12] This project received ethical approval from the University of Michigan (HUM00102228).

Selection of panel members

One author (HH) made a list of corresponding authors of all shoulder trials indexed in PubMed from 2006 to 2015 (identified from our previous literature review).[18] Three authors (RB, JJG and APV) nominated other researchers with expertise in shoulder disorders not identified from PubMed, and clinicians from different disciplines who have clinical experience in managing people with shoulder disorders. Clinicians could invite patients to participate if they had a current or had a previous episode of shoulder pain and understood written English. Patients were compensated US$25 for completion of each Delphi round. The final list of selected participants was kept confidential and known only to the project team members.

Generation of a list of potential core domains

We refined the initial list of 32 potential core domains identified from our previous literature review,[18] through discussion and debate, by considering if there were any important domains missing and if certain domains were too broad or should be aggregated. This resulted in 41 domains being generated for inclusion in round one (table 1; definitions in online supplementary file 1, table S1). The framework endorsed by OMERACT ('OMERACT Filter 2.0') was used to structure the potential core domains under four core areas of health (death, life impact, pathophysiological manifestations, resource use[10]). However, when selecting the threshold for inclusion in the COS, there was no requirement that at least one core domain needed to exist in each of the core areas. To generate definitions for each domain, terminology used in a previous Delphi study to develop a COS for non-specific low-back pain[13] was consulted.

Delphi procedure

Participants were invited via email to complete two Delphi rounds administered via the online software SurveyMonkey (SurveyMonkey, Palo Alto, USA; surveys in online supplementary files S2 and S3). In both rounds, we kept the survey available for 2 weeks, and sent a reminder email 1 week after the initial invitation. Individuals not participating in the first round, who did not explicitly express their desire to opt out, were invited to the subsequent round. This modified Delphi approach has been used in previous studies.[12 13] The study was fully anonymised, in that participants did not know the identities of other individuals in the group, nor the specific answers that any individual provided.

The first round ran from 4 to 18 March 2016. Participants were asked to judge the importance of each potential core domain by answering questions phrased as follows, for example: 'Is the domain "Overall pain" important enough to be included in a core domain set for clinical trials for shoulder conditions?' A definition of the domain was provided under each question. Possible responses included 'Yes', 'No' and 'Unsure/I do not know'. Participants were encouraged to provide a

Table 1 Potential core domains considered for shoulder disorder clinical trials (administered in Delphi round one)

Core area	Domain
Life impact	Pain intensity
	Overall
	At rest
	With activity
	On resisted movement
	During the day
	During the night
	Temporal aspects of pain
	Analgesic use
	Physical functioning
	Health-related quality of life
	Social functioning
	Recreation and leisure activity
	Work ability
	Depression
	Anxiety
	Fear avoidance beliefs
	Sleep functioning
	Fatigue
	Satisfaction with treatment services
	Global assessment of treatment success
	Severity of the main complaint
Resource use	Work productivity
	Healthcare services
	Non-healthcare services
	Requiring reoperation or revision surgery
Pathophysiological manifestations	Range of motion
	Muscle strength
	Muscle tone
	Pain on palpation
	Testing positive on specific tests during physical examination
	Shoulder instability
	Scapular dysfunction
	Proprioception
	Weakness on movement
	Shoulder swelling
	Shoulder posture
	Radiographic outcomes
	Failure of surgery
	Surgical process outcomes
	Haemodynamic variables
Death	Number of deaths

rationale for their answer. We also asked participants whether certain domains should be aggregated because of conceptual overlap, and to suggest core domains that may be missing.

We analysed round one data by calculating frequencies of 'Yes', 'No' and 'Unsure/I do not know' for each domain, based on responses from the total panel and separately for each of the four stakeholder groups (patients, researchers, clinicians, professionals working as clinician and researcher). Those domains to which more than two-thirds (at least 67%) of the total panel chose the response option 'Yes' were considered for inclusion in round two. Domains rated as important to include by fewer than 67% of the total panel were excluded unless there were any significant arguments provided in the rationale for responses that were against the overall trend in frequencies. Participants' suggestions for the aggregation of certain domains and addition of missing core domains were discussed and debated by all authors via teleconference until consensus was reached. One author (MJP or HH) summarised the decisions made at each teleconference and fed this back to all authors for review and refinement.

The second round ran from 4 to 15 April 2016. A feedback report (see online supplementary file S4) summarising the level of consensus for each domain (overall and

Table 2 Characteristics of participants completing the first Delphi round

Characteristic	Number (%)
General characteristics (total panel)	n=91
Gender	
Women	49 (54)
Age	
Median (IQR)	51 (40–58)
Highest education	
Postgraduate degree (MSc, PhD)	56 (62)
Graduated from college/university	18 (20)
Graduated from high school	10 (11)
Graduated from primary school	1 (1)
Other	6 (7)
Stakeholder group	
Clinician	19 (21)
Researcher	13 (14)
Professional working as clinician and researcher	18 (20)
Patient	41 (45)
Clinician/researcher characteristics	n=50
Country of work	
UK	11 (22)
The Netherlands	8 (16)
Australia	7 (14)
USA	7 (14)
Canada	6 (12)
Sweden	2 (4)
Ireland	2 (4)
Turkey	2 (4)
Other (Brazil, Finland, Germany, Italy, Norway)	5 (10)
Field of work	
Orthopaedics	17 (34)
Physiotherapy	15 (30)
Rheumatology	5 (10)
Epidemiology	4 (8)
General practice/family medicine	3 (6)
Other (emergency medicine, physical medicine and rehabilitation)	6 (12)
Research experience	
Conducted at least one clinical trial for a shoulder disorder	40 (80)
Conducted at least one systematic review for a shoulder disorder	23 (46)
Developed or tested a measurement instrument for a shoulder disorder	15 (30)
Participated in development of other COSs	20 (40)
Patient characteristics	n=41
Country of residence	
USA	18 (44)
UK	6 (15)
The Netherlands	4 (10)
Austria	3 (7)
Australia	2 (5)
Germany	2 (5)
Other (Belgium, Canada, Poland, Slovakia)	4 (10)
Shoulder disorder history	
Have sought healthcare for a shoulder disorder at least once	36 (88)
Currently have a shoulder disorder	27 (66)
Duration of current shoulder disorder	
Less than 6 weeks	2 (7)*
Between 6 and 12 weeks	3 (11)*
More than 3 months	15 (56)*
Not stated	7 (26)*

*Denominator is the number of patients who currently have a shoulder disorder (n=27).

per stakeholder group) from round one, but not the individual comments to justify ratings, was emailed to all participants, together with the second survey. Participants were asked to judge whether each of the domains brought forward from round one should indeed be considered core. Participants were informed that we chose to aggregate certain domains based on round one comments, and were then asked to rate whether the new (aggregated) domain should be included in the COS. Participants were also asked whether they agreed that all of the domains excluded following round one should be excluded from the COS. Possible responses were the same as those in the round one survey. We deemed (a priori) that consensus was reached if at least 67% of the total panel and 67% of all stakeholder groups agreed that a domain was core. The results of round two were reviewed and discussed via teleconference.

RESULTS

Participants

In both rounds, 335 individuals were invited to participate (268 clinicians/researchers and 67 patients). In the first round, we received responses from 91 participants (50 clinicians/researchers and 41 patients; response rate 27%) and in the second round, we received responses from 96 participants (55 clinicians/researchers and 41 patients; response rate 29%). We could not determine how many participants completed both Delphi rounds due to the anonymity of the survey.

Demographic characteristics were collected only in round one (table 2), so as not to burden respondents with questions they may have previously answered. In round one, the median age of the 91 participants was 51 years, and 49/91 (54%) were women. Most (56/91 (62%)) had a postgraduate degree. The 50 clinicians/researchers worked in 13 countries (most were from the UK and The Netherlands), and in several fields (orthopaedics and physiotherapy were most common). The majority of clinicians/researchers (40/50 (80%)) had conducted at least one clinical trial for shoulder disorders. Of the patient respondents, almost half lived in the

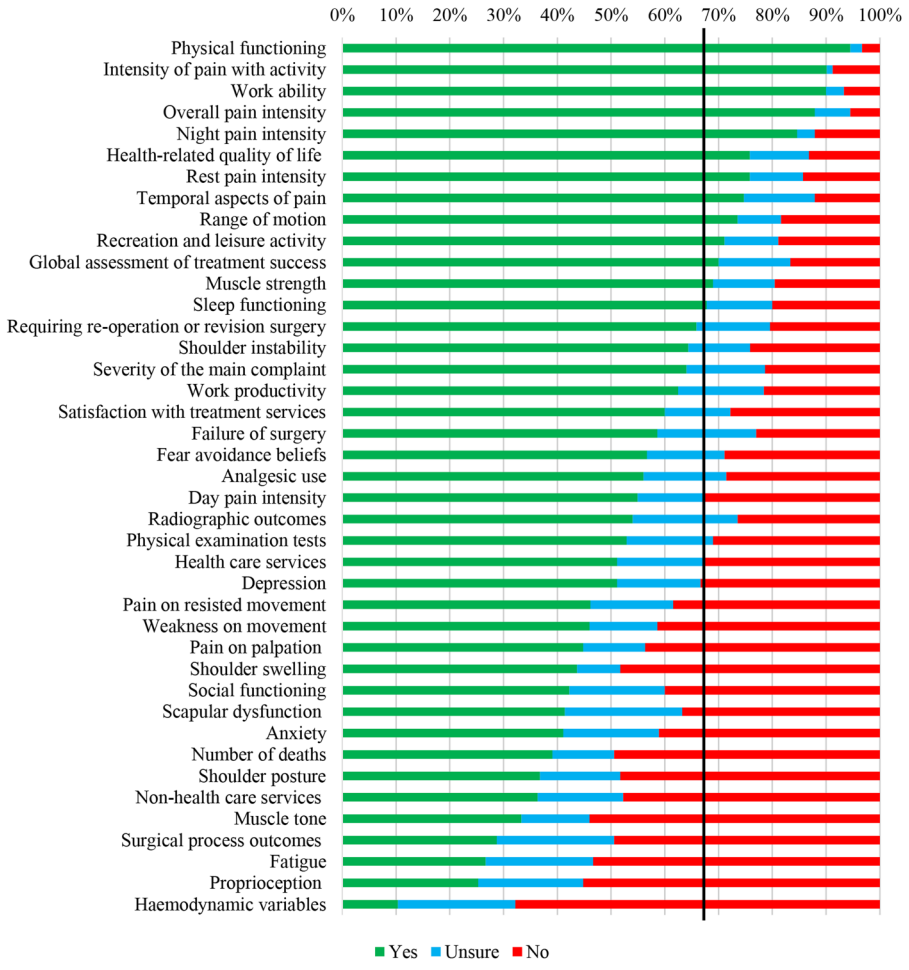

Figure 1 Ratings by total panel (n=91) of potential core domains in the first Delphi round. The solid black line depicts the consensus threshold (67%).

USA (18/41 (44%)), while other participants lived in nine other countries. Most patients (27/41 (66%)) were currently experiencing shoulder pain (table 2).

Delphi round one

In total, 13 domains met our a priori criteria for inclusion (figure 1). They consisted of 'physical functioning', 'intensity of pain with activity', 'work ability', 'overall pain intensity', 'night pain intensity', 'health-related quality of life', 'rest pain intensity', 'temporal aspects of pain', 'range of motion', 'recreation and leisure activity', 'global assessment of treatment success', 'muscle strength' and 'sleep functioning'. However, a clear discrepancy between the patients' perspective and the total panel responses was observed, since 22 domains met the criteria for inclusion in the patient subgroup (see online supplementary file 1, figure S1 and S2). Further, those who worked as a researcher only were more likely than all other stakeholders to consider most domains as not core (see online supplementary file 1, figures S3–S5).

Several patients, clinicians and researchers emphasised the overlap between all of the pain domains (eg, 'intensity of pain with activity', 'night pain') (see all comments to justify ratings in online supplementary file 5). To address this, a proposal was formulated for the second round to combine them all under one domain titled 'pain', defined as 'how much a person's shoulder hurts, reflecting the overall magnitude of the pain experience (ie, at rest, during and after activity, at night)'. Several clinicians and researchers suggested there was overlap between activities of daily living (eg, bathing), recreational activities and work ability. These comments were addressed by creating a single domain called 'physical functioning', with the following revised definition: 'a person's ability to carry out daily physical activities required to meet basic needs (eg, bathing, combing hair), more complex activities that require a combination of skills (eg, driving a car), recreational/leisure activities (eg, sports) and work tasks'. In addition, for round two we combined different measures of psychological functioning (ie, depression, anxiety) into a single domain based on comments from several clinicians and researchers.

We proposed to exclude all domains that were considered by fewer than 67% of the total panel as core. An exception to this was the domain 'number of deaths'. Despite being considered core by only 39% of the total panel, we wanted participants to consider it again given its obvious importance as an outcome event in shoulder disorder trials, and due to the recommendation in the OMERACT Filter 2.0 that deaths be measured in all trials.[10] Also, despite being rated highly by panellists in

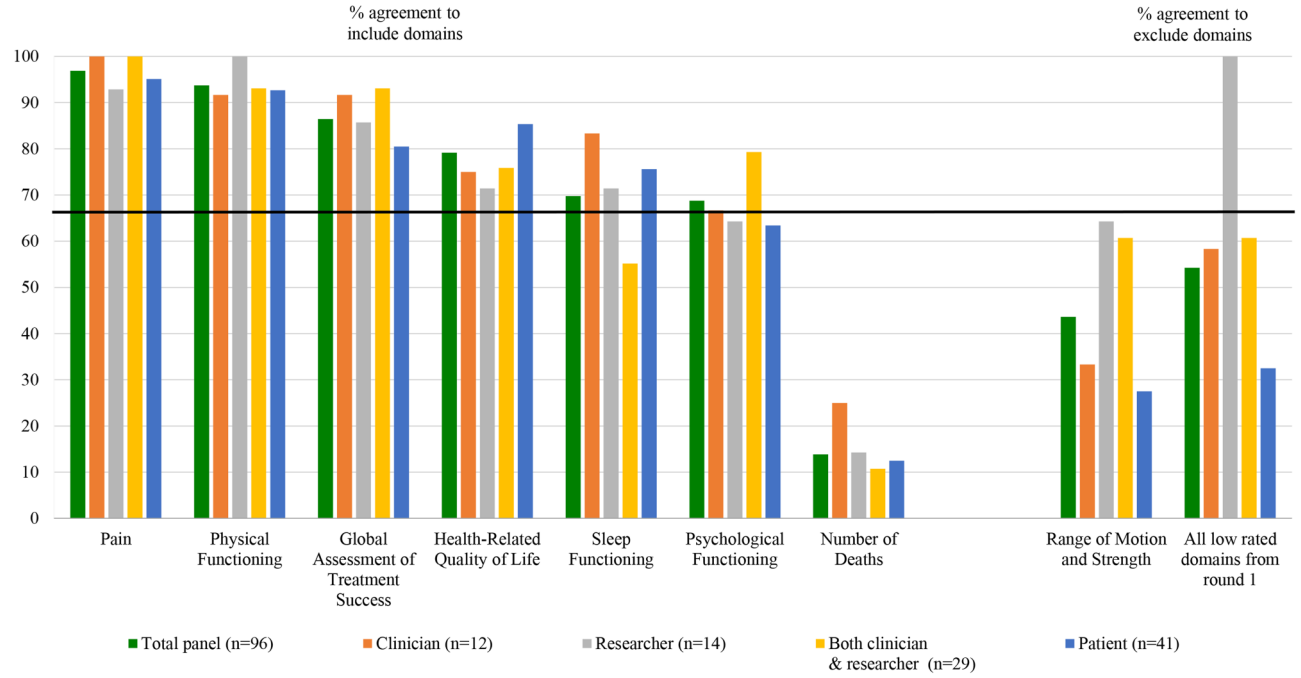

Figure 2 Ratings of core domains stratified by stakeholder group in the second Delphi round. The percentage of participants agreeing that a domain is important to include is presented in the left panel. The percentage of participants agreeing that a domain (or set of domains) should be excluded is presented in the right panel. The solid black line depicts the consensus threshold (67%).

round one, many comments emphasised that 'range of motion' and 'muscle strength' are surrogate measures of functional ability and do not necessarily relate well to functional tasks participants report being able to perform. These comments prompted a proposal for the second round to exclude both domains. Participants did not suggest any new domains.

Delphi round two

Across all stakeholder groups of participants, at least 67% indicated that 'pain' (97%), 'physical functioning' (94%), 'global assessment of treatment success' (86%) and 'health-related quality of life' (79%) should be included as core domains (figure 2). 'Sleep functioning' (70%) and 'psychological functioning' (69%) also met the criteria for inclusion based on responses from the total panel. However, the threshold for inclusion of these two domains was not met by all stakeholders, as 35% of clinicians/researchers were unsure about or preferred to exclude 'sleep functioning', while 37% of patients were unsure about or preferred to exclude 'psychological functioning' (see online supplementary file 1, figure S6–S8). Few participants (14%) considered it important to include 'number of deaths' as a core domain. Comments emphasised that death could be recorded as a serious adverse event rather than as a separate domain.

Only 44% of the total panel agreed that 'range of motion' and 'muscle strength' should be excluded from the COS. Few participants provided rationale for their responses. Of those that did, beliefs about the lack of added value that these measures provide over self-reported functional ability were most commonly expressed. Further, 54% of the total panel agreed that none of the domains from round one that we excluded belonged in the COS. However, there was variation across stakeholder groups on this question, with 100% of those who worked only as researchers agreeing that they should all be excluded, compared with 33% of patients. A recurring comment was that 'requiring reoperation or revision surgery' and 'failure of surgery' should be retained.

DISCUSSION

From a list of 41 potential core domains, we identified four that the majority of participants (across all stakeholder groups) considered core: 'pain', 'physical functioning', 'global assessment of treatment success' and 'health-related quality of life'. Two additional domains —'sleep functioning' and 'psychological functioning'— also met the threshold for inclusion based on responses from the total panel, but not among all stakeholders. There was consensus that 'number of deaths' should not be a core domain, but insufficient agreement on whether or not several other domains, including 'range of motion' and 'muscle strength', were core domains.

There are several strengths of this research. We generated a comprehensive list of potential core domains based on a prior literature review, as recommended by the OMERACT[8] and COMET[14] initiatives. We obtained participation from patients, and clinicians and researchers representing various scientific and clinical disciplines, which may enhance the generalisability of the

results. Also, we allowed Delphi participants to provide rationale for their domain ratings, which yielded convincing arguments to aggregate certain domains. A limitation is that the response rates to both Delphi rounds were low, particularly for clinicians/researchers. The majority of individuals in the sampling frame were corresponding authors of shoulder trials indexed in PubMed from 2006 to 2015. Not all of these authors may have had a longstanding interest in shoulder disorders, and those who switched their focus to other clinical conditions may have been less likely to respond. Also, we do not know how many participants completed both rounds of the survey, and therefore cannot tell if the characteristics of participants were similar in both rounds. While this can be viewed as a minor drawback, we still maintained close to equal numbers of our stakeholder groups in both rounds. Further, we did not pilot the survey with patients beforehand. Given that the concept of 'outcomes/endpoints' can be obscure,[15] some patients may have misunderstood the purpose of the Delphi survey, or the wording of certain domains and definitions, which could have resulted in unreliable ratings from them. Another limitation is that the sample size for some of the stakeholder groups was small (eg, in round one, only 13 participants worked as a researcher only, and in round two, only 12 participants worked as a clinician only). Therefore, the comparison of ratings by the different stakeholder groups should be interpreted with caution.

The finding that nearly all Delphi participants rated 'pain' and 'physical functioning' as core domains is consistent with what has been measured in published shoulder trials (these were measured in 90% and 71% of trials, respectively.[18] However, the high ratings of 'global assessment of treatment success' and 'health-related quality of life' contrasts with published literature, where both domains were measured in only 15% of trials.[18] The infrequent measurement of these domains may have occurred because of unfamiliarity with appropriate measurement instruments or assumptions that 'pain' and 'physical functioning' are the only PROMs that matter. If 'global assessment of treatment success' and 'health-related quality of life' are included in the final COS for shoulder disorders, education on the value of these domains will need to be prioritised during implementation strategies.

Based on suggestions from round one participants, the definition of 'physical functioning' that we used in round two covered activities of daily living (eg, bathing), recreation/leisure activities and work ability. Despite being considered core by 94% of participants, the revised 'physical functioning' definition is incompatible with some measurement instruments that purport to measure this construct (eg, the disability subscale of the Shoulder Pain And Disability Index[16] only addresses activities of daily living). It is possible that treating 'physical functioning' as a multidimensional domain may not be feasible in practice, and instead the different components may need to be treated as separate subdomains. This will become clearer once we select measurement instruments for the COS.

Implementation of a COS into clinical trials may depend on the brevity of the set and the feasibility of measurement.[14] With this in mind, there are several reasons why it may be appropriate to exclude domains where no consensus was reached; for example, 'range of motion', 'muscle strength' and surgical outcomes. Inclusion of 'range of motion' and 'muscle strength' could lead to increased research costs for shoulder disorder trials, since staff and equipment are needed to perform the assessments. Also, surgery-relevant outcomes should only be measured in surgical trials (not *all* intervention trials, which, by definition, a COS should apply to). We plan to discuss these issues further during face-to-face meetings with various stakeholders before reaching a final decision.

We are aware of only one other study which has sought perspectives on the most important domains to measure in shoulder trials.[17] Researchers asked 225 UK orthopaedic surgeons, general practitioners and physiotherapists to suggest which domain should be the primary outcome in trials for adhesive capsulitis. The most frequently suggested primary outcomes were function (59%), pain (48%) and range of motion (46%), while less than 5% of respondents thought quality of life, period to resolution, patient satisfaction, cost-effectiveness and adverse events should be primary. Having to nominate domains (without seeing a list of potentially important ones) and being able to select only one domain may explain why 'global assessment of treatment success' and 'health-related quality of life' were rarely or never considered primary by clinicians in the study by Rodgers *et al*. To the best of our knowledge, no other study has been as comprehensive in scope as ours, with its aim to identify a core set of domains for all trials across various shoulder disorders, and consideration of researcher and patient perspectives.

CONCLUSION

A majority of patients, clinicians and researchers identified 'pain', 'physical functioning', 'global assessment of treatment success' and 'health-related quality of life' as important domains to include in a COS for shoulder disorders. The value of several other domains, including 'sleep functioning', 'psychological functioning', 'range of motion' and 'muscle strength', needs further consideration. Our findings will be discussed in consensus meetings at OMERACT 2016, where we plan to reach agreement on a preliminary core set of domains for shoulder disorder trials.[18] In future, we plan to seek endorsement of this core set from OMERACT, and to select measurement instruments that must be administered to cover each corresponding domain.

Author affiliations
[1]School of Public Health and Preventive Medicine, Monash University, Melbourne, Victoria, Australia
[2]School of Social and Community Medicine, University of Bristol, Bristol, UK

[3]Department of Orthopaedic Surgery, University of Michigan, Ann Arbor, Michigan, USA
[4]Department of General Practice, Erasmus MC, University Medical Center Rotterdam, Rotterdam, The Netherlands
[5]Monash Department of Clinical Epidemiology, Cabrini Institute and Monash University, Melbourne, Victoria, Australia
[6]Department of Epidemiology and Preventive Medicine, School of Public Health and Preventive Medicine, Monash University, Melbourne, Victoria, Australia
[7]Department of Epidemiology, School of Public Health, University of Michigan, Ann Arbor, Michigan, USA

Contributors RB and JJG conceived the study design and obtained funding. MJP, HH and APV provided input on the study design. MJP and HH undertook the statistical analyses. MJP wrote the first draft of the manuscript. All authors contributed to revisions of the manuscript. All authors approved the final version of the submitted manuscript.

Funding This project was supported by funding from a Patient-Centered Outcomes Research Institute (PCORI) Eugene Washington Engagement Award #2072, and Outcome Measures in Rheumatology (OMERACT). MJP is supported by an Australian National Health and Medical Research Council (NHMRC) Early Career Fellowship (1088535). RB is supported by an Australian NHMRC Senior Principal Research Fellowship.

Competing interests None declared.

Ethics approval University of Michigan (HUM00102228).

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement All relevant data are within the paper and its Supporting Files.

REFERENCES

1. Linsell L, Dawson J, Zondervan K, *et al*. Prevalence and incidence of adults consulting for shoulder conditions in UK primary care; patterns of diagnosis and referral. *Rheumatology (Oxford)* 2006;45:215–21.
2. Ostör AJ, Richards CA, Prevost AT, *et al*. Diagnosis and relation to general health of shoulder disorders presenting to primary care. *Rheumatology (Oxford)* 2005;44:800–5.
3. Green S, Buchbinder R, Glazier R, *et al*. Systematic review of randomised controlled trials of interventions for painful shoulder: selection criteria, outcome assessment, and efficacy. *BMJ* 1998;316:354–60.
4. Tekavec E, Jöud E, Rittner R, *et al*. Population-based consultation patterns in patients with shoulder pain diagnoses. *BMC Musculoskelet Disord* 2012;13:238.
5. van der Heijden GJ. Shoulder disorders: a state-of-the-art review. *Baillieres Best Pract Res Clin Rheumatol* 1999;13:287–309.
6. Largacha M, Parsons IM, Campbell B, *et al*. 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.
7. Page MJ, McKenzie JE, Green SE, *et al*. Core domain and outcome measurement sets for shoulder pain trials are needed: systematic review of physical therapy trials. *J Clin Epidemiol* 2015;68:1270–81.
8. Boers M, Kirwan JR, Tugwell P, *et al*. The OMERACT Handbook. http://www.omeract.org/pdf/OMERACT_Handbook.pdf. 2016.
9. Williamson P, Altman D, Blazeby J, *et al*. Driving up the quality and relevance of research through the use of agreed core outcomes. *J Health Serv Res Policy* 2012;17:1–2.
10. Boers M, Kirwan JR, Wells G, *et al*. Developing core outcome measurement sets for clinical trials: OMERACT filter 2.0. *J Clin Epidemiol* 2014;67:745–53.
11. Whittle S, Buchbinder R. In the clinic. Rotator cuff disease. *Ann Intern Med* 2015;162:ITC1–15.
12. 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;8:e1000393.
13. Chiarotto A, Deyo RA, Terwee CB, *et al*. Core outcome domains for clinical trials in non-specific low back pain. *Eur Spine J* 2015;24:1127–42.
14. Williamson PR, Altman DG, Blazeby JM, *et al*. Developing core outcome sets for clinical trials: issues to consider. *Trials* 2012;13:132.
15. Keeley T, Williamson P, Callery P, *et al*. The use of qualitative methods to inform Delphi surveys in core outcome set development. *Trials* 2016;17:230.
16. Roach KE, Budiman-Mak E, Songsiridej N, *et al*. Development of a shoulder pain and disability index. *Arthritis Care Res* 1991;4:143–9.
17. Rodgers S, Brealey S, Jefferson L, *et al*. Exploring the outcomes in studies of primary frozen shoulder: is there a need for a core outcome set? *Qual Life Res* 2014;23:2495–504.
18. Buchbinder, Page, Huang, *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* 2016. In press.