Eur Spine J (2015) 24:1127–1142
DOI 10.1007/s00586-015-3892-3

ORIGINAL ARTICLE

Core outcome domains for clinical trials in non-specific low back pain

Alessandro Chiarotto[1] · Richard A. Deyo[2] · Caroline B. Terwee[3] · Maarten Boers[3] · Rachelle Buchbinder[4] · Terry P. Corbin[5] · Leonardo O. P. Costa[6,7] · Nadine E. Foster[8] · Margreth Grotle[9] · Bart W. Koes[10] · Francisco M. Kovacs[11] · Chung-Wei Christine Lin[7] · Chris G. Maher[7] · Adam M. Pearson[12] · Wilco C. Peul[13] · Mark L. Schoene[14] · Dennis C. Turk[15] · Maurits W. van Tulder[1] · Raymond W. Ostelo[1,3]

Received: 19 December 2014 / Revised: 17 March 2015 / Accepted: 19 March 2015 / Published online: 5 April 2015

Abstract

Purpose Inconsistent reporting of outcomes in clinical trials of patients with non-specific low back pain (NSLBP) hinders comparison of findings and the reliability of systematic reviews. A core outcome set (COS) can address this issue as it defines a minimum set of outcomes that should be reported in all clinical trials. In 1998, Deyo et al. recommended a standardized set of outcomes for LBP clinical research. The aim of this study was to update these recommendations by determining which outcome domains should be included in a COS for clinical trials in NSLBP.

Methods An International Steering Committee established the methodology to develop this COS. The OMERACT Filter 2.0 framework was used to draw a list of potential core domains that were presented in a Delphi study. Researchers, care providers and patients were invited to participate in three Delphi rounds and were asked to judge which domains were core. A priori criteria for consensus were established before each round and were analysed together with arguments provided by panellists on importance, overlap, aggregation and/or addition of potential core domains. The Steering Committee discussed the final results and made final decisions.

✉ Alessandro Chiarotto
a.chiarotto@vu.nl

[1] Department of Health Sciences, Faculty of Earth and Life Sciences, EMGO+ Institute for Health and Care Research, VU University Amsterdam, de Boelelaan 1085, room U-601, 1081 HV Amsterdam, The Netherlands

[2] Departments of Family Medicine, Internal Medicine, Public Health and Preventive Medicine, Oregon Institute of Occupational Health Sciences, Oregon Health and Science University, Portland, USA

[3] Department of Epidemiology and Biostatistics, EMGO+ Institute for Health and Care Research, VU University Medical Center, Amsterdam, The Netherlands

[4] Monash Department of Clinical Epidemiology, Cabrini Institute and Department of Epidemiology and Preventive Medicine, School of Public Health and Preventive Medicine, Monash University, Melbourne, Australia

[5] Cochrane Collaboration Back Review Group, Maple Grove, USA

[6] Masters and Doctoral Programs in Physical Therapy, Universidade Cidade de Sao Paulo, Sao Paulo, Brazil

[7] The George Institute for Global Health, Sydney Medical School, The University of Sydney, Sydney, Australia

[8] Arthritis Research UK Primary Care Centre, Institute of Primary Care and Health Sciences, Keele University, Keele, UK

[9] Oslo and Akershus University College of Applied Sciences, Faculty of Health Sciences and FORMI, Oslo University Hospital, Oslo, Norway

[10] Department of General Practice, Erasmus MC, University Medical Center, Rotterdam, The Netherlands

[11] Spanish Back Pain Research Network, Fundacion Kovacs, Palma de Mallorca, Spain

[12] Department of Orthopaedic Surgery, Dartmouth-Hitchcock Medical Center, Lebanon, USA

[13] Department of Neurosurgery, Leiden University Medical Center, Leiden, The Netherlands

[14] Cochrane Collaboration Back Review Group, Newbury, USA

[15] Department of Anesthesiology and Pain Medicine and Center for Pain Research on Impact, Measurement and Effectiveness, University of Washington, Seattle, USA

Results A set of 280 experts was invited to participate in the Delphi; response rates in the three rounds were 52, 50 and 45 %. Of 41 potential core domains presented in the first round, 13 had sufficient support to be presented for rating in the third round. Overall consensus was reached for the inclusion of three domains in this COS: 'physical functioning', 'pain intensity' and 'health-related quality of life'. Consensus on 'physical functioning' and 'pain intensity' was consistent across all stakeholders, 'health-related quality of life' was not supported by the patients, and all the other domains were not supported by two or more groups of stakeholders. Weighting all possible argumentations, the Steering Committee decided to include in the COS the three domains that reached overall consensus and the domain 'number of deaths'.

Conclusions The following outcome domains were included in this updated COS: 'physical functioning', 'pain intensity', 'health-related quality of life' and 'number of deaths'. The next step for the development of this COS will be to determine which measurement instruments best measure these domains.

Keywords Core outcome set · Domains · Clinical trials · Non-specific low back pain

Introduction

The Global Burden of Disease study has highlighted that low back pain (LBP) is the leading global contributor to years lived with disability and the sixth global contributor to disability-adjusted life years [1, 2]. The global prevalence of activity-limiting LBP was recently estimated to be approximately 39 % for lifetime prevalence and 18 % for point-prevalence [3]. Only a small proportion of people experiencing LBP seek health care but these account for high costs that represent an important burden to society [4, 5]. The large majority of patients with LBP are labelled as having non-specific LBP (NSLBP) because no underlying pathology or cause can be found [6–8]. A wide range of health interventions exists for patients with NSLBP and related clinical trials are often summarized in systematic reviews [9, 10]. However, authors of these reviews report that outcomes are inconsistently measured and reported across trials [11–13]. This inconsistency may limit the comparison of findings among trials and hinder statistical pooling [14]. In addition, inconsistent reporting can be due to selective reporting bias (e.g. reporting only favourable outcomes in a publication), which may strongly affect the conclusions of systematic reviews [15].

The development and use of core outcome sets (COS) for specific health conditions has been suggested to reduce inconsistency in outcomes measured and reported across clinical trials [14]. A COS represents an agreed set of outcomes that should be measured and reported, as a minimum, in all clinical trials for specific health conditions [16]. Such a set does not restrict measurement or the choice of the primary outcome, but mandates collection and reporting of the COS alongside the outcomes of interest [16]. A COS thus creates a minimum standard of outcomes reported, reducing the risk of selective reporting bias and increasing the validity and statistical power of meta-analyses [17].

The recently launched Core Outcome Measures in Effectiveness Trials (COMET) initiative fosters methodological research and provides methodological guidance on the development of a COS [16]. The expertise accumulated by the Outcome Measures in Rheumatology (OMERACT) initiative is also a fundamental guidance in COS development [18]. A stepwise approach is suggested by both initiatives: first, the core outcome domains should be selected (i.e. 'what' to measure), and then the measurement instruments for each domain (i.e. 'how' to measure) [16, 19].

In the field of LBP, recommendations for standardized reporting of outcome measurement instruments in clinical studies were formulated at an expert panel discussion held at the 1997 International Forum on LBP in Primary Care (The Hague, The Netherlands) [20]. Specific recommendations were made for five outcome domains (i.e. 'pain symptoms', 'back-related function', 'generic well-being', 'disability social role' and 'satisfaction with care') [20, 21]. A workshop discussion among LBP researchers during the 2012 International LBP Forum (Odense, Denmark) agreed on the need of updating the existing recommendations [22]. This was motivated by recent advances in understanding of construct development and measurement properties that stress the need to explore whether relevant domains are missing and to critically appraise recommended instruments [22]. Deyo et al. [20] proposed also a parsimonious set of six questions covering the five domains suggested for measurement in LBP clinical research. These questions were extracted from existing questionnaires and were proposed as the minimum to be used in a wide variety of settings, including routine clinical care [20]. This brief set was labelled as 'Core Outcome Measures Index' (COMI) by other investigators who assessed its measurement properties and feasibility of implementation [23, 24]. However, it is out of the scope of this study to update the set of questions included in the COMI for LBP.

The aim of this study is to update the existing standardized set of outcome domains and measurement instruments recommended for LBP [20, 21], through the development of a COS. This COS is intended for the measurement of efficacy or effectiveness of health interventions assessed in all clinical trials for patients with NSLBP. We defined NSLBP as "low back pain not

attributable to a recognizable, known specific pathology (e.g. infection, tumour, fracture, axial spondyloarthritis)" [25]. The first step in the development of this COS and focus of this manuscript was to perform a Delphi study to reach international consensus on core outcome domains.

Methods

A detailed description of the methods of this Delphi study is presented elsewhere [26]. An International Steering Committee with members from four continents, including researchers, care providers and patients' representatives, worked on the development of this COS. The day-to-day conduction of the study was performed by a project team of four people (AC, CT, MB, RO) working at the same institution (VU University/VU Medical Center, Amsterdam) who designed and addressed key aspects of the study. The other members of the Committee were regularly consulted by e-mail regarding critical decisions.

The Steering Committee decided to involve four groups of stakeholders in the Delphi study: health care researchers, health care providers, professionals working both as researchers and providers, and patients with NSLBP. Professionals from many fields of clinical research relevant for NSLBP (e.g. orthopaedics, physiotherapy, epidemiology, psychology, rheumatology, rehabilitation medicine) were involved. Patients are judged to be essential in developing COSs as they can bring the perspective of those living with a health condition [16, 18]. Previous COS efforts involving patients or the public identified core outcome domains that were not previously identified by other stakeholders [27–29].

The main advantages of a Delphi method include the involvement of informed individuals, anonymity of responses that reduces influence of prominent personalities, and the possibility for Delphi panellists to reconsider their views based on feedback reports of previous rounds [30, 31]. As this project did not involve experiments with patients or study subjects, according to the Dutch Medical Research in Human Subjects Act (WMO), it was exempt from ethical approval. All patients involved were asked for their consent prior to participation and all procedures were conducted according to the Declaration of Helsinki.

Selection of panellists

A list of health care researchers who had extensively published on LBP over the last 10 years (2003–2013) was made by one reviewer (AC) through a structured search in Web of Science (accessed October 7, 2013) and PubMed [26]. Other researchers and health care providers were added to this list through convenience sampling. Patients were recruited through the Steering Committee, seeking people who sought care for a present or past episode of NSLBP and had a fluent understanding of written English. When patients willing to participate were identified, they were contacted by email, given further information on the study and asked for consent to participate. Patients agreeing to participate were sent an information document giving simplified explanations of the terminology used in the study. Members of the Committee were also selected to participate in the Delphi so that they could express their vote on core domains. The final list of potential panellists was managed by the project team and names in the list remained blinded to all those selected for participation.

Generation of a list of potential core domains

The Steering Committee took responsibility for drawing a list of potential core domains that was used in the Delphi study. This list resulted from a search of outcome domains measured in clinical trials included in five recent systematic reviews [12, 13, 32, 33] (one of which not published yet) with addition of the (sub) domains included in the comprehensive International Classification of Functioning (ICF) core set for LBP [34], and in a conceptual model developed to characterize the burden of LBP [35]. This conceptual model and the ICF core set were adopted to account for the patients' perspective in this early phase. The model on the burden of LBP was developed by asking different stakeholders (including patients) which aspects of health were the most relevant to them [35]; the comprehensive ICF core set was shown to cover all health issues identified by patients with LBP [36]. The OMERACT Filter 2.0 framework was used to structure the list of potential core domains, subdividing it into four core areas that encompass the complete content of what is potentially measurable in a clinical trial ("Appendix I") [19]. To determine wording and definitions of the potential core domains, terminology used in existing health frameworks or COSs were consulted: ICF [37], Patient Reported Outcomes Measurement Information System (PROMIS) [38], Wilson and Cleary Model [39] and IMMPACT [40, 41].

Delphi procedure

Three Delphi rounds, including open- and close-ended questions, were used to reach consensus on core outcome domains. Individuals not participating in one round, and who did not explicitly express their desire to opt-out, were invited to each subsequent round. The Delphi study was conducted using SurveyMonkey software and invitations to participate were sent by email.

In the first round, panellists were asked to judge whether each potential core domain was important enough to be

included in this COS with possible answers 'yes', 'no' and 'unsure/not my expertise'. Panellists were given the opportunity to propose changes of wording and definitions of domains, to indicate if some domains had major conceptual overlap or had to be aggregated, and to suggest the inclusion of missing potential core domains. A question was asked about the ideal number of domains for this COS and another about reporting of adverse events (AEs). Panellists were always encouraged to provide a rationale for their answers. A priori cut-off criteria were established for excluding domains that were rejected by more than 60 % and favoured by less than 20 % of respondents.

In the second round, a proposal was made for exclusion of domains that did not have at least 67 % of the first round respondents answering 'yes' or 'unsure/not my expertise'. Other proposals were made for excluding or retaining domains suggested as having large conceptual overlap. Consensus for the second round was a priori set at 67 % of respondents agreeing with a proposal. Panellists were also asked to judge whether the potential core domains suggested as missing were important enough to be included in the COS, as done for the other domains in the first round.

The remaining potential core domains were presented in the third round to ask the panellists if each was indeed core. A priori consensus was set at 67 % of the panel agreeing that a domain is core. In each round, descriptive statistics were used to summarize all the questions. All rationales provided by panellists were checked against the quantitative results to evaluate whether substantial inconsistencies emerged. Responses of the patients' group were always analysed separately to assess whether discrepancies were emerging with the rest of the panel. In the third round, frequencies of responses for each domain were calculated for the whole panel and separately for each of the stakeholder groups.

Final decisions

The project team made some proposals to the Steering Committee regarding the interpretation of the final results of the Delphi. Committee members expressed their opinion on each proposal and the opinion supported by more than 50 % of members was followed. Some proposals concerned the inclusion of a 'death' and a 'pathophysiological manifestations' domain in the COS (as recommended by the OMERACT initiative for all COSs [19]), and what would be an appropriate approach for the reporting of adverse events (AEs).

Results

Panellists

We selected a sample of 280 experts to participate: 139 researchers, 108 care providers, 15 patients, and 18 members of the Steering Committee. A flowchart of the response rate in each round is presented in Fig. 1; 79 of the selected panellists (29 %) participated in all three rounds. People from five continents participated, with the United States, The Netherlands, Australia and the United Kingdom being the most represented countries (Table 1). Socio-demographic characteristics, panellists' disciplines of expertise and experience with NSLBP clinical research were not substantially different between rounds (Table 1). Fourteen patients (seven men and seven women) participated in the first round: three had current and past episodes of NSLBP, six had only a current episode, and three had NSLBP only in the past. Among the nine with current NSLBP: seven sought care for their back problem, three were off-work

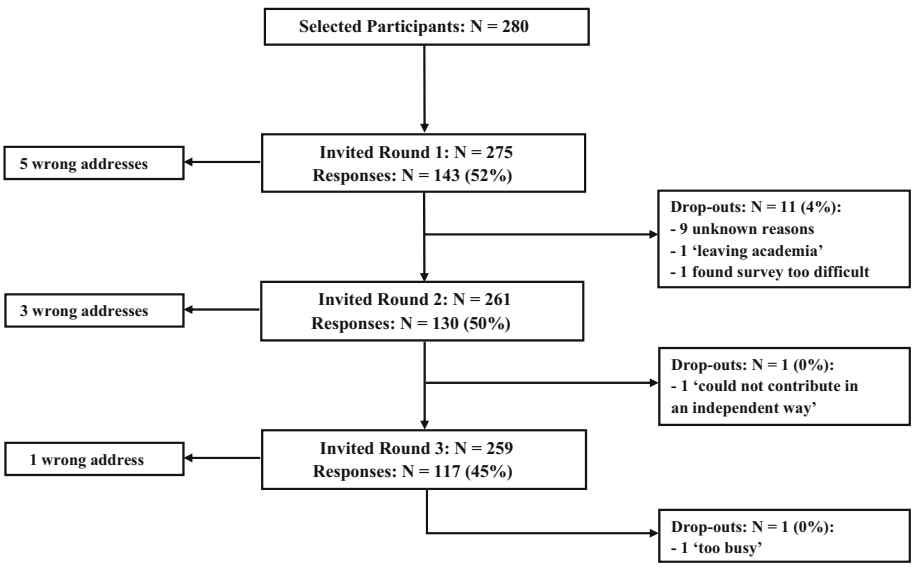

Fig. 1 Flowchart of participation rates per round

Table 1 Characteristics of participants of the Delphi study

Characteristics	Round 1	Round 2	Round 3
Number (*n*)			
Total number of participants	143	130	117
Complete answers on domains	131	127	115
Stakeholder group (%)			
Health care researchers	45	44	47
Health care providers	15	15	19
Health care researchers and providers	30	28	25
Patients	10	8	9
Missing information	0	5	0
Gender (%)			
Female	28	24	26
Male	72	74	74
Missing information	0	2	0
Nationality (%)			
United States of America	22	21	22
The Netherlands	17	19	20
Australia	11	11	11
United Kingdom	10	12	9
Brazil	6	5	5
Italy	6	5	3
Norway	4	4	2
Canada	3	2	5
Spain	3	3	3
Belgium	2	3	2
Germany	2	2	3
France	1	1	1
Finland	1	1	2
Other[a]	12	11	12
Work country (%)			
United States of America	22	21	23
The Netherlands	17	20	20
Australia	13	11	13
United Kingdom	11	11	9
Brazil	6	5	4
Italy	6	5	3
Norway	4	4	2
Canada	4	3	6
Spain	3	3	3
Belgium	3	2	2
Germany	2	2	3
Denmark	1	2	3
France	1	2	1
Finland	1	1	2
Switzerland	1	1	0
Other[a]	5	7	6
Educational background[b] (%)			
Physiotherapy	36	32	34
Epidemiology	28	26	29
Orthopaedics	12	12	14
Rheumatology	8	9	7
Human movement science	8	8	10
Internal medicine	7	8	8
Psychology	6	6	9
Physical medicine and rehabilitation	6	7	14
Anesthesiology	5	4	3
Chiropractic	4	4	5
Osteopathy	4	2	3
Neurosurgery	3	2	2
Other	17	18	21
Missing information	0	5	0
Field of work[b] (%)			
Physiotherapy	32	29	30
Epidemiology	29	28	30
Orthopaedics	22	23	19
Rheumatology	12	13	9
Physical medicine and rehabilitation	6	7	24
Anesthesiology	4	6	5
Psychology	6	5	5
Chiropractic	4	4	1
Human movement science	4	5	3
Internal medicine	2	3	1
Neurosurgery	4	3	3
Osteopathy	4	3	1
Other	47	22	29
Missing information	0	5	0
Clinical trials in NSLBP[c] (%)			
None	19	19	17
1–3 clinical trials	36	37	41
>4 clinical trials	45	37	42
Missing information	0	7	0
Systematic reviews in NSLBP[c] (%)			
None	42	44	31
1–3 systematic reviews	39	33	43
>4 systematic reviews	19	16	26
Missing information	0	7	0
Development of measurement instruments for NSLBP[c] (%)			
None	48	44	48
1 measurement instrument	26	25	18
>2 measurement instruments	26	24	34
Missing information	0	7	0
Testing of measurement instruments for NSLBP[c] (%)			
None	32	28	31
1–3 studies	47	44	46

Table 1 continued

Characteristics	Round 1	Round 2	Round 3
>4 studies	21	21	23
Missing information	0	7	0
COS development[c] (%)			
None	70	63	65
>1 COS	30	30	35
Missing information	0	7	0

[a] Participants with more than one nationality or working in more than one country are included in this category

[b] Percentages are calculated on the whole sample because each participant could indicate more than one field

[c] Participation in the design, analysis and/or conduction of the mentioned type of study. These questions were not asked to patients, percentages are calculated on the sample of potential respondents

due to their LBP, two had acute NSLBP (i.e. pain for less than a month), three chronic NSLBP from three months to a year, four chronic NSLBP for more than a year. None of the patients underwent a surgical operation for current and/or past episodes of LBP. In total, forty-six panellists of the first round (32 %) sought care for a present or past episode of LBP but only those specifically invited as patients were considered part of this stakeholder group.

List of potential core domains

The list of potential core domains generated by the Steering Committee included 41 outcome domains, subdivided as follows: 1 in the core area 'death', 21 in 'life impact', 6 in 'resource use/economical impact' and 13 in 'pathophysiological manifestations'. The list with all definitions used in the Delphi study is presented in Table 2.

Delphi round 1

The first round ran from February 18 to March 24, 2014. The results on inclusion of the 41 domains are presented in Fig. 2. Six domains met a priori criteria for exclusion: 'legal services', 'body structures', 'muscle tone', 'structural stability', 'proprioception' and 'urination'. For 12 of the other domains, at least 67 % of respondents indicated that they should be included in the COS or were unsure about it (Fig. 2). The remaining 23 domains did not reach this threshold and their exclusion was proposed in the second round. No clear discrepancies between the patients' perspective and overall panel responses were identified.

One hundred and thirty-one panellists answered the question on the ideal number of domains and 106 (81 %) indicated a specific number; the suggested median number of domains was 7 (interquartile range 5–10) and the majority of the comments were in favour of a small COS. The majority of respondents to the question on AEs (72 %) agreed that only AEs occurring outside of core outcome domains should be reported as AEs.

Several panellists emphasized the overlap of 'health-related quality of life' with other more specific domains (e.g. 'physical functioning', 'psychological functioning') (see "Appendix II"). To address this, a proposal was formulated for the second round to exclude 'health-related quality of life' from the list. Panellists also remarked that 'work ability' and 'work productivity' should not be included in all trials because they are not applicable to non-working populations, and because they overlap ("Appendix II"). These comments had to be balanced against favourable comments for inclusion and prompted a proposal for the second round to retain these two domains in the list with an adapted definition that includes also non-paid workers (e.g. students, housewives). Several panellists commented about the overlap of 'pain interference' with other domains ("Appendix II") and these comments were addressed in a proposal to retain it in the list despite the overlap. Despite disagreements on inclusion of 'non-health care services' (Fig. 2), substantial arguments were put forward in its favour. Two patients emphasized that these services (e.g. alternative health care) can be very important, others highlighted that what constitutes 'non-health care services' can differ between countries and that they can be relevant cost-drivers ("Appendix II"). Based on these comments, a proposal for the second round was made to incorporate the content of this domain into 'health care services'. In total, 16 new potential core domains were suggested by panellists for inclusion in the list. Appropriate definitions were searched for these domains and they were presented in the second round for rating ("Appendix III").

Delphi round 2

The second round ran from April 27 to May 26, 2014. Consensus was reached for the exclusion of all but one domain (i.e. 'social functioning', 64 % consensus) that did not have at least 67 % support from the first round. No substantial arguments favoured the retention of these domains.

Consensus was not obtained for excluding the domain 'health-related quality of life' (55 % of the panel recommended its exclusion). Some substantial arguments (e.g. "Construct overlap can only be answered empirically. It is just as likely that the entire question set loads on a single factor, or that there are a few higher order factors. Pain, pain interference, physical functioning, QOL, work, sleep, self-rated health have all been showing to share variance in previous studies. […].") explained the lack of consensus. Consensus was obtained (i.e. 85 %) for incorporating

Table 2 Definitions of potential core domains considered for NSLBP clinical trials

Core area	Domain	Definition
Death	Number of deaths	Reporting of number of deaths occurred within a clinical trial
Life impact	Health-related quality of life	Impact on physical, psychological and social domains of health, seen as distinct areas that are influenced by a person's experiences, beliefs, expectations and perceptions
Life impact	Illness perception	Impact on cognitive and emotional representations of the illness that patients develop to respond to a perceived health threat and that will give rise to problem-based and emotion-focused coping procedures
Life impact	Work ability	Impact on a worker ability to meet physical and/or psychological work demands
Life impact	Individual work performance	Impact on work behaviours or actions that are relevant to the goals of the organization
Life impact	Physical functioning	Impact on patient's ability to carry out daily physical activities required to meet basic needs, ranging from self-care to more complex activities that require a combination of skills
Life impact	Psychological functioning	Impact on patient's levels of anxiety, depression, anger, or other types of psychological distress. Anxiety refers to fear, extreme worrying and hyperarousal symptoms. Depression refers to negative mood, loss of self-confidence, loss of motivation and enjoyment. Anger refers to irritability and frustration
Life impact	Cognitive functioning	Impact on patient's levels of attention, memory, concentration and perception. Attention refers to the ability to focus on tasks, memory refers to the ability to recall information, concentration to the ability to sustain attention, and perception to the ability to interpret information
Life impact	Social functioning	Impact on patient's ability to interact with people in a contextually and socially appropriate manner (e.g. showing consideration and esteem when appropriate, responding to the feeling of others), to create and maintain close relationships with others (excluding members of the family), to engage in desired aspects of community social life (e.g. charitable organizations, service clubs or professional social organizations)
Life impact	Sexual functioning	Impact on patient's ability to conduct physical and mental functions related to intimacy and sexual acts
Life impact	Recreation and leisure activity	Impact on patient's ability to engage in any form of play, recreational or leisure activity
Life impact	Satisfaction with social roles and activities	Impact on patient's satisfaction in performing usual social roles and activities (including family and work)
Life impact	Satisfaction with treatment services	Impact on patient's satisfaction with care received, including treatment and care providers
Life impact	Sleep functioning	Impact on sleep functions like onset, maintenance, quality, amount of sleep, and functions involving the sleep cycle. This domain should also include the impact on perceptions of alertness and sleepiness during usual waking hours
Life impact	Fatigue	Impact on fatigue, ranging from mild subjective feelings of tiredness to an overwhelming, debilitating, and sustained sense of exhaustion that it is likely to decrease one's ability to carry out daily activities (including work activities) and to function at one's usual levels in family or social roles
Life impact	Pain intensity	Impact on how much a patient hurts, reflecting the overall magnitude of the pain experience
Life impact	Pain quality	Impact on sensory and affective qualities of the pain a patient experiences
Life impact	Temporal aspects of pain	Impact on variability of pain over time, namely frequency and duration of pain episodes
Life impact	Pain behaviour	Impact on external manifestations of experiencing pain, such as verbal or nonverbal, involuntary or deliberate actions and reactions
Life impact	Self-rated health	Impact on the subjective rating of patients regarding their general health perceptions, including all the existing health concepts
Life impact	Pain interference	Impact on consequences of pain on relevant aspects of a patient's life and may include the impact of pain on social, cognitive, emotional, physical and recreational activities
Life impact	Independence	Impact on ability to get things the patient wants to do, without the help of others
Resource use/ economical impact	Work productivity	Economical impact on paid or unpaid job employment due to low back pain, including absenteeism and presenteeism

Table 2 continued

Core area	Domain	Definition
Resource use/ economical impact	Health care services	Utilization of health care services within the formal health care system for treating low back pain, including: visits for primary and secondary care, laboratory tests, days of admission to a hospital, medications
Resource use/ economical impact	Informal care	Utilization of unpaid care given to patients with low back pain by members of the family, friends, neighbours, etc
Resource use/ economical impact	Non-health care services	Utilization of health care services not included in the formal health care system for treating low back pain, including: visits to professionals of alternative medicine, "over-the-counter" medications, patients' time and travel expenses
Resource use/ economical impact	Societal services	Utilization of public services, systems or policies aimed at providing support to people who require assistance that is funded by general tax revenues or contributory schemes
Resource use/ economical impact	Legal services	Utilization of services, systems and policies concerning the legislation and other law of a country
Pathophysiological manifestations	Pain biomarkers	Indicators aimed at providing insight into peripheral and central neurobiological mechanisms of pain
Pathophysiological manifestations	Body structures	Bones, joints, muscles, tendons, nerves and other body structures localized on the lumbar spine and/or on other adjacent body parts (i.e. thoracic spine, pelvis, rib cage or lower limbs)
Pathophysiological manifestations	Muscle strength	Force generated by the contraction of a muscle or of a group of muscles of the lumbar spine and/or of other adjacent body parts (i.e. thoracic spine, pelvis, rib cage or lower limbs)
Pathophysiological manifestations	Range of motion	Quantity of movement of the lumbar spine and/or of other adjacent body parts (i.e. thoracic spine, pelvis, rib cage or lower limbs)
Pathophysiological manifestations	Muscle endurance	Capability of sustaining contractions for a required period of time of a muscle or of a group of muscles of the lumbar spine and/or of other adjacent body parts (i.e. thoracic spine, pelvis, rib cage or lower limbs)
Pathophysiological manifestations	Muscle tone	Tension present in the resting muscles (i.e. resistance offered when trying to move them passively) of the lumbar spine and/or of other adjacent body parts (i.e. thoracic spine, pelvis, rib cage or lower limbs)
Pathophysiological manifestations	Structural stability	Maintenance of structural integrity of joint structures of the lumbar spine and/or of other adjacent body parts
Pathophysiological manifestations	Proprioception	Sensory capability of sensing position and movement of the lumbar spine and/or of other adjacent body parts
Pathophysiological manifestations	Spinal control	Capability of performing all aspects related to the control of movement (i.e. motivation to move, sensory inputs, integration of inputs and planning of outputs, motor output to the muscles and mechanical properties of the tissues of the lumbar spine) of the lumbar spine and/or other adjacent body parts
Pathophysiological manifestations	Physical endurance	Respiratory and cardiovascular capacity for enduring physical exercise
Pathophysiological manifestations	Urination	Capability of discharging the urinary bladder
Pathophysiological manifestations	Gait	Movement patterns associated with walking, running or other whole body movements
Pathophysiological manifestations	Neurological signs	Impairments of nerves, spinal cord or brain functions that affect a specific region of the body

"non-health care services" into "health care services", for retaining 'work ability' and 'work productivity' as independent domains (72 %), and for retaining 'pain interference' in the list (68 %).

However, relevant arguments were made against including 'health care services' and 'work productivity' in the list of potential core domains. These arguments outlined that, given the scope of the COS, it might not be appropriate to include these domains in efficacy trials (e.g. "[…] Often in trials patients are requested not to undertake/receive any other treatments during the intervention period, which means differences in use depend on things other than the patient's health state […]"). Several panellists also questioned whether there are valid and reliable methods to assess these domains in all clinical trials (e.g. "[…] during follow-up the acquisition of accurate and

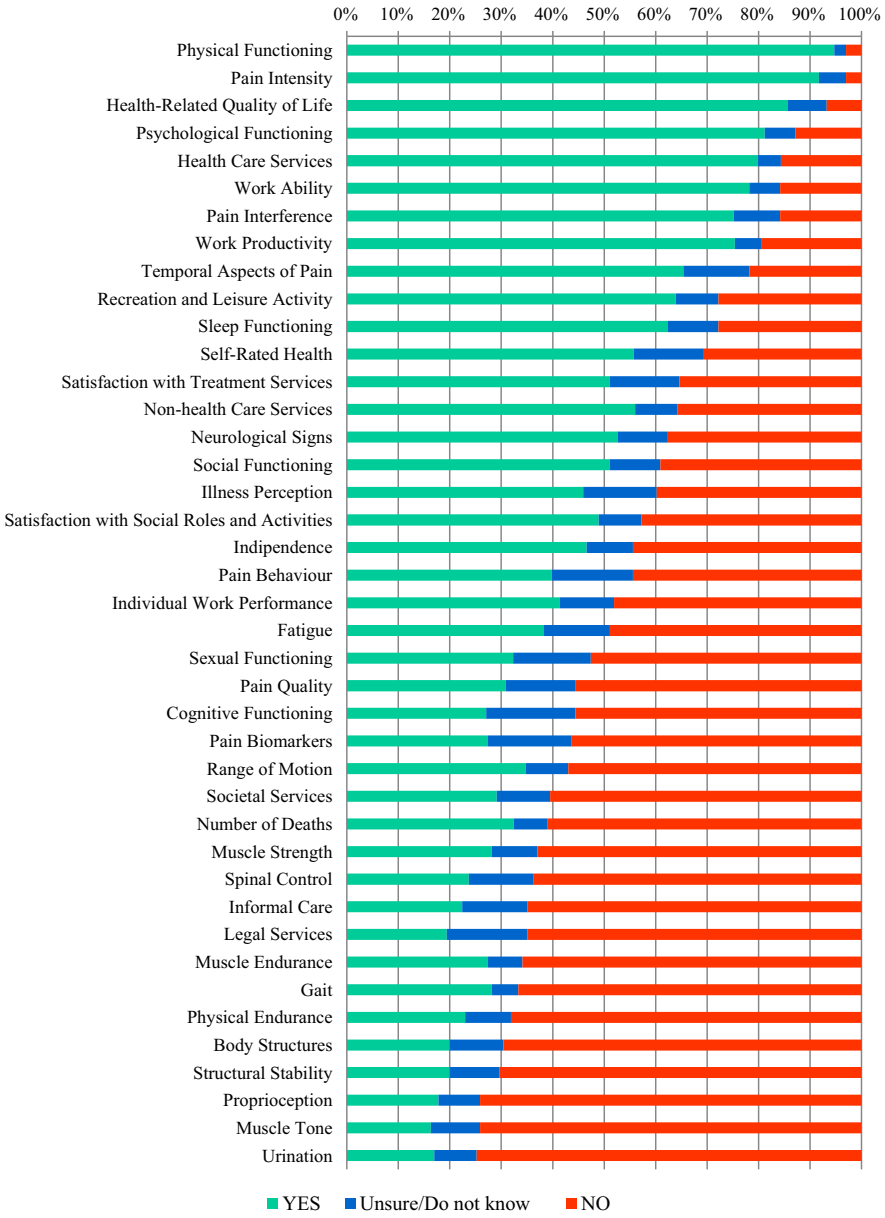

Fig. 2 Ratings of 41 potential core domains in the first Delphi round

Domains are ordered on decreasing % of 'Yes' and 'Unsure/Do not know' summed together

reliable health care services data is questionable", or "[…] both are difficult to assess, may be influenced by factors other than the presence of LBP, and I am not sure of the reliability of the assessment methods"). These domains were kept in the list but these arguments were highlighted in the third round.

None of the new potential core domains suggested in the first round reached consensus for inclusion. Votes for inclusion ranged from 60 % for 'satisfaction with the outcome of treatment' to 13 % for 'travel and transportation'. No substantial differences between patients' responses and the rest of the panel emerged. A total of 13 domains were retained in the list of potential core domains and presented in the last round.

Delphi round 3

The third round ran from June 23 to July 17, 2014. Three domains exceeded the a priori threshold for inclusion in the COS: 'physical functioning' (96 % of respondents indicating it as core), 'pain intensity' (90 %) and 'health-related quality of life' (73 %) (Fig. 3). These ratings were consistent across stakeholder subgroups with the only exception that the patients' group that did not reach agreement (55 %) on 'health-related quality of life' (Fig. 3). 'Work ability' was rated as a core domain by 76 % of health care providers but only by 64 % of the whole panel and 36 % of the patients (Fig. 3). 'Psychological functioning' was considered a core domain by 76 % of care

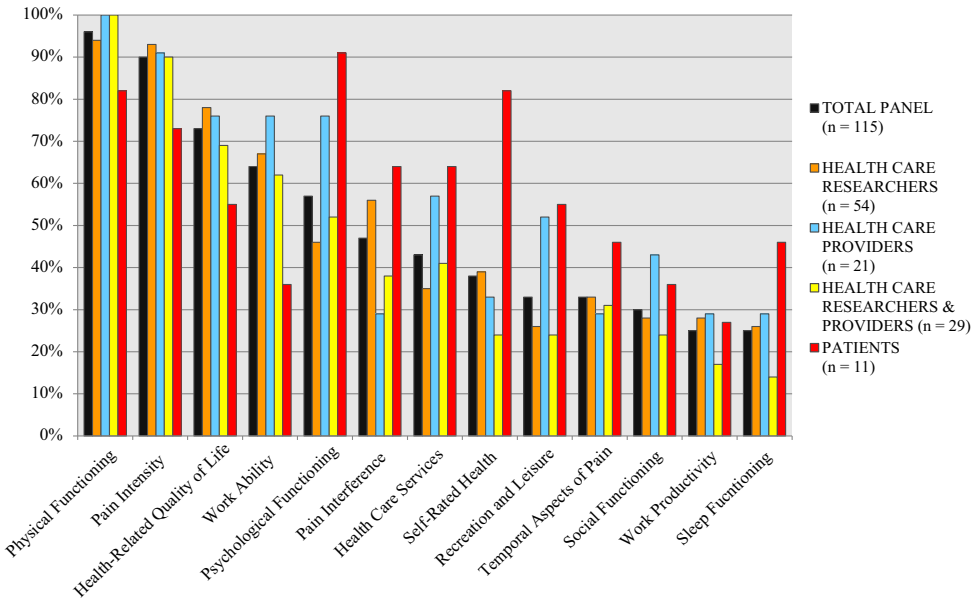

Fig. 3 Ratings of 13 potential core domains in the third Delphi round

The percentage of people rating a domain as core is presented

providers and 91 % of patients but not by the whole panel (Fig. 3). While providers and patients provided ten comments in favour of its inclusion, half of these supported its inclusion as a confounder or moderator, being these not appropriate arguments to support inclusion as an outcome domain. The other eight potential core domains did not reach consensus for inclusion in the COS for any stakeholder group, except 82 % of the patients that rated 'self-rated health' as a core domain (Fig. 3).

Final decisions

Based on the Delphi results, the majority of the Steering Committee members agreed on including 'physical functioning', 'pain intensity' and 'health-related quality of life' in this COS. The Steering Committee considered the inclusion of 'health-related quality of life' because there were strong arguments in its favour: overall consensus was reached, three groups of stakeholders were in favour, and its definition (Table 2) incorporated the excluded domains 'psychological functioning' and 'self-rated health' that were rated as core by some groups of stakeholders (Fig. 3). The Steering Committee also agreed on the exclusion of 'work ability' as overall agreement for inclusion was not reached, as several arguments for inclusion were weak and as it was not considered core by three groups of stakeholders (Fig. 3). These decisions were also taken with the intention of keeping this COS as short as possible to facilitate its implementation.

The majority of Steering Committee members agreed on including the domain 'number of deaths' in the COS as this emphasizes the need to report on the occurrence of deaths in every clinical trial. The Steering Committee acknowledges that death is a rare event for NSLBP clinical trials but a short statement, such as "no deaths occurred in this clinical trial", would suffice to cover this outcome domain. The Steering Committee did not agree with the inclusion of a generic pathophysiological manifestation domain in this COS, as recommended by OMERACT [19]. The main rationale for this decision was that not all interventions for NSLBP are targeting a pathophysiological manifestation, as this disorder is characterized by the absence of a known pathophysiology [6–8, 25]. Furthermore, its inclusion could create unnecessary increases in research costs and impact upon the brevity of the COS. This recommendation does not imply that measuring pathophysiological manifestations is unimportant in relevant NSLBP clinical trials and researchers are encouraged to include them when appropriate for their individual studies.

In the first round of the Delphi, consensus was reached on the reporting of AEs only for those domains not already included in the COS. This approach ensures that, where appropriate, AEs that occur within a core outcome domain (e.g. an increase in 'pain intensity' or a decrease in 'health related quality of life') are included in the statistical analysis. However, taking into account some comments by Delphi panellists, the Steering Committee decided to adopt a flexible approach to the reporting of AEs. This approach leaves the option open to trialists to also report, as separate AEs, those negative outcomes occurring within core domains.

Discussion

Using the methodological guidance of initiatives like COMET and OMERACT [16, 19], we performed a Delphi study to provide an international, multidisciplinary and

multistakeholder consensus-based update of an earlier standardized set of outcome domains for LBP research [20, 21]. Sufficient agreement was reached on core outcome domains that are part of a COS intended for clinical trials assessing efficacy or effectiveness of health interventions in patients with NSLBP. The domains included in this COS are 'physical functioning', 'pain intensity', 'health-related quality of life' and 'number of deaths' (see definitions in Table 2).

The domain 'physical functioning' reached the highest level of consensus in this study and the definition focuses on ability to engage in daily physical activities (Table 2). Our definition of 'physical functioning' will be fundamental to determine which measurement instrument would best measure this domain. IMMPACT recommendations for chronic pain clinical trials also suggest measuring 'physical functioning' as a core outcome domain [40, 41], and this convergence strengthens its inclusion.

'Pain intensity' also reached a very high level of consensus for inclusion in this COS. The inclusion of a pain domain is in line with the original core set [20, 21] and IMMPACT recommendations [40, 41]. 'Pain intensity' for this COS refers to the magnitude of the pain experience, whereas other pain (sub)domains were suggested for consideration by the previous core set and/or IMMPACT (e.g. 'bothersomeness of pain', 'pain quality', 'temporal aspects of domains', 'pain medications') [20, 21, 40, 41]. Some of those pain domains and others (i.e. 'Pain behaviour', 'pain interference') were presented as potential core domains in this Delphi but not sufficient agreement was reached to consider them as core (Figs. 2, 3).

'Health-related quality of life' included in this COS could be considered as the 'successor' of 'general well-being' included in the previous set [20, 21]. However, a definition of 'general well-being' was not given for the previous set and this makes a clear comparison of the two constructs challenging. Taking into account the widely accepted bio-psycho-social model for LBP [42, 43], it may be appropriate to have a domain like 'health-related quality of life' in this COS as its definition includes all components of the model (Table 2). The inclusion of all components of the bio-psycho-social model is also in line with the domains included in a conceptual framework developed to characterize the burden of LBP [35] and with the results of a review that attempted to summarize qualitative research conducted on the impact of LBP on people's lives [44]. However, it will be clear only when choosing measurement instruments for this COS if the different components of 'health-related quality of life' can be treated as separate domains or as one multidimensional domain. The choice of instruments will also be guided by the intention of minimizing redundancy of measurement, to avoid large overlap of instruments and promote brevity of the COS.

Another key aspect in the development of a COS is the definition of contextual factors (i.e. potential confounders and effect modifiers) that should be measured alongside core outcome domains [19]. However, it was beyond the scope of this study to address contextual factors and for the measurement of these factors a reference is made to the prominent work of the National Institutes of Health (NIH) Task Force [45]. This Task Force recently published a report on minimum baseline standards that should be collected in clinical studies for chronic LBP, to standardize their assessment [45].

This COS includes refined versions of three domains included in the previous standardized set but does not incorporate the other two: 'disability social role' and 'satisfaction with care' [20, 21]. 'Disability social role' referred to work absenteeism and could be replaced by the domain 'work productivity' used in this study, while 'satisfaction with care' was formulated as 'satisfaction with treatment services' in this study, but neither was supported by the Delphi panel (Figs. 2, 3). 'Work productivity' refers to indirect non-medical costs that are the first cost drivers for LBP [5] and it is an undoubtedly important outcome for clinical trials with economic evaluations alongside. However, this domain poses the challenge of its measurement in clinical trials aimed at assessing efficacy of interventions, in which an economic evaluation might be out of the scope of the trial. To support the exclusion of 'satisfaction with treatment services' several panellists underlined that it could be highly influenced by factors unrelated to an intervention (e.g. waiting list, amiability of providers, unfriendly receptionist, parking difficulty) and, consequently, that it could say relatively little about efficacy or effectiveness of that intervention.

This is the first Delphi study conducted to explore international, multistakeholder, and multidisciplinary consensus on core outcome domains to be reported in NSLBP clinical trials. This study highlighted diverging opinions on the importance of some domains and reinforced the wisdom of a comprehensive exercise to determine which outcome domains are felt by the majority to be core. The strengths of this study include methods that followed guidance of initiatives like COMET and OMERACT [16, 19], having a large expert panel of varied stakeholders representing various disciplines and countries, giving the opportunity to Delphi panellists to provide comments for each choice, allowing panellists to reconsider their views after considering other panellists' reasoning, attempting to address strong arguments emerging from the Delphi panel, and rigorous reporting of methods [26] and results. One limitation of this study could be the relatively small number of patients involved in the Delphi rounds, which could have led to under or overestimation of the importance of certain domains from their perspective. However,

the goal of this study was not to develop a comprehensive range of outcome domains important to all stakeholders, but rather a core set for inclusion in all clinical trials. Patients can also be involved in trial management teams where they can shape the range of outcomes measures collected in individual trials and this should represent good practice. Finally, the definition of COSs places emphasis on the concept of a minimum set [16, 19] and the four domains included in this COS seem to fit perfectly within this definition. The existence of a small COS for NSLBP should facilitate its inclusion in clinical trials, alongside trial-specific outcomes.

The development of a COS is a stepwise approach [16, 19] and this study determined core outcome domains for clinical trials on NSLBP. The next step will be to reach consensus on which measurement instruments should be used to measure these outcome domains. The selection of instruments will be focused on those that have demonstrated adequate measurement properties for these domains with the least participant burden. Recently published methodological guidance on this topic [46, 47] will help to conduct the next step for this COS in NSLBP.

Conclusions

A consensus-based COS for NSLBP was developed and included the domains 'physical functioning', 'pain intensity', 'health-related quality of life' and 'number of deaths'. This COS represents the update of the standardized set proposed by Deyo et al. in 1998 [20, 21]. The brevity of this COS should facilitate its implementation in clinical trials assessing efficacy or effectiveness of health interventions for NSLBP. Future research should establish which measurement instruments are the most appropriate to measure these core outcome domains.

Acknowledgments We would like to acknowledge all the people who participated in at least one round of the Delphi study. These people (excluding members of the Steering Committee) are listed here in alphabetical order: William A. Abdu, Luc Ailliet, Marcelo Anderson Bracht, Gunnar Andersson, Adri T. Apeldoorn, Majid Artus, Julie Ashworth, Steven J. Atlas, Roxy Azoory, Marco Barbero, Heinz Dieter Basler, David Baxter, Ramsin Benyamin, Mark D. Bishop, Paul Bishop, David Borenstein, Lex Bouter, Hilary Bradbury, Alan Breen, Jens Ivar Brox, Elaine Buchanan, Alex Burdorf, Eugene J. Carragee, John David Cassidy, Roger Chou, Aldo Ciuro, Kris Clark, Steven P. Cohen, Pierre Côté, Peter Croft, Vinicius Cunha Oliveira, Wim Dankaerts, Gavin Davis, Ric Day, Rob de Bie, Henrica C. W. de Vet, Clermont E. Dionne, Wendy T. Enthoven, Hege R. Eriksen, Felipe Fagundes, Carmen Fernandez, Silvano Ferrari, Manuela Ferreira, Paulo H. Ferreira, Timothy W. Flynn, Victoria Franzinetti, Robert Froud, Andrea Furlan, Diego Galace, Robert J. Gatchel, Steven George, Sergio Gimenez Basalotte, Hedley Griffiths, Lars Grovle, Andrew John Haig, Murray Hames, Mark Hancock, Ian Harris, Jan Hartvigsen, Anne Julsrud Haugen, Elaine Hay, Rowland G. Hazard, Standiford Helm, Rob Herbert, Jan Hildebrandt, Deirdre A. Hurley, Eric L. Hurwitz, Julia Hush, Frank Huygen, Wilco C. Jacobs, Matthew Jennings, Johan Juch, Steven J. Kamper, Jaro Karppinen, Peter Kent, Suraj Kumar, Charlotte Leboeuf-Yde, Myeong So Lee, Martyn Lewis, Patrick Loisel, Pim A. J. Luijsterburg, Jon D. Lurie, Luciana Macedo, Luciana Machado, Laxmaiah Manchikanti, Anne F. Mannion, Lynn March, Norman Marcus, Teresa Marin, James McAuley, Alison McGregor, Luciola Menezes Costa, Jan Mens, Stephan Milosavljevic, Shail K. Mirza, Marco Monticone, Lorimer Moseley, Paulo Nascimento, Stefano Negrini, Colin Nelson, Jo Nijs, Oystein Nygaard, John O'Dowd, Teddy Oosterhuis, Richard Osborne, Peter O'Sullivan, Adriano Pezolato, Michael Pfingsten, Serge Poiraudeau, Jan Pool, Pina Porzio, Kristen Radcliff, James Rainville, Francois Rannou, Lisa Roberts, Michael E. Robinson, Myron Rogers, Martin Roland, Ana Royuela, Tamer Sabet, Petry Saeys, Marcus Schiltenwolf, Gay Schoene, Jesus Seco Calvo, Ruth Sephton, William S. Shaw, Karen J. Sherman, Rob Smeets, Anne J. Smith, Matthew Smuck, Bart Staal, Kjersti Storheim, Liv Inger Strand, Michael Sullivan, Simo Taimela, Kazuhisa Takahashi, Judith A. Turner, Martin Underwood, Alexander Vaccaro, Allard van der Beek, Bob van der Meiracker, Danielle van der Windt, Hans van Helvoirt, Willem van Mechelen, Rodrigo Vasconcelos, Arianne Verhagen, Johan W. S. Vlaeyen, Michael von Korff, Debra K. Weiner, Harriet Wittink, Ian Wright, Gustavo Zanoli.

Conflict of interest None of the authors has any potential conflict of interest.

Appendix I

See Table 3.

Table 3 OMERACT Filter 2.0 framework specifying all aspects of a health condition that should be considered in clinical trials (adapted from Boers et al. [19])

Core area	Specification
Death[a]	This core area includes possible specifications of death, as generic or disease-specific, i.e. all cause vs disease-specific morality; and intervention-specific (e.g. death due to surgery)
Life impact[a]	This core area can include domains of the ICF (3) (e.g. activity and participation) and domains within the concept of health-related quality of life (14) (e.g. functional status, general health perceptions, overall quality of life)
Resource use/economical impact[a]	This core area describes the economical impact of health conditions both on society and on the individual. In fact, the presence of a health condition and its treatment incur resource use
Pathophysiological manifestations[b]	This core area is to assess whether or not the effect of the intervention specifically targets the pathophysiology of the health condition. Pathophysiology can include psychosocial manifestations. Example domains are: ICF body function, reversible manifestations (including modifiable risk factors and actual manifestations of ill health), and irreversible manifestations (including unmodifiable risk factors and damage). This area can also encompass all biomarkers and surrogate outcomes

[a] These core areas belong to the concept 'impact of health conditions' that includes all aspects of health or a health condition that are important to the patient and society

[b] This core area belong to the concept 'pathophysiological manifestations of health conditions'

Appendix II

See Table 4.

Table 4 Relevant comments regarding some domains with good level of consensus in the first round

Domain	Comments inconsistent with the overall group response
Health-related quality of life	*"Overlaps with other domains listed in this core area"*
	"Too broad domain"
	"Too generic"
	"Conceptual and measurement overlap with other core variables above and list below"
	"Too broad"
	"This is the overarching domain with other encompasses many of the previous domains discussed"
	"Strongly agree that HR-QoL is a critical domain to include in the COS. It encompasses relevant aspects of psychological, physical and social functioning that patient's identify as key domains for their recovery from LBP (1). If this is included, then there is no need to include items 22, 27 and 33 as they become redundant"
	"Would be a way to capture important aspects of many of the preceding domains"
	"This may be duplicating other measures of mood, social, pain, and function. There is also a cost associated with these with licensees"
	"This is an important one, but seems it would fit within other domains"
	"For me, this would depend on the exact measure, and how much it overlaps with other domains"
	"This is vague, or maybe too broad"
	"Health-related QOL has substantial overlap, as stated in its definition, with the components listed elsewhere of physical, psychological and social functioning"
	"Quality of life is unnecessary given all the other measures"
	"In my mind, HRQoL overlaps with all the other domains"
	"Many of the items overlap by example HRQL and self-rated"
	"HR QOL picks up a lot of aspects of these domains–for trials and getting them done it might be preferred for efficiency reasons"

Table 4 continued

Domain	Comments inconsistent with the overall group response
Work ability and work productivity	"Not all subjects are working"
	"Work is not relevant to ALL patients with LBP (some are retired, some are sick listed / on long-term sick and disability benefits, some are students etc). Whilst work ability is a key outcome for the working population with LBP, it seems unreasonable to include it in a COS. Interference with life and physical functioning should 'cover' this domain, as for those who are working, their work takes up a lot of their time"
	"Too much overlap with work productivity (this could replace it though)"
	"Work ability seems to overlap with absenteeism and presenteeism"
	"Work ability could be combined with presenteeism in the prior list of factors"
	"Work ability is important for most people, but should perhaps be extended to include other activities for people who do not have a job"
	"Actual work is the bottom line success in this area, however economic factors may vary, causing undue variation in study outcomes from time to time and location to location. Also the person's interest in work may vary. They may have children, retire, or the episode may help them realize that they need to go to school or do something else in life. So measuring the ability is definitely second in importance to actual work success. But it can be useful"
	"Yes but unsure if work ability should be within productivity and health QOL. Needs very strict description of what is ability and outcome is more related to whether do or do not participate rather than ability to participate in work/ activity"
	"Work ability could be relevant from the patient's perspective–but may not be applicable to everyone (eg those that do not work etc)"
	"Work productivity should be given very high priority. The problem with this domain comes with handling people who are not in the labor force (students, retired persons, homemakers), it would be helpful if there were clear guidance on how work outcomes are handled when not all of the sample is in the labor force. This contingency can create difficulties in reporting trial results"
	"Work productivity is very important but should also encompass not just work as many unemployed and more and more retired as population ages. Why not just productivity related to all paid or unpaid activity and include ADLs"
Pain interference	"Work ability should be covered by pain interference"
	"Psychological functioning covers the important impact of pain on patient life, it overlaps with pain interference"
	"Recreation and leisure would be captured under the HRQOL and pain interference"
	"Pain interference is covered by physical functioning"
	"Would fold pain interference into functional restoration and/or psychosocial domains"
	"Pain interference is too broad, captures a number of the other domains"
	"Pain interference covers a broad range of constructs (disability, depression, social participation, to name a few) and so interpretation may not be straightforward"
	"I think that pain interference overlaps with physical functioning and HRQoL domains"
	"In a way this can overlap with physical functioning, i.e. physical functioning might be impaired partly because of 'pain interference'"
	"This would be a good one to combine others into–social functioning, recreational functioning etc"
	"Can pain inference be separated from other variables?"
	"Obviously this is very important, not sure that it is core, depends on the other domains chosen and their interrelationship"
	"Physical functioning and pain interference overlap"
Non-health care services	*"Absolutely. Much of low back care occurs outside of the formal healthcare system. Patients, for instance, often engage in both mainstream and alt/comp medicine. And since most back is coping issue rather than a 'treat and cure' disease, it is important to understand the full array of services patients utilize. Think this question should be reworded. The types of services mentioned are all healthcare services so shouldn't be termed ""non-healthcare services"*
	"Non-health care services are commonly used by patients with LBP. They are often quite expensive"
	"Non-health care services are used by the vast majority of back pain patients"
	"This category will vary substantially between cultures/regions, but informal services can represent a significant fraction of HC services sought by patients with LBP. Could be a very brief question/assessment (# visits)"
	"Alternative therapy is fundamental for a better life style, especially when the patient requires non official cures, recognized by protocols, for example patients with allergies to medicines"

In italics are presented those comments that were made by patients

Appendix III

Missing potential core domains suggested by panellists in the first round

1. Decision quality = shared decision making that focuses on two areas: (1) the extent to which patients are informed about treatment options, and (2) the extent to which treatments match what is most important to patients.
2. Satisfaction with the outcome of treatment = extent to which the patient is satisfied with the results of treatment.
3. Abuse/misuse of drugs = misuse or abuse of prescribed or not prescribed drugs.
4. Pain self-efficacy = degree of confidence a patient has in performing regular activities despite the presence of pain.
5. Pain catastrophizing = tendency to misinterpret and/or exaggerate actual or anticipated pain experiences.
6. Bodily extent of pain = number of painful body areas besides the pain experienced in the back.
7. Size of painful area = dimension of the body area in which the patient experiences pain.
8. Compliance with treatment = degree to which a patient follows treatment in all its components.
9. Coping = purposeful use of cognitive and behavioral techniques to manage demands that are perceived as stressful or taxing the resources of the individual.
10. Patient nominated goals = subjective judgment about the achievement of pre-set individual goals.
11. Participation = involvement in a life situation.
12. Recovery = an individual's determination of his/her recovery which involves cognitive appraisal of the impact of symptoms on his/her life, the capacity to perform relevant daily activities, and achievement of an acceptable quality of life through readjustment and other adaptive strategies.
13. Acceptance of pain = willingness to have uncomfortable experiences when the actions that bring about those experiences serve important purposes for the individual.
14. Kinesiophobia = excessive, irrational, and debilitating fear of physical movement and activity resulting from a feeling of vulnerability to painful injury or reinjury.
15. Patient impression of change = patient's assessment of the change related to low back pain since beginning of treatment.
16. Travel and transportation = patient's difficulty with travel and transportation.

References

1. Murray CJ, Vos T, Lozano R, Naghavi M, Flaxman AD, Michaud C, Ezzati M, Shibuya K, Salomon JA, Abdalla S et al (2013) Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet 380(9859):2197–2223
2. Vos T, Flaxman AD, Naghavi M, Lozano R, Michaud C, Ezzati M, Shibuya K, Salomon JA, Abdalla S, Aboyans V et al (2013) Years lived with disability (YLDs) for 1160 sequelae of 289 diseases and injuries 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet 380(9859):2163–2196
3. Hoy D, Bain C, Williams G, March L, Brooks P, Blyth F, Woolf A, Vos T, Buchbinder R (2012) A systematic review of the global prevalence of low back pain. Arthritis Rheum 64(6):2028–2037
4. Dagenais S, Caro J, Haldeman S (2008) A systematic review of low back pain cost of illness studies in the United States and internationally. Spine J 8(1):8–20
5. Lambeek LC, van Tulder MW, Swinkels IC, Koppes LL, Anema JR, van Mechelen W (2011) The trend in total cost of back pain in The Netherlands in the period 2002–2007. Spine 36(13):1050–1058
6. Cohen SP, Argoff CE, Carragee EJ (2008) Management of low back pain. BMJ 337:a2718
7. Deyo RA, Weinstein JN (2001) Low back pain. N Engl J Med 344(5):363–370
8. Koes B, Van Tulder M, Thomas S (2006) Diagnosis and treatment of low back pain. BMJ 332(7555):1430
9. Artus M, van der Windt DA, Jordan KP, Hay EM (2010) Low back pain symptoms show a similar pattern of improvement following a wide range of primary care treatments: a systematic review of randomized clinical trials. Rheumatology 49(12):2346–2356
10. Furlan AD, Pennick V, Bombardier C, van Tulder M (2009) 2009 updated method guidelines for systematic reviews in the Cochrane Back Review Group. Spine 34(18):1929–1941
11. Hayden J, Van Tulder MW, Malmivaara A, Koes BW (2005) Exercise therapy for treatment of non-specific low back pain. Cochrane Database Syst Rev 3:CD000335
12. Kamper SJ, Apeldoorn AT, Chiarotto A, Smeets RJ, Ostelo RW, Guzman J, van Tulder MW (2014) Multidisciplinary biopsychosocial rehabilitation for chronic low back pain. Cochrane Database Syst Rev 9:CD000963
13. Rubinstein SM, Terwee CB, Assendelft W, de Boer MR, van Tulder MW (2012) Spinal manipulative therapy for acute low-back pain. Cochrane Database Syst Rev 9:CD008880
14. Clarke M (2007) Standardising outcomes for clinical trials and systematic reviews. Trials 8(1):39
15. Kirkham JJ, Dwan KM, Altman DG, Gamble C, Dodd S, Smyth R, Williamson PR (2010) The impact of outcome reporting bias in randomised controlled trials on a cohort of systematic reviews. BMJ 340:c365
16. Williamson PR, Altman DG, Blazeby JM, Clarke M, Devane D, Gargon E, Tugwell P (2012) Developing core outcome sets for clinical trials: issues to consider. Trials 13(1):132
17. Kirkham JJ, Gargon E, Clarke M, Williamson PR (2013) Can a core outcome set improve the quality of systematic reviews?–a survey of the Co-ordinating Editors of Cochrane review groups. Trials 14(1):21
18. Tugwell P, Boers M, Brooks P, Simon L, Strand V, Idzerda L (2007) OMERACT: an international initiative to improve outcome measurement in rheumatology. Trials 8(1):38
19. Boers M, Kirwan JR, Wells G, Beaton D, Gossec L, d'Agostino M-A, Conaghan PG, Bingham CO III, Brooks P, Landewé R,

March L, Simon LS, Singh JA, Strand V, Tugwell P (2014) Developing core outcome measurement sets for clinical trials: OMERACT filter 2.0. J Clin Epidemiol 67(7):745–753
20. Deyo RA, Battie M, Beurskens A, Bombardier C, Croft P, Koes B, Malmivaara A, Roland M, Von Korff M, Waddell G (1998) Outcome measures for low back pain research: a proposal for standardized use. Spine 23(18):2003–2013
21. Bombardier C (2000) Outcome assessments in the evaluation of treatment of spinal disorders: summary and general recommendations. Spine 25(24):3100–3103
22. Lin C-WC, Ostelo RW, Maher CG, Grotle M, Koes BW Core outcomes for clinical research in low back pain: revisited. In: Abstract Book, XII International Low Back Pain Forum, Odense (Denmark), 16–19 October 2012
23. Ferrer M, Pellisé F, Escudero O, Alvarez L, Pont A, Alonso J, Deyo R (2006) Validation of a minimum outcome core set in the evaluation of patients with back pain. Spine 31(12):1372–1379
24. Mannion AF, Porchet F, Kleinstück F, Lattig F, Jeszenszky D, Bartanusz V, Dvorak J, Grob D (2009) The quality of spine surgery from the patient's perspective. Part 1: the Core Outcome Measures Index in clinical practice. Eur Spine J 18(3):367–373
25. Balagué F, Mannion AF, Pellisé F, Cedraschi C (2012) Non-specific low back pain. Lancet 379(9814):482–491
26. Chiarotto A, Terwee CB, Deyo RA, Boers M, Lin CW, Buchbinder R, Corbin TP, Costa LO, Foster NE, Grotle M, Koes BW, Kovacs FM, Maher CG, Pearson AM, Peul WC, Schoene ML, Turk DC, van Tulder MW, Ostelo RW (2014) A core outcome set for clinical trials on non-specific low back pain: study protocol for the development of a core domain set. Trials 15(1):511
27. Kirwan J, Heiberg T, Hewlett S, Hughes R, Kvien T, Ahlmèn M, Boers M, Minnock P, Saag K, Shea B, Suarez Almazor M, Taal E (2003) Outcomes from the patient perspective workshop at OMERACT 6. J Rheumatol 30(4):868–872
28. Kirwan JR, Hewlett SE, Heiberg T, Hughes RA, Carr M, Hehir M, Kvien TK, Minnock P, Newman SP, Quest EM, Taal E, Wale J (2005) Incorporating the patient perspective into outcome assessment in rheumatoid arthritis–progress at OMERACT 7. J Rheumatol 32(11):2250–2256
29. Sanderson T, Morris M, Calnan M, Richards P, Hewlett S (2010) What outcomes from pharmacologic treatments are important to people with rheumatoid arthritis? Creating the basis of a patient core set. Arthritis Care Res 62(5):640–646
30. Binkley J, Finch E, Hall J, Black T, Gowland C (1993) Diagnostic classification of patients with low back pain: report on a survey of physical therapy experts. Phys Ther 73(3):138–150
31. Sinha IP, Smyth RL, Williamson PR (2011) 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 8(1):e1000393
32. Henschke N, Ostelo R, van Tulder MW, Vlaeyen J, Morley S, Assendelft W, Main CJ (2010) Behavioural treatment for chronic low-back pain. Cochrane Database Syst Rev 7:CD002014
33. Oosterhuis T, Costa LO, Maher CG, de Vet HC, van Tulder MW, Ostelo R (2014) Rehabilitation after lumbar disc surgery. Cochrane Database Syst Rev 3:CD003007
34. Cieza A, Stucki G, Weigl M, Disler P, Jackel W, van der Linden S, Kostanjsek N, de Bie R (2004) ICF Core Sets for low back pain. J Rehabil Med 44 Suppl:69–74
35. Buchbinder R, Batterham R, Elsworth G, Dionne CE, Irvin E, Osborne RH (2011) A validity-driven approach to the understanding of the personal and societal burden of low back pain: development of a conceptual and measurement model. Arthritis Res Ther 13(5):R152
36. Mullis R, Barber J, Lewis M, Hay E (2007) ICF core sets for low back pain: do they include what matters to patients? J Rehabil Med 39(5):353–357
37. ICF Classification Browser. http://www.apps.who.int/classifications/icfbrowser/. Accessed 26 September 2014
38. PROMIS Domain Framework. http://www.nihpromis.org/measures/domainframework/. Accessed 26 September 2014
39. Wilson IB, Cleary PD (1995) Linking clinical variables with health-related quality of life: a conceptual model of patient outcomes. JAMA 273(1):59–65
40. Dworkin RH, Turk DC, Farrar JT, Haythornthwaite JA, Jensen MP, Katz NP, Kerns RD, Stucki G, Allen RR, Bellamy N et al (2005) Core outcome measures for chronic pain clinical trials: IMMPACT recommendations. Pain 113(1–2):9–19
41. Turk DC, Dworkin RH, Allen RR, Bellamy N, Brandenburg N, Carr DB, Cleeland C, Dionne R, Farrar JT, Galer BS et al (2003) Core outcome domains for chronic pain clinical trials: IMMPACT recommendations. Pain 106(3):337–345
42. Pincus T, Kent P, Bronfort G, Loisel P, Pransky G, Hartvigsen J (2013) Twenty-five years with the biopsychosocial model of low back pain—is it time to celebrate? A report from the twelfth international forum for primary care research on low back pain. Spine 38(24):2118–2123
43. Waddell G (1987) 1987 Volvo award in clinical sciences: a new clinical model for the treatment of low-back pain. Spine 12(7):632–644
44. Froud R, Patterson S, Eldridge S, Seale C, Pincus T, Rajendran D, Fossum C, Underwood M (2014) A systematic review and meta-synthesis of the impact of low back pain on people's lives. BMC Musculoskelet Disord 15:50
45. Deyo RA, Dworkin SF, Amtmann D, Andersson G, Borenstein D, Carragee E, Carrino J, Chou R, Cook K, DeLitto A (2014) Report of the NIH Task Force on research standards for chronic low back pain. J Pain 15(6):569–585
46. Macefield RC, Jacobs M, Korfage IJ, Nicklin J, Whistance RN, Brookes ST, Sprangers MA, Blazeby JM (2014) Developing core outcomes sets: methods for identifying and including patient-reported outcomes (PROs). Trials 15(1):49
47. Prinsen CA, Vohra S, Rose MR, King-Jones S, Ishaque S, Bhaloo Z, Adams D, Terwee CB (2014) Core Outcome Measures in Effectiveness Trials (COMET) initiative: protocol for an international Delphi study to achieve consensus on how to select outcome measurement instruments for outcomes included in a 'core outcome set'. Trials 15(1):247