


BMJ Open Development of CORE-CM core outcome domain sets for trials of Chinese medicine for lumbar spinal stenosis

Ya-Nan Sun,¹ Yi An,^{2,3} Zhi-Wen Weng,⁴ Yan-Ji Zhou,^{2,3} Xi-You Wang,⁴ Chang-He Yu⁴ 

To cite: Sun Y-N, An Y, Weng Z-W, *et al.* Development of CORE-CM core outcome domain sets for trials of Chinese medicine for lumbar spinal stenosis. *BMJ Open* 2023;**13**:e075856. doi:10.1136/bmjopen-2023-075856

► Prepublication history and additional supplemental material for this paper are available online. To view these files, please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2023-075856>).

Y-NS, YA, Z-WW and Y-JZ contributed equally.

Received 20 May 2023
Accepted 22 September 2023



© Author(s) (or their employer(s)) 2023. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

For numbered affiliations see end of article.

Correspondence to
Dr Chang-He Yu;
yakno2@163.com and
Dr Xi-You Wang;
dwxy658@163.com

ABSTRACT

Objectives Most Asian countries have employed Chinese medicine (CM) and Western medicine to treat lumbar spinal stenosis (LSS). Evidence synthesis and comparison of effectiveness are difficult since outcomes examined and presented through trials possess heterogeneity. This study aimed to solve the outcome problems for CM clinical trials in LSS by building a core outcome set (COS).

Methods To achieve an agreement on a set of core outcome domains, a four-phase study was carried out. First, we identified candidate outcome domains by systematically reviewing trials. In addition, we identified outcome domains associated with patients by conducting semistructured interviews with patients. Next, outcome domains were processed through a national two-round Delphi survey, in which 18 patients and 21 experts were recruited. Finally, the above domains were converted as a core outcome domain set based on a consensus meeting, in which 24 stakeholders were recruited.

Results Seventeen outcome subdomains were identified by the systematic review and interviews. The Delphi survey assigned a priority to four outcome domains in the first round and four outcomes additionally in the second round. The core outcome domains were determined through discussion and redefinition of outcomes in the consensus meeting: pain and discomfort, health-related quality of life, lumbar function, activities of daily living, measures of walking, patient global assessment, adverse events and CM-specific outcomes.

Conclusion COS-CM-LSS is likely to enhance the consistency of outcomes reported in clinical trials. In-depth research should be conducted for the exploration of the best methods to examine the above outcomes.

INTRODUCTION

Lumbar spinal stenosis (LSS) arises from spinal anatomical or functional narrowing with a negative effect on the spinal cord and nerve roots, characterised by pain and discomfort in legs, buttocks and lumbar spine, as well as disability of walking capacity.¹ The above discomfort and pain can be increased by walking and alleviated through sitting or lumbar flexion.² LSS affects a global

STRENGTHS AND LIMITATIONS OF THIS STUDY

- ⇒ We used a mixed-method approach to determine which outcomes would be included in the CORE outcome domain sets for trials of Chinese medicine for lumbar spinal stenosis (CM-LSS).
- ⇒ In this investigation, we incorporated the perspectives of different stakeholders, including patients, physicians and researchers.
- ⇒ The participants were sampled based on duration and socioeconomic status, disease severity, as well as LSS manifestations, which ensures that the core domains are generalisable to LSS people.
- ⇒ CM or integrated medicine studies have been mostly used there, which limits the results since stakeholders are distributed in different geographical areas.

population of nearly 103 million³ and 11% of the elderly in the USA.⁴ Most LSSs are treated non-operatively, with physical therapy, analgesia and activity modification as the first-line therapies, whereas patients subjected to limited activities and continuous pain are likely to be an alternative in terms of surgery.⁵

Chinese medicine (CM), a non-surgical treatment, is critical in the treatment of LSS. Acupuncture and acupotomy contribute to the LSS patients on pain, symptoms and functional outcomes up to 6 months post-treatment.^{6 7} Moreover, CM alone or combined treatment is likely to more pronouncedly alleviate pain and ameliorate functional outcomes than conventional therapies.⁸ Furthermore, manual therapy in combination with exercises under supervision can improve walking capacity, symptoms and pain in comparison to exercises.⁹

A review of clinical trials of LSS found inconsistency between results reporting or measuring instrument application under one outcome and poorly defined outcomes.¹⁰ An important effect of the above inconsistencies is to limit the potential of robust

meta-analysis. In a network meta-analysis of conservative treatment of LSS, only four results were analysed, while the other results could not be analysed due to the limited data or no meta-analysis to determine the outcome, or the variety of definitions of an outcome.⁸ Existing problems, supported by most CM trials, include poorly defined outcomes, insufficient evidence of instruments, selective reporting of outcomes or no criteria for selection for core outcomes.¹¹ Data that cannot be interpreted or used cause unacceptable and unethical waste of research. Selective reporting of results and associated reporting biases may also occur if consistent results are not specified in advance.¹²

The core outcome set (COS) includes standardised outcomes. It has been found as the minimal measurement and report criterion in terms of the respective trial for a specific health area,¹³ increasing outcome reporting consistency, accountability and transparency. Outcomes, which conform with certain standards and are examined in studies under a particular condition, can reduce this research waste, such that the bias of reporting can be prevented. The above outcomes can ensure that existing research reporting outcomes is able to be integrated into meta-analyses with certain significance.¹⁴ The review of the Core Outcome Measures in Effectiveness Trials (COMET) database and searching OMERACT for COSs of trauma and orthopaedics ensured the lack of COS on LSS.¹⁵

This study presents a multiple-stakeholder, Chinese nationally endorsed, consensus-based CORE outcome set suitable for CM intervention trials in adults with LSS (CORE-CM-LSS), as well as its development.

METHOD

The study protocol was registered in the COMET database (<https://www.comet-initiative.org/Studies/Details/1363>), whereas the protocol was not published. The development of our COS was reported and consistent with the COS-STAndards for Reporting¹⁶ as well as COS-STAndards for Development¹⁷ guidelines (online supplemental table S1 and S2). This is a further study underlying COS development for low back pain (LBP), and the COS focused on specific LBP due to LSS which is treated by CM.

Scope and design

Study advisory group (SAG) was formed, in which a wide variety of stakeholders, two orthopaedists, one acupuncture and Tuina expert, one patient, one methodologist, one clinical trial researcher, as well as one statistician were invited. SAG confirmed the outcome set that serves as a candidate in terms of data analyses and explanation, process coordination and Delphi survey. Furthermore, some of them participated in the consensus process.

Following SAG, this COS's scope is clarified as follows: setting: randomised controlled trials (RCTs); health condition: symptomatic LSS.¹ Target interventions are

CM for LSS, which comprise acupuncture, Tuina (CM massage), Gongfa (CM exercise), bloodletting, cupping, oral herbal medicine, local washing or compressing with CM. Furthermore, CM alone or CM combined with other conventional treatments were involved.

This study fell into three vital sections to obtain consensus on the outcome domains that were to be examined, which were completed in the proper sequence. The following inquiries were answered, including which outcome domains are likely to benefit LSS patients, which outcomes are more important, as well as which results should be included in the COS.

All participants declared no interest conflict during the study. Patients contributed to the design of the study and were involved in the stages of patients' interview and consensus meeting.

Systematic literature review

A list of outcomes was established through Systematic literature review (SLR). Moreover, the results of the SLR were partly published to assess the effectiveness of non-pharmaceutical Chinese medical therapies alone or in combination for the treatment of LSS.¹⁸

Eligible trials

The RCTs of the LSS patients diagnosed by clinical symptoms of neurogenic claudication and imaging findings were included, no matter whether LSS patients have complicating diseases. Interventions included the treatment with CM alone or treatment including CM. The control intervention involved routine treatment (eg, injection therapy, physical therapy, exercise therapy, health education, self-management), or a combination of the above. There were no restrictions on publication type, language or status.

Literature search and selection

The trials were identified by searching RCT and spinal stenosis terms from CNKI, VIP, WangFang, Sinomed, PubMed, Cochrane Library and EMBASE online databases, from their inception to 1 January 2022 (search strategy in online supplemental table S3). Grey literature and reference lists of included literature were searched. Furthermore, the authors of included literature were contacted to identify eligible trials.

The EndNote V.20 managed literature and excluded the duplicate ones. Eligibility was evaluated initially by two independent reviewers (including Y-NS and Y-JZ) through reading abstracts and titles, and the trials were included after the full texts were read. Any disagreements would be addressed through discussions when the full text was critically reviewed, or through consultation with a third author (CY).

Data collection and analyses

The data from eligible trials were extracted independently and inputted into Microsoft EXCEL for management. Extracted data included the first authors, contact information, outcome measurement instruments (OMI) (name

and measuring time frame), comparator, intervention, sample size, country and year of publication. If response rate or composite index outcomes existed in trials, the criteria and classification of them were recorded.

After data extraction, the measurement instruments were categorised by SAG into outcome subdomains and domains, and the respective outcome was defined by SAG following the COMET criteria.^{19 20} Besides, SAG removed the duplicates and standardised the similar or overlapping outcomes. Information and purpose of an instrument (ie, to evaluate physical function, or pain intensity) was confirmed by original prescription, from either method or results parts, and considered into right subdomains. Any disagreements were resolved by consulting a third author (CY).

The number of instruments of the respective trial and subdomain and outcome domains of all trials was obtained. The frequency and percentages of categorical instruments and outcomes were conducted with SPSS V.18.0.

The semistructured interview

The additional associated outcome domains were elicited through qualitative semistructured interviews of patients.

Participants

The LSS patients previously or currently under CM treatment were recruited. While the LSS patients due to trauma or congenital spinal disease, having hearing or communication problems, or refusing to join the interviews were excluded.

We employed convenient and purposeful sampling methods based on several ages, gender, years of LSS and imaging findings related to the hospital outpatients from seven territories of China (predefined features in online supplemental table S4). Features were defined by the SAG to ensure diversity represented. The qualitative data were analysed, while the interviews continued, and the sampling was ended following data saturation criteria, based on the definition from two consecutive interviews without any additional subdomain.

Interview process

Interviews were carried out face to face in outpatient or via remote video software (WeChat) and recorded by qualified researchers (CY). Explanation and information consent should be given to patients before the interviews. We initiated the interview with questions (eg, 'what outcomes are important or most concern to you, or how do you determine the effectiveness of treatment, or what aspect they would like to get better improvement'). A list of subdomains from SLR was provided as the outline when patients could not answer or had no ideas about the important outcomes. After patients completed reading the list, another open-ended question was asked to allow patients to provide additional outcomes.

Data analysis

The additional outcomes and the demographic and medical information of patients were collected. The words expressed by patients were analysed through qualitative content analysis. For an overall perspective and familiarity with the content, the recorded interviews were listened to and the transcripts were reviewed and reread. The two researchers (Y-NS and YA) first carried out the initial assessment individually before being mapped into the initiative list in three steps. Specifically, sentences and paragraphs were found, abstracted and then coded as meaning units. The codes were organised into subjects under the context of COMET outcomes subdomains; the codes of each topic fell into initial COMET outcome domains. Subsequently, the draft outcomes domains of the two researchers were combined and compared. Afterwards, outcomes subdomains with similar names were examined, and those with the same content were grouped together. Any discrepancies were resolved with discussion.

Expert consensus

Panel participants

A group of participants specialised in CM, integrated Chinese and Western medicine, nursing, orthopaedics, acupuncture, Tuina, pain management, rehabilitation and clinical researchers were recruited in the Delphi survey, and the professional and geographical distribution of panellists was considered. Furthermore, all SAG members engaged in the consensus meeting via WeChat conference instead of face to face due to the COVID-19 pandemic.

It was expected to select 30 participants based on a snowball sampling method. The experts were preliminarily identified by reviewing the authors of high-impact papers and recommended by the preliminary stakeholders. The patients were selected following a pool of outpatients. All participants completed round 1 were invited to join round 2 of Delphi.

Identifying important outcomes in Delphi Survey

In rounds 1 and 2, for the respective outcome, panellists were recruited for assigning scores between 1 (of no importance) and 9 (of high importance), where 1–3 represents that it is 'of no importance to be included in the COS', 4–6 represents that it is 'of importance but no critical importance to be included in the COS' and 7–9 represents that it is 'of critical importance to be included in the COS'.²¹ In round 1, participants were recruited to add new outcome(s), if they regarded it/them as important.

We removed outcomes reaching consensus thresholds between rounds for the minimisation of attrition. Predefined 'consensus in' thresholds are reached if >80% of the panellists scores 7–9 and ≤15% scores 1–3; 'consensus out' thresholds are met if >80% of the panellists scores 1 to –3 and ≤15% of the panellists scores 7–9. This threshold is consistent with those set for other core

outcomes, protecting minority stakeholders' different views from the rejection by a greater stakeholder group.²²

The outcomes that scored neither consensus in nor out were retained to the next round. The newly added outcomes by the participants that existed in the preliminary list were removed. Otherwise, the new outcomes were entered in the next round for scoring. Feedback was presented between the 1st and 2nd rounds, with average scores of outcomes.

Identifying core outcomes in consensus meetings

A total of 9 LSS patients and 15 experts, most from previous study stages, were recruited in an online consensus meeting. One author (CHY), who is independent of the discussion and voting poll, moderated the meeting using the nominal group technique (NGT). The NGT refers to a meeting with a rigorous structure, which is carried out for allowing key stakeholders' identification and rating of a list of priorities; it also aims to ensure that the opinions of all participants are included.²³ The meeting aimed to reach an agreement in terms of a preliminary core set of 7–10 domains.

The NGT process started with the discussion of domains that were in consensus out or not a consensus with the purpose to discard them or move them into consensus in. Subsequently, the rest outcomes were investigated, redefined, kept or integrated into greater categories if an agreement was reached by most panellists. Anonymous votes were made in terms of agreements with domain placement. When the meeting was about to be completed, a draft preliminary core set of domains was made and then shown to the participants.

After the Delphi survey was completed, the outcomes of 'consensus in' and 'no consensus' were scored by using yes, no, or unsure for inclusion of the COS (yes for selected; no for not selected). In terms of outcomes to be included in the core domain set, a prespecified threshold of >80% on yes was set.

Patient and public involvement

None.

RESULTS

Identification of candidate outcomes

Outcomes from the SLR

A total of 5674 trials were identified through the SLR, 86 trials could be included after duplicates were removed, and abstract, title and full-text were screened (Preferred Reporting Items for Systematic Reviews and Meta-Analyses flow diagram in online supplemental figure S1). Eighty-six trials involved 6892 LSS patients (range 26–200), with 80% (2980/6658, two trials did not report gender) female, aged from 33 to 72 years. Most trials compared a wide variety of CM treatments alone with placebo or routine treatment, and others compared the combination of CM treatment versus CM treatment alone

or western treatment. Online supplemental table S4 lists the characteristics of the included trials in detail.

Table 1 lists a total of 86 trials that reported 54 different OMI. The number of OMIs was applied and reported, ranging from 1 to 6 (median 3). The most used OMI included response rates (64/86, 74.42%), various versions of JOA (42/86, 48.84%), Visual Analogue Scale (37/86, 43.02%), adverse events (AEs) (18/86, 20.93%), as well as measures of walking (12/86, 13.95%) (online supplemental table S5). Fifty per cent of OMI were patient-reported outcomes, and 30% were performance-based measurements. While the rest were clinician-based measurements (eg, CT and MRI).

SAG reviewed 54 OMI and identified 20 subdomain outcomes and 10 COMET domains (table 1). Among 86 trials, pain (98.8%; n=85) and function (97.7%; n=84) were the most frequently evaluated COMET domains, followed by AEs (22.1%; n=19), and Physiological index (12.8%; n=11). Three COMET domains (including resource use, mortality and infection) were not reported in any trial.

Patients interview

In this study, 18 interviews were carried out with LSS patients from seven territorial regions around China. Eight of the 18 interviews with them were done via the WeChat app. Online supplemental table S6 presents the demographic details of the participants. The content analyses of interview transcript and outcomes from open-ended questions indicated that 16 subdomain outcomes were identified and then classified into 11 COMET domains (table 2).

SAG identified subdomain outcomes as candidate outcomes from SLR and interviews, defined outcomes and constructed a final inventory of 17 outcomes^{20 24–31} for the Delphi survey (table 3). Among candidate outcomes, physiological outcome was separated by SAG into biomarkers and radiographic changes. The biomarkers outcome was identified by SAG by combining inflammatory markers, haemorrhological markers, immunological markers and physiological outcomes (figure 1).

Important outcomes identified from Delphi surveys

A total of 25 experts and 18 patients were recruited for online Delphi survey, and 21 experts and all patients responded and completed the first-round survey (participant characteristics detailed in online supplemental table S7). Delphi survey identified four outcome domains (including pain, function, activities of daily living (ADL) and quality of life (QOL)) in the first round, and another four outcome domains (including symptoms, measures of walking, global rating of change and AE) in the second round, all of which met the consensus threshold. Table 4 lists the scores for all candidate outcomes and 'consensus-in' outcomes. The 'consensus-in' outcomes drew the Delphi consensus threshold and employed the above for the development of several initial outcome domains to be covered in the core outcome domain set.

Table 1 Classification of outcome measurement instruments into subdomains and COMET outcomes

No	COMET domain outcome	No of 86 RCTs reporting COMET outcomes (%)	Subdomain outcome	No of 54 OMI into subdomain outcomes (%)	OMI (n=54)
1	Function	85 (98.8)	Function ADL ROM Symptoms Measure of walking Global rating of change	10 (18.5) 3 (5.6) 1 (1.9) 1 (1.9) 3 (5.6) 3 (5.6)	JOA/improvement of JOA/mJOA/ODI/ mRMDQ/RMDQ/ ZCQ/SSS/self-made lumbar function evaluation scale/ physical function, role-physical, vitality of SF-36 ADL subscale of JOA, ODI or RMDQ Rang of lumbar spine extension Self-made symptoms rating scale SPWT/walking capacity/Pain-free walking distance Responder rates/Global Rating of Change Scale/general health of SF-36
2	Pain	84 (97.7)	Pain	6 (11.1)	VAS/NRS/UAB-PBS/pain subscale of JOA/pain subscale of SSS/bodily pain subscale of SF-36
3	AEs	19 (22.1)	AE	1 (1.9)	AE
4	Physiological	11 (12.8)	Inflammatory markers Haemorrhological marker Immunological markers Physiological markers Radiographic changes	8 (14.8) 3 (5.6) 4 (7.4) 3 (5.6) 1 (1.9)	IL-6\IL-1B\TNF\CRP\IL-1\IL-4\IL-10\ ESR Blood viscosity/plasma viscosity/RBC haematocrit Changes in T lymphocyte subsets Hepatic and renal function tests/ Serum endothelin Parameters of CT
5	CM indicator	5 (5.8)	CM meridian/CM Zheng	2 (3.7)	Near-infrared imaging system on meridian CM Zheng scores
6	Mental health	4 (4.7)	Mental health	2 (3.7)	HADS/ mental health subscales of SF-36
7	Satisfaction	4 (4.7)	Satisfaction index	1 (1.9)	Satisfaction subscale of SSS
8	Quality of life	3 (3.5)	Quality of life	1 (1.9)	SF-36
9	Psychosocial	3 (3.5)	Psychosocial	1 (1.9)	Social function, role-emotional subscales of SF-36
10	Compliance	1 (1.2)	Adherence and attrition	1 (1.9)	Treatment Adherence index
11	Resource use	0 (0)	Resource use	0 (0)	NR
12	Mortality	0 (0)	Mortality	0 (0)	NR
13	Infection	0 (0)	Infection	0 (0)	NR

ADL, activities of daily living; AE, adverse event; CM, Chinese medicine; COMET, Core Outcome Measures in Effectiveness Trials; CRP, C reactive protein; HADS, Hospital Anxiety and Depression Scale; JOA, Japanese Orthopedic Association Score; mJOA, modified JOA; mRMDQ, modified Roland Morris Disability Questionnaire; NR, not reported; NRS, Numerical Rating Scale; ODI, Oswestry Disability Index; OMI, Outcome Measurement Instruments; RBC, red blood cell; RMDQ, Roland Morris Disability Questionnaire; ROM, range of movement; SF-36, 6-Item Short Form Survey; SPWT, Self-Paced Walk Test; SSS, Spinal Stenosis Scale; TNF, tumour necrosis factor; UAB-PBS, University of Alabama at Birmingham- Pain Behavior Scale; VAS, Visual Analogue Scale; ZCQ, Zurich Claudication Questionnaire.

COS determined by consensus meetings

Consensus meeting summary

The participants redefined some outcomes from the list of 17 domains (table 4) in the NGT process. LSS patients were subjected to the pain accompanied by numbness or tingling in the lower legs or feet. Some severe limitations in activity resulted in the gradual worsening of pain over

time. The severity of pain, walking disability underlying definition of symptoms outcome may overestimate or underestimate outcomes. Thus, the experts suggested that the overall symptom outcome can be replaced by the outcomes of pain, lumbar function, walking disability and ADL, respectively, which were evaluated easily and adequately.

Table 2 Subdomain and COMET outcomes identified from interviews

COMET domain outcome	Subdomain outcomes	No of 18 patients (%)	Example of interview transcript (Chinese words presented in English)
Function	Function	17 (94.4)	<i>"This waist does not seem to be as flexible as before....."</i>
	ADL	14 (77.8)	<i>"I felt hard to get dressed, brush teeth, wash face, or go to toilet....."</i>
	ROM	3 (16.7)	<i>"I felt hard to back straight or bend over (in some degrees)"</i>
	Symptoms	16 (88.9)	<i>"Pain is on my low back and legs, makes me hard to move anymore (or in certain distance)"</i>
	Measure of walking	16 (88.9)	<i>"I can't walk long way, I felt my legs do not work, and then I have to stop for a rest"</i>
	Global rating of change	2 (11.1)	<i>"I would like to feel wellbeing,, even for a while"</i>
Pain	pain	18 (100)	<i>"I cannot get into sleep due to pain when I tried to turn over on bed"</i>
Adverse events	AE	16 (88.9)	<i>"Is that (the treatment) safe? Are there any side effects?"</i>
Physiological	Inflammatory markers	0 (0)	nr
	Haemorheological markers	0 (0)	nr
	Immunological markers	0 (0)	nr
	Physiological markers	0 (0)	nr
	Radiographic changes	2 (11.1)	<i>"(Treatment) helps me release the narrowing of space and pressure of nerves, I would feel well"</i>
CM indicator	CM meridian	0 (0)	nr
	CM Zheng	1 (5.5)	<i>"Can Chinese medicine help to treat blood stasis pattern?"</i>
Mental health	Mental health	12 (66.7)	<i>"It always hurts and pain seems not to be relieved, so I felt some irritable, and worried as it is getting more serious"</i>
Satisfaction	Satisfaction index	1 (5.5)	<i>"I felt satisfied if it (treatment) can relieve my pain and help me walk long"</i>
Quality of life	Quality of life	5 (27.8)	<i>"Low back pain affects life, and the most impact of pain is on my quality of life"</i>
Psychosocial	Psychosocial	4 (22.2)	<i>"I was really worried because I was younger and worried about my professional longevity.....I couldn't hang out with my family, it was always a drag on my family, and I had to let them take care of me"</i>
Compliance	Adherence and attrition	1 (5.5)	<i>"You have to listen to the doctor, and the efficacy would be guaranteed"</i>
Resource use	Resource use	3 (16.7)	<i>"Treatment wastes lots of time on transportation and waiting, but I have to (do it) due to pain"</i>

ADL, activities of daily living; AE, adverse event; CM, Chinese medicine; COMET, Core Outcome Measures in Effectiveness Trials; nr, not reported; ROM, range of movement.

For pain outcome, several experts suggested that some patients felt discomfort rather than pain, so pain outcome was redefined 'pain' to 'pain and discomfort'. Furthermore, the physical function of LSS was redefined as lumbar function and walking function, and the latter referred to measures of walking or walking performance.

QOL, a board definition, was brought up for discussion. First, experts redefined QOL to health-related QoL (HRQoL), consisting of mental and physical health perceptions (eg, mood, energy level) and their correlates

(eg, socioeconomic status, social support, functional status, as well as health conditions and risks). The concept HRQoL presented potentially overlapping with some of the above domains. Thus, participants agreed and favoured the inclusion of physical, emotional and social life were covered in HRQoL for LSS patients.

Global rating of change was also discussed. The concept was felt to reflect disease activity and overarching global health status of the patient, specific to that patient. Based on the above discussion, the global rating of change was

Table 3 Candidate lumbar spinal stenosis outcomes and definitions

No	Candidate outcome	Definition	Resources
1	Pain	Experiencing an unpleasant physical sensation that aches, hurts in one or more joints or the spine; an unpleasant sensory and emotional experience associated with actual or potential tissue damage, or described for such damage. ²⁹	SLR+Int
2	Function	Being able to perform physical activities (includes lower extremity functioning, balance); 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. ²⁴	SLR+Int
3	ADL	Fundamental skills required to independently care for oneself, such as eating, bathing and mobility. ²⁵	SLR+Int
4	ROM	Quantity of movement of the lumbar spine and/or of other adjacent body parts (ie, thoracic spine, pelvis, rib cage or lower limbs). ²⁴	SLR+Int
5	Symptoms	Presence of symptoms on back, leg and walking. ²⁷	SLR+Int
6	Measure of walking	Measuring ability, capability, distance, performance of walking. ²⁴	SLR+Int
7	Global rating of change	Considering the ways that the health condition affects the individual on a given day. ³⁰	SLR+Int
8	AE	Any untoward medical occurrence associated with the use of an intervention in humans, whether or not considered intervention related. ³¹	SLR+Int
9	Biomarkers	Indicators aimed at providing insight into peripheral and central neurobiological mechanisms of pain. ²⁴	SLR
10	Radiographic changes	medical imaging such as MRI, CT, X-ray detecting the changes of bones, joints, muscles, tendons, nerves and other body structures localised on the lumbar spine and/or on other adjacent body parts (ie, thoracic spine, pelvis, rib cage or lower limbs). ²⁴	SLR+Int
11	CM-specific outcomes	CM outcomes related to CM Zheng or meridians based on CM theory. ²⁸	SLR
12	Mental health	A person's condition with regard to their psychological, social and emotional well-being. ²⁶	SLR+Int
13	Satisfaction index	Satisfaction with care received, including of the process and outcomes of the treatment experience and care providers. ²⁴	SLR+Int
14	Quality of life	Broad multidimensional concept that usually includes subjective evaluations of both positive and negative aspects of life, including health-related quality of life. ²⁶	SLR
15	Adherence and attrition	Withdrawal from treatment. ²⁰	SLR
16	Psychosocial	An individual's interactions with their environment and the ability to fulfil their role within such environments as work, social activities, and relationships with partners and family. ²⁴	SLR
17	Resource use	Treatment burden such as impact of treatment and monitoring of disease or treatment (ie, financial loss due to treatment cost, work loss or time commitment). ²⁰	Int

ADL, activities of daily living; AE, adverse events; CM, Chinese medicine; Int, interview; ROM, range of movement; SLR, systematic literature review.

renamed and defined as patient global assessment (PGA) of disease-related health status and kept as a core domain.

COS identified by final voting

According to the list of outcomes, an agreement was reached on the core set based on an electronic voting programme in consensus meetings. Table 4 lists the scores for the respective outcome domain. An agreement was reached on eight domains of importance and inclusion in the core domain set for clinical trials (including pain and discomfort, HRQoL, lumbar function, ADL, walking

function, PGA, AE and CM-specific outcomes). Online supplemental table S8 lists the sensitive analysis of score of outcomes between patients and experts.

DISCUSSION

Summary of the main results

This study presents the development of the CORE-CM-LSS and the steps involved for reaching a consensus of patients and experts. The patient perspective was

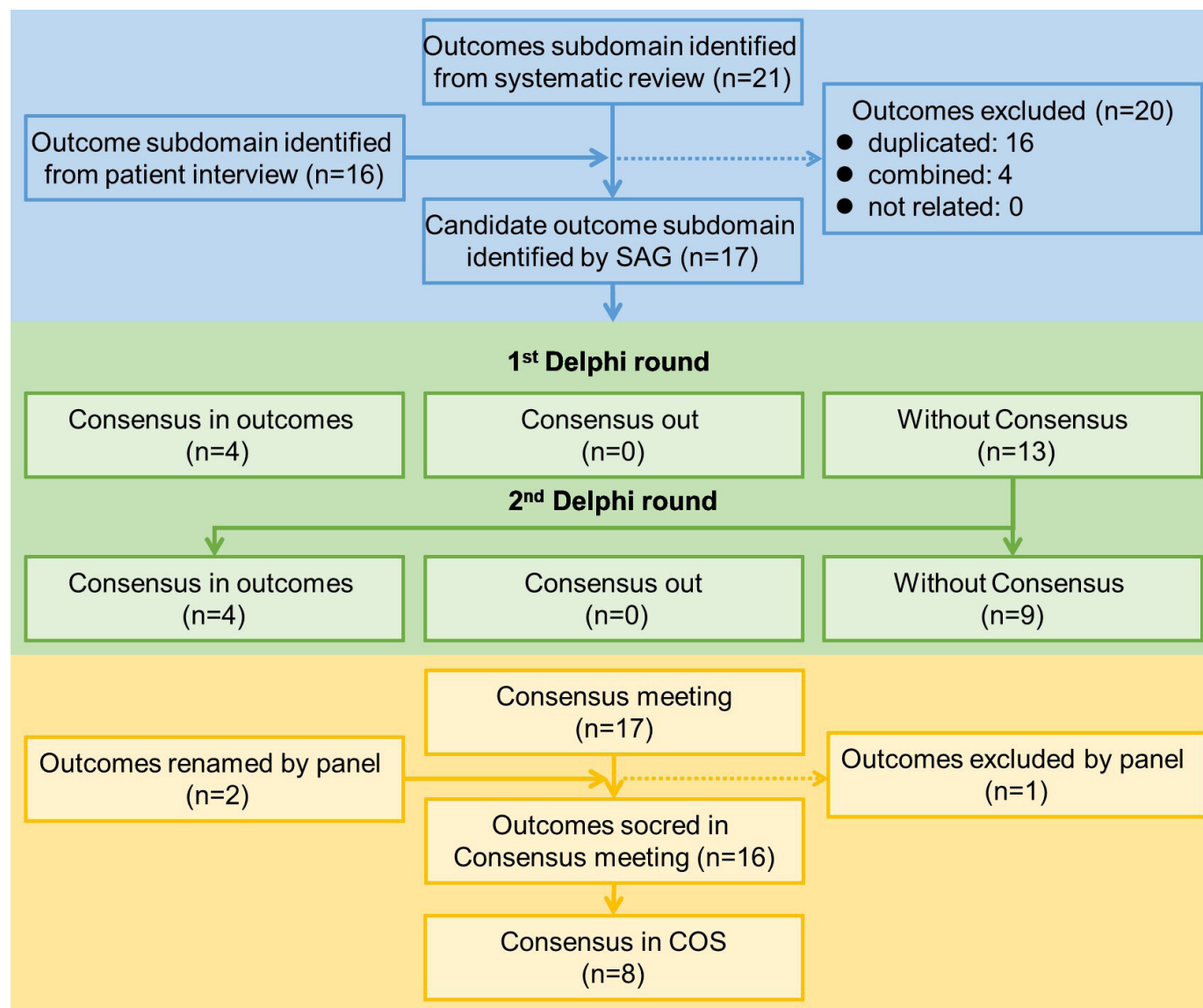


Figure 1 Flow chart of core outcomes selection process. COS, core outcome set; SAG, study advisory group.

integrated in the respective research phase. The sampling process included panellists nationally, which can endow our findings with greater generalisability in China.

Outcomes included in the COS

Our review and consensus results confirm that the pain/discomfort, function, walking disability and ADL of LSS patients arouse the main concern of patients and physicians and have been most reported in trials. The above outcomes were common symptoms of LSS or impacts of symptoms. The AEs are required for the assessment of the harms of all interventions, and they arouse the most concern of patients. The HRQoL is vital outcomes for the trials on pain for its generic construct, which is beneficial to compare populations from different diseases. However, the LSS-associated HRQoL is necessary but scarce, which can precisely indicate the outcomes changes and should replace the generic HRQoL. PGA, counterparting to the physician's global assessment, was first developed to measure self-assessed pain in rheumatoid arthritis. PGA

scales were employed in a broad range of diseases over the past years. The application of PGA in clinical practice covered two different concepts. One of the concepts is concerned with global health. The other concept is relating to overall changes of disease activity or severity.³²

The CM-specific outcomes were covered in COS in terms of its specific for CM, which may deviate from that employed in Western medicine,^{28 33} which are attributed to a general agreement to not discuss instruments. However, CM-specific outcome measures have been rarely investigated. The CM pattern (syndromes or Zheng in Chinese) is a diagnostic conclusion based on pathological changes in a disease, at a certain stage.³⁴ A pattern often contains several CM symptoms (eg, tongue manifestation or pulse condition). CM physicians should measure patterns and CM symptom changes during the treatment of patients. The meridian detection and CM pattern is a diagnostic and outcome assessment tool for one health condition. However, the definitions and measurement instruments

Table 4 Candidate outcomes ratings in two rounds Delphi and voting at consensus meetings

Candidate outcomes	Round 1 (n=39)			Round 2 (n=39)			Redefined outcomes	Consensus meeting voting (n=24)
	% score 1–3	% score 4–6	% score 7–9	% score 1–3	% score 4–6	% score 7–9		% yes
Pain	0%	5%	95%	nr	nr	nr	Pain/discomfort	100%
Function	3%	5%	92%	nr	nr	nr	Lumbar function	100%
ADL	0%	5%	95%	nr	nr	nr	ADL	92%
ROM	10%	26%	64%	5%	18%	77%	ROM	50%
Symptoms	5%	15%	79%	5%	8%	87%	nr	nr
Measure of walking	5%	28%	67%	0%	8%	92%	Walking function	96%
Global rating of change	3%	26%	72%	3%	15%	82%	PGA	88%
AE	0%	21%	79%	0%	15%	85%	AE	100%
Biomarkers	28%	41%	31%	21%	51%	28%	Biomarkers	4%
Radiographic changes	5%	33%	62%	3%	31%	67%	Radiographic changes	38%
CM-specific outcomes	8%	28%	64%	3%	21%	77%	CM-specific outcomes	88%
Mental health	5%	36%	59%	3%	38%	59%	Mental health	63%
Satisfaction index	0%	28%	72%	0%	26%	74%	Satisfaction index	21%
Quality of life	3%	13%	85%	nr	nr	nr	HRQoL	96%
Adherence and attrition	8%	33%	59%	5%	31%	64%	Adherence and attrition	4%
Psychosocial	49%	38%	13%	49%	41%	10%	Psychosocial	21%
Resource use	8%	26%	67%	5%	23%	72%	Resource use	58%

ADL, activities of daily living; AE, adverse events; CM, Chinese medicine; HRQoL, health-related QOL; nr, not reported; PGA, patient global assessment; QOL, quality of life; ROM, range of movement.

of CM-specific outcomes varied in LSS RCTs. It is likely to be a solution to develop a scientific, standard CM pattern scale or a more specific outcome to evaluate the effect of patterns.³⁵

Recent SLR of outcomes reporting in RCTs of LSS has suggested that among 29 trials, function and pain were the most common outcomes, followed by AEs.¹⁰ The results supported the results of our study from SLR and consensus-COS though differences were identified in the trials with comparisons among Western medicine (eg, surgery, physical therapy, medication), as well as the trials identified from six SLRs from Cochrane Central Register of Controlled Trials database and PubMed during 2016 and 2021. Furthermore, function, pain, HRQoL and AE are reported as vital outcomes for LSS in Cochrane SLRs. If LSS was considered specific LBP, several studies consistently recommended pain, function and HRQoL as core outcome domains for LBP.^{24 36–38} Furthermore, additional core domains may be examined alongside the above outcomes to capture condition-specific characteristics.

Strengths and limitations

Strengths of this study include a China national representation of LSS patient and physician stakeholders participating in the consensus meeting, surveys and candidate outcome generation. We followed rigorous research methods and had nearly equal representation of patients and physicians at each step of the process. The response rates were 100% from two rounds in Delphi, avoiding attrition bias. The participants were sampled following duration and socioeconomic status, disease severity and LSS manifestations. This ensured that we captured broad content early in the process of data collection and obtained domains that are generalisable to LSS people.

However, this study also had limitations. First, some of the experts participated in the consensus meeting via WeChat conference instead of face to face due to the COVID-19 pandemic. This may have led to insufficient discussion and affected the consensus results. However, we ensured that every participant had sufficient time for making statements and voting. Each electronic voting was confirmed by reminder before submission. Second, the

number of patients who participated in Delphi rounds and consensus meetings was relatively small. Thus, the importance of certain areas may be underestimated from their perspectives. It is worth mentioning that the goal of this study was to develop a core set of outcomes to be included in all clinical trials, instead of a set of outcome areas important to all stakeholders. Third, participants were not asked to assign relative priority to any domain, whereas all outcome domains that met the consensus threshold of 80% for consensus, should be considered with equal importance.

Implication for clinical practices and research

This study primarily aimed to collect core outcome measures for use in reliable prospective studies related to LSS patients. This COS might potentially be incorporated into LSS registries and used as a reference for data collecting in clinical practice as a list of significant outcomes to monitor during any therapy. When the COS's external validation would be confirmed, the findings can be extrapolated to an adequate population. Next, the psychometric features of each core set domain's outcome measure will be assessed, and a core set of outcome measures that is sufficient and not redundant will be selected.

CONCLUSION

The COS for CM in LSS was first established. Pain and discomfort, HRQoL, lumbar function, ADL, walking function, PGA, AE and CM-specific outcomes should be measured and reported in all future research trials that have evaluated CM for the treatment of LSS to increase consistency in the report of the result. The COS can facilitate the synthesis of the evidence relating to LSS patient-associated outcomes and support overall field development and research.

Author affiliations

¹Traditional Chinese Medicine Department, Xuanwu Hospital Capital Medical University, Beijing, China

²First Clinical College, Beijing University of Chinese Medicine Affiliated Dongzhimen Hospital, Beijing, China

³Health Management Department, Aerospace Central Hospital, Beijing, China

⁴Tuina and Pain Management Department, Beijing University of Chinese Medicine Affiliated Dongzhimen Hospital, Beijing, China

Acknowledgements We would like to acknowledge all the people who participated in the study. We thank all organisations that funded our research.

Contributors Y-NS, X-YW and C-HY all contributed to the study's conception and design. Y-NS, Z-WW, YA and Y-JZ conducted the literature review, interview and had primary responsibility in writing this article. Y-NS, Y-JZ and C-HY conducted survey and consensus meeting, performed analysis and interpreted them. X-YW and C-HY critically revised the manuscript. All authors have read and approved the final manuscript. C-HY is responsible for the overall content as guarantor.

Funding This review was funded by the National Natural Science Foundation of China (No. 81803956); Capital Health Development Research Project (No. 2020-4-4195); Seed Funding of Golden Bridge Project of Beijing Municipal Science and Technology Commission (No. ZZ21053); and Capital Medical University Scientific Research Foundation (Natural Science PYZ22052).

Disclaimer These funding sources had no role in the design of the study, data collection and analysis, or preparation of the manuscript.

Competing interests None declared.

Patient and public involvement Patients and/or the public were involved in the design, or conduct, or reporting, or dissemination plans of this research. Refer to the Methods section for further details.

Patient consent for publication Not applicable.

Ethics approval This study involves human participants and was approved by the ethics committee of Dongzhimen hospital (DZMEC-KY-2020-60). Participants gave informed consent to participate in the study before taking part.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available on reasonable request.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

ORCID iD

Chang-He Yu <http://orcid.org/0000-0002-2764-6229>

REFERENCES

- Jensen RK, Lauridsen HH, Andresen ADK, *et al.* Diagnostic screening for lumbar spinal stenosis. *CLIN EPIDEMIOL* 2020;12:891–905.
- Ishimoto Y, Yoshimura N, Muraki S, *et al.* Associations between radiographic lumbar spinal stenosis and clinical symptoms in the general population: the Wakayama spine study. *Osteoarthritis and Cartilage* 2013;21:783–8.
- Ravindra VM, Senglaub SS, Rattani A, *et al.* Degenerative lumbar spine disease: estimating global incidence and worldwide volume. *Global Spine J* 2018;8:784–94.
- Quality Agency For Healthcare Research. *Overview of the National (nationwide) inpatient sample (NIS). health care cost and utilization project (H-CUP)*. 2022.
- Katz JN, Zimmerman ZE, Mass H, *et al.* Diagnosis and management of lumbar spinal stenosis: A review. *JAMA* 2022;327:1688–99.
- Ammendolia C, Hofkirchner C, Plener J, *et al.* Non-operative treatment for lumbar spinal stenosis with Neurogenic Claudication: an updated systematic review. *BMJ OPEN* 2022;12:e057724.
- Bussi res A, Cancelliere C, Ammendolia C, *et al.* Non-surgical interventions for lumbar spinal stenosis leading to Neurogenic Claudication: A clinical practice guideline. *J Pain* 2021;22:1015–39.
- Chen X, Zheng Z, Lin J. Clinical effectiveness of conservative treatments on lumbar spinal stenosis: A network meta-analysis. *FRONT PHARMACOL* 2022;13:859296.
- Jacobi S, Beynon A, Dombrowski SU, *et al.* Effectiveness of conservative Nonpharmacologic therapies for pain, disability, physical capacity, and physical activity behavior in patients with degenerative lumbar spinal stenosis: A systematic review and meta-analysis. *Arch Phys Med Rehabil* 2021;102:2247–60.
- de Luca K, Anderson D, Dutt A, *et al.* Outcome domain and measurement instrument reporting in randomized controlled trials of interventions for lumbar spinal stenosis: A systematic review. *J Orthop Sports Phys Ther* 2022;52:446–A2.
- Dai X, Zi M, Yang Q, *et al.* Analysis and reflection on the literature status of TCM clinical research outcome indicators. *J Tradit Chin Med* 2022;12:1122–7.
- Kirkham JJ, Dwan KM, Altman DG, *et al.* The impact of outcome reporting bias in randomised controlled trials on a cohort of systematic reviews. *BMJ* 2010;340:bmj.c365.
- Tunis SR, Clarke M, Gorst SL, *et al.* Improving the relevance and consistency of outcomes in comparative effectiveness research. *J Comp Eff Res* 2016;5:193–205.

- 14 Kirkham JJ, Gargon E, Clarke M, *et al.* Can a core outcome set improve the quality of systematic Reviews?--A survey of the Coordinating editors of Cochrane review groups. *TRIALS* 2013;14:21.
- 15 Ollivere BJ, Marson BA, Haddad FS. Getting the right answer: core outcome SETS in Orthopaedics. *BONE JOINT J* 2019;101-B:233–5.
- 16 Kirkham JJ, Gorst S, Altman DG, *et al.* Core outcome set-standards for reporting: the COS-STAR statement. *PLoS Med* 2016;13:e1002148.
- 17 Kirkham JJ, Davis K, Altman DG, *et al.* Core outcome set-standards for development: the COS-STAD recommendations. *PLoS Med* 2017;14:e1002447.
- 18 Sun Y, An Y, Zhou Y-J, *et al.* Non-pharmaceutical Chinese medical therapies for degenerative lumbar spinal stenosis: A systematic review and meta-analysis of randomized controlled trials. *Complementary Therapies in Medicine* 2023;74:102949.
- 19 Williamson PR, Altman DG, Bagley H, *et al.* The COMET Handbook: version 1.0. *Trials* 2017;18(Suppl 3):280.
- 20 Dodd S, Clarke M, Becker L, *et al.* A Taxonomy has been developed for outcomes in medical research to help improve knowledge discovery. *J Clin Epidemiol* 2018;96:84–92.
- 21 Guyatt GH, Oxman AD, Kunz R, *et al.* GRADE guidelines: 2. framing the question and deciding on important outcomes. *J CLIN EPIDEMIOL* 2011;64:395–400.
- 22 Marson BA, Manning JC, James M, *et al.* Development of the CORE-kids core set of outcome domains for studies of childhood limb fractures. *The BONE & JOINT Journal* 2021;103-B:1821–30.
- 23 McMillan SS, King M, Tully MP. How to use the nominal group and Delphi techniques. *Int J Clin Pharm* 2016;38:655–62.
- 24 Chiarotto A, Deyo RA, Terwee CB, *et al.* Core outcome domains for clinical trials in non-specific low back pain. *EUR SPINE J* 2015;24:2097.
- 25 Edemekong PF. *BDSS: Activities of Daily Living*. StatPearls Publishing LLC, 2019.
- 26 The Whoqol Group. The world health organization quality of life assessment (WHOQOL): development and general Psychometric properties. *Social Science & Medicine* 1998;46:1569–85.
- 27 Orbai A-M, de Wit M, Mease P, *et al.* International patient and physician consensus on a Psoriatic arthritis core outcome set for clinical trials. *ANN RHEUM DIS* 2017;76:673–80.
- 28 Qiu R-J, Li M, Hu J-Y, *et al.* Methods for development of a core outcome set for clinical trials integrating traditional Chinese medicine and Western medicine. *J Integr Med* 2021;19:389–94.
- 29 Raja SN, Carr DB, Cohen M, *et al.* The revised International Association for the study of pain definition of pain: concepts, challenges, and compromises. *PAIN* 2020;161:1976–82.
- 30 Singh JA, Yang S, Strand V, *et al.* Validation of pain and patient global scales in chronic gout: data from two randomised controlled trials. *ANN RHEUM DIS* 2011;70:1277–81.
- 31 U. S. Food And Drug Administration. *IND application reporting: safety reports*. 2022.
- 32 Nikiphorou E, Radner H, Chatzidionysiou K, *et al.* Patient global assessment in measuring disease activity in rheumatoid arthritis: a review of the literature. *ARTHRITIS RES THER* 2016;18:251.
- 33 Zhang L, Zhang J, Chen J, *et al.* Shang H: clinical research of traditional Chinese medicine needs to develop its own system of core outcome SETS. *Evidence-Based Complementary and Alternative Medicine* 2013;2013:1–4.
- 34 World Health Organization. *WHO International Standard Terminologies on Traditional Medicine in the Western Pacific Region*. Geneva, Switzerland, 2007.
- 35 Li G, Han R, Lin M, *et al.* Developing a core outcome set for clinical trials of Chinese medicine for hyperlipidemia. *FRONT PHARMACOL* 2022;13:847101.
- 36 Chapman JR, Norvell DC, Hermsmeyer JT, *et al.* Evaluating common outcomes for measuring treatment success for chronic low back pain. *Spine* 2011;36(21 Suppl):S54–68.
- 37 Clement RC, Welander A, Stowell C, *et al.* A proposed set of Metrics for standardized outcome reporting in the management of low back pain. *ACTA ORTHOP* 2015;86:523–33.
- 38 Deyo RA, Battie M, Beurskens AJ, *et al.* Outcome measures for low back pain research. A proposal for standardized use. *Spine (Phila Pa 1976)* 1998;23:2003–13.