ORIGINAL RESEARCH

# Diagnostic Screening for Lumbar Spinal Stenosis

Rikke Krüger Jensen [1,2]
Henrik Hein Lauridsen [1]
Andreas Duch Kiilerich Andresen [3]
Rune Mygind Mieritz [4]
Berit Schiøttz-Christensen [5,6]
Werner Vach [7,8]

[1]Department of Sports Science and Clinical Biomechanics, University of Southern Denmark, Odense, Denmark; [2]Nordic Institute of Chiropractic and Clinical Biomechanics, Odense, Denmark; [3]Spine Surgery and Research, Spine Centre of Southern Denmark, University Hospital Lillebaelt, Middelfart, Denmark; [4]Department of Neurosurgery, University Hospital Odense, Odense, Denmark; [5]Spine Centre of Southern Denmark, University Hospital Lillebaelt, Middelfart, Denmark; [6]Department of Regional Health Research, University of Southern Denmark, Odense, Denmark; [7]Department of Orthopaedics and Traumatology, University Hospital Basel, Basel, Switzerland; [8]Basel Academy, Basel, Switzerland

**Purpose:** To develop a self-administered diagnostic screening questionnaire for lumbar spinal stenosis (LSS) consisting of items with high content validity and to investigate the diagnostic value of the questionnaire and the items.

**Patients and Methods:** A self-reported diagnostic LSS screening questionnaire was developed based on items from the existing literature describing key symptoms of LSS. The screening questionnaire (index test) was to be tested in a cohort of patients with persistent lumbar and/or leg pain recruited from a Danish publicly funded outpatient secondary care spine clinic with clinicians performing the reference test. However, to avoid unnecessary collection of data if the screening questionnaire proved to be of limited value, a case–control design was incorporated into the cohort design including an interim analysis. Additional cases for the case–control study were recruited at two Danish publicly funded spine surgery departments. Prevalence, sensitivity, specificity and diagnostic odds ratio (OR) were calculated for each individual item, and AUC (area under the curve) was calculated to examine the performance of the full questionnaire.

**Results:** A 13-item Danish questionnaire was developed and tested in 153 cases and 230 controls. The interim analysis was not in favour of continuing the cohort study, and therefore, only results from the case–control study are reported. There was a positive association for all items except the presence of back pain. However, the association was only moderate with ORs up to 3.3. When testing the performance of the whole questionnaire, an AUC of 0.72 was reached with a specificity of 20% for a fixed sensitivity of 95%.

**Conclusion:** The items were associated with LSS and therefore have some potential to identify LSS patients. However, the association was not strong enough to provide sufficient accuracy for a diagnostic tool. Additional dimensions of symptoms of LSS need identification to obtain a reliable questionnaire for screening purposes.

**Keywords:** lumbar spinal stenosis, neurogenic claudication, diagnostic screening, questionnaire

## Introduction

Neurogenic claudication is the main symptom of lumbar spinal stenosis (LSS). It is described as lower limb pain and neurological symptoms increased by walking and with symptom relieve when sitting or flexing the spine.[1] Additional symptoms can include numbness, fatigue, impaired balance, muscle weakness and back pain.[2] LSS refers to spinal osteoarthritis that narrows the central spinal canal, the lateral recess or the intervertebral foramen. The decrease in the total area of the spinal canal can lead to compression of the associated neurovascular structures and inflammation of soft tissue structures which can cause neurogenic claudication.[3] The anatomical changes are seen on Magnetic Resonance Imaging (MRI) which is often used in the diagnosis of LSS, as it can provide information on the presence and extent of

Correspondence: Rikke Krüger Jensen
Department of Sports Science and Clinical Biomechanics, University of Southern Denmark, Campusvej 55, Odense M 5230, Denmark
Tel +45 40 41 97 73
Email rikkekruger@nikkb.dk

degenerative spinal disease.[4] However, radiologically verified narrowing of the spinal canal is often found in asymptomatic people, ie the association between MRI findings and clinical symptoms is weak.[5,6]

LSS seems to be a highly prevalent condition although the precise prevalence is unknown. In a study estimating the number of people over 60 years of age with LSS in the general population, the prevalence ranged from 18–49% depending on the diagnostic criteria used.[7] The prevalence of LSS increases with age due to the degenerative nature of the condition.[8] With a growing senior population in most western countries, the number of people with pain and disability due to LSS will expand and simple ways to detect patients as early as possible become more and more necessary to keep the impact on patients, society and health care costs at a minimum. One of the challenges in establishing the prevalence is the lack of a valid and reliable gold standard for LSS.

In 2007, Konno et al[9] developed a 10-question diagnostic support tool (SSHQ) to distinguish LSS from other specific diagnoses of leg pain by means of a self-reported history questionnaire. In a more recent study from 2016, Aizawa et al[10] developed a 15-question diagnostic support tool designed to distinguish between LSS or lumbar disc herniation in patients with leg symptoms. The questionnaire was reported to perform well on this task with sensitivity and specificity for LSS calculated to be 93% and 85% and an area under the curve (AUC) of 0.95 (95% CI 0.93 to 0.97). To our knowledge, no diagnostic screening questionnaire designed to identify LSS patients in a general clinical population of patients with chronic low back pain and/or leg pain has been developed.

The aim of this project was to develop and test a self-administered diagnostic screening questionnaire for patients with LSS. The specific objectives were 1) to define a set of items characterized by high content validity and applicability in a self-administered questionnaire, 2) to investigate the diagnostic value of these items when included in a diagnostic screening questionnaire.

## Methods
This study was reported according to Standards for the Reporting of Diagnostic Accuracy Studies (STARD).[11,12] The protocol was registered on ClinicalTrials.gov (https://clinicaltrials.gov/show/NCT03910335).

## Development of the Screening Questionnaire
The screening questionnaire was developed based on a literature search and qualitative semi-structured interviews.

### Literature Search
Screening questionnaires including symptoms of LSS have already been developed[10,13] and an International Delphi Study has reached a consensus on clinical diagnosis of LSS.[2] Therefore, we did not repeat the process of identifying relevant history questions of LSS. A literature search was performed to identify papers reporting on diagnostic criteria of LSS and a broad spectrum of existing diagnostic screening questions was identified.

### Item Development
The total number of existing diagnostic questions from the identified literature was reduced by merging questions covering the same topic, removing questions with direct overlap and excluding questions not directly associated with LSS symptoms or requiring a physical examination. The resulting list of pre-final screening questions was translated from English to Danish.

### Pilot Test
Face validity of the questionnaire was assessed by conducting semi-structured interviews with LSS patients. Patients assessed the final questions and they were interviewed about their understanding of the phrasing, the appropriateness of the questions in relation to LSS symptoms and if symptoms were covered sufficiently. Patients participating in the interviews were recruited by the examining clinicians at an outpatient public spine clinic, and one of the authors (RKJ) conducted the interviews. Interview duration was on average 30 minutes and was conducted person-to-person. After interviewing four patients no new answers emerged and the interviews were discontinued. Based on the feedback from this process, a final version of the screening questionnaire was developed.

## Evaluation of the Screening Questionnaire
### Design
The screening questionnaire was planned to be evaluated by designing a prospective cohort study in a relevant clinical population. However, due to an expected low prevalence of LSS (approx. 10%), a large sample size was required to recruit a sufficient number of cases. We regarded it as unethical and inconvenient to force patients to answer questionnaires and

medical staff to collect data on hundreds of patients if the screening questionnaire proved to be of limited value. Therefore, a "proof of concept" study with a case–control design was built into the cohort study based on sampling additional cases. This allows for a first judgement of the usefulness of the items with a rather small sample, even if we overestimate the accuracy due to adding more extreme cases than to be expected within the cohort. In case of the failure of such a proof, the cohort study could be terminated. The study design is illustrated in Figure 1.

## Study Setting
The cohort study was performed at a Danish publicly funded outpatient secondary care spine clinic ("Spine Centre of Southern Denmark, medical department"). Patients with persistent spinal pain or suspicion of pain related to the spine are referred to the Spine Centre from general physicians and chiropractors in primary care if they have received relevant non-surgical treatment without achieving the expected effect. Patients who visit the medical department fill out a standardised evaluation form in a clinical registry (SpineData).[14] SpineData is an internet-based system that captures patient data electronically at the point of clinical contact. The patient fills out the evaluation form either the day before their visit to the medical department or in the waiting room before the first contact with a clinician.

For the case–control study, patients with LSS (additional cases) were recruited from the two spine surgery



**Figure 1** Illustration of study design. The arrow (dark green) illustrates the cohort study conducted at the medical department and the area with the dotted line illustrates the case–control study build into the cohort study. The light green area illustrates the additional cases collected at surgical departments for the purpose of the case–control study only.

departments in the Region of Southern Denmark. One was located at the Spine Centre of Southern Denmark, and the other was located at the department of neurosurgery at Odense University Hospital (OUH).

## Inclusion and Exclusion Criteria
We used three inclusion criteria to define the target population:

- Pain in the lumbar spine and/or leg(s) as primary complaint
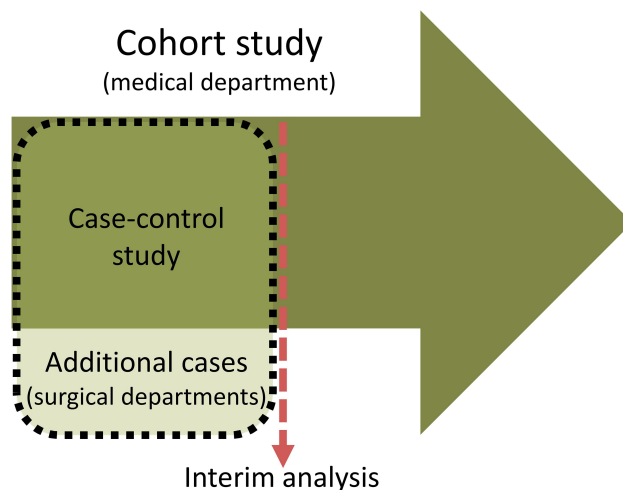- Able to read and write Danish
- Age above 50

Due to the degenerative nature of LSS, the prevalence increases with age and is not often seen in patients under the age of 50[7] which is why an age restriction was chosen.

## Patient Eligibility and Data Collection
### Medical Department
All patients were screened for potential inclusion, based on the data registered in the standard electronic questionnaire. From this data we were able to check the age, the ability to read and write Danish and the primary complaint which the patient indicated by choosing the primary area of pain in an electronic pain drawing. Any patient found eligible for participation in the study, was automatically referred to the LSS screening questionnaire (index-test). Written consent for participation was collected electronically.

A group of volunteering clinicians participated in the data collection. They covered approximately 20% to 25% of the workload of the medical department and had from 3 to 20 years of experience in diagnosing patients with leg and/or back pain. The clinicians were instructed on how to perform the reference test. Patients referred to the medical department were randomly assigned to clinicians' consultation schedule and were only included in the study if they were examined by one of the selected clinicians. Therefore, not all patients who filled out an index test were included in the study population. However, the consecutive referral of patients from primary care to the medical department at the Spine Centre as well as the standard procedure for allocating patients to clinicians' consultation schedule was not changed during the study period. If a patient fulfilled the inclusion criteria, an automatic pop-up message would appear on the clinician's screen as a reminder to fill out the reference test when examining the patient. The results of the index test were not available to the clinicians performing the reference test.

After the clinical contact with the patient, the clinician filled out a questionnaire with a predefined list of clinical symptoms including a diagnosis using a module of the electronic database "SpineData" developed for the purpose of this study.

## Surgical Departments

The two surgical departments were able to offer participation in the study to patients if they were scheduled for decompression surgery due to LSS. Cases recruited from the surgical departments filled out a paper version of the questionnaire and gave written consent for participation. A project nurse at each of the surgical departments distributed and collected the questionnaires and ensured that patients filled them out independently of hospital staff.

## Patient Categorisation

**Reference Standard.** There is no available gold standard for diagnosing LSS and the diagnosis is therefore based on expert opinion. Patients were categorised as LSS-patients by the clinician if they had clinical symptoms of LSS (neurogenic claudication) based on diagnostic criteria in line with the recent suggestions of clinical diagnosis of LSS.[2,15,16] The diagnosis of neurogenic claudication is a clinical assessment primarily based on key items of the case history and ruling out the differential diagnosis by a clinical examination.

Key symptoms for neurogenic claudication included i) pain, motor or sensory disturbance in one or both legs or buttock(s) while walking and standing, ii) forward flexion of the spine relieves symptoms and, iii) symptom relieve when sitting. In case a clinical examination indicates i) vascular claudication, ii) cox arthrosis, iii) radiculopathy due to disc herniation, iv) greater trochanteric pain syndrome or v) peripheral neuropathy, LSS is typically ruled out.

No data on the reference standard were collected at the surgical departments, as it was implicit that patients awaiting surgery for LSS would have clinical symptoms of LSS.

**Case History.** In a clinical consultation, the case history is obtained by a dialogue between patient and clinician, with the clinician asking open questions about symptoms, eg "Can you describe your symptoms?", "Are there factors or activities that will worsen your symptoms?", "Are there situations or positions that relieve your symptoms?", etc. If the patient does not describe a classic picture of neurogenic claudication such as relieving symptoms by leaning forward or flexing the spine, the clinician will often ask more specific questions about the movement but without providing the patient with the "right" answer, eg "Does it change your symptoms when you bend forward?", "Are the symptoms worse, better or unchanged when you bend forward?". Also, the clinician would show the patient the movement or ask the patient to flex the spine and to notice any change in the symptoms when doing so.

**Clinical Examination.** Patients with neurogenic claudication will often have normal examination findings. Sometimes extension will provoke leg symptoms and if the symptoms are severe, there can be neurological deficits such as sensory disturbance or reduced muscle strength in the legs. The clinical examination is also used to rule out differential diagnosis. Vascular claudication can mimic the symptoms and are considered in the case history (relieve of symptoms with discontinuation of walking without sitting or leaning forward, worsening of symptoms when biking) but also by examining pulses in legs and feet. Other potential differential diagnoses such as hip arthritis and greater trochanteric pain syndrome were tested by specific clinical examination tests. In addition, peripheral neuropathy was evaluated based on the case history and examination findings, and additional neurophysiological tests were obtained if necessary.

## Variables of Interest
### Data Collected from the Patient
The following data were extracted from SpineData: Date, Index test, age (in 5-year intervals), sex, physical function measured by Oswestry Disability Index (ODI),[17,18] back- and leg pain measured on an 11-point numerical rating scale (NRS),[19] and level of physical activity measured by means of the question "What is the intensity of the physical activity in your recreational activities?" with four options (1: "I normally sit and read, watch television, go to the movies, or spend my leisure time with quieter activity", 2: "I walk, take short trips on the bicycle or do other kinds of physical activities at least 4 hours a week", 3: "I am an active sportsperson, I run, swim, play tennis/ badminton at least 3 hours a week. If I do not participate in sports, I often do heavy gardening or other demanding leisure time activities, so I belong to this group too", or 4: "I participate in competition sports or swim, bike, or run long distances several times a week"). Also, information on the work situation was collected with the following question: "What is your present work situation?" (1: "Regular job – fulltime or part time", 2: "Subsidised

job", 3: "Studying", 4: "Undertaking work rehabilitation", 5: "Unemployed", 6: "Receiving disability pension", 7: "Receiving retirement pension", 8: "Housewife or house-husband", 9: "Other").

### Data Collected from the Clinician

Following the examination, the clinician provided clinical information and a diagnosis by answering "yes" or "no" to the questions described below.

Clinical information: "Does the patient have i) low back pain, ii) pain or sensory disturbance in one leg, iii) pain or sensory disturbance in both legs?"

Diagnosis: "Does the patient have symptoms of neurogenic claudication as described in the reference standard?". In addition, the clinician provided information on the level of the diagnostic certainty measured in percentages by marking a visual analogue scale from 0% to 100% ("How sure are you of the diagnosis?").

MRI: As neurogenic claudication is a clinical symptom, the diagnosis of LSS did not require confirmation by MRI. However, although MRI is not a gold standard for LSS (due to a high number of asymptomatic patients with narrowing of the spinal canal[5,6]), it can influence the diagnostic certainty by confirming the clinical diagnosis, and information on MRI was therefore collected by asking: 1) "Has an MRI of the lumbar spine been performed?" and 2) "Is the clinical diagnosis confirmed by the MRI?".

### Analysis

The plan for the statistical analysis was developed after the development of the questionnaire was finished and documented in the study protocol (ClinicalTrials.gov Identifier: NCT03910335). According to this plan, the primary analysis was the application of a logistic regression model combining all items in a linear manner (with age as a continuous variable) and the determination of the corresponding AUC by a 10-fold cross-validation. In addition, the specificity corresponding to a cut-point with 95% sensitivity was determined. The continuation of the cohort study would be regarded as justified if we could reach a specificity of at least 68%. With this, we could expect a positive predictive value above 25% and a frequency of positive test results of 39% with a prevalence of 10% in the cohort setting.

The study protocol included a plan to prepare a publication about the negative results from the case–control study, which could include additional analyses planned for the cohort study, in case it was decided not to continue the cohort study.

We decided to include the following analyses planned for the cohort study:

- A description of the diagnostic value of each binary item by prevalence, sensitivity, specificity and the diagnostic odds ratio. All conditional items will be reported only in the subpopulation where the item can be filled in. Age will be dichotomized at the median.
- A description of the diagnostic value of each binary item beyond age by the diagnostic odds ratio adjusted for age.
- A description of the independent value of each binary item by reporting the odds ratio from a multivariate logistic model with adjustment for all other items (full model). Since the condition (item 3) triggering the conditional items (item 4–9) is included in the model, all items can be included. For this analysis the conditional items are recoded as −0.5 and 0.5 within patients with leg pain in order to ensure that the effect of the conditioning item 3 refers to the difference between an average patient with leg pain and a patient without leg pain.

The study protocol also included a plan to describe the distribution of additional variables in LSS and non-LSS patients in tabular form.

In addition, we decided post hoc to perform a factor analysis with varimax rotation to check the dimensionality of the constructs we measured implicitly with our items and whether all items relate to these constructs. We further decided to report not only the specificity related to a cut-point referring to 95% sensitivity but also the specificity of cut-points related to a sensitivity of 90% and 85%. In addition to ordinary logistic regression, we also applied the Lasso technique[20] in order to obtain a suggestion for a parsimonious model. To obtain a better idea about the overall value of our items, we performed a comparison based on the 10 items of the ODI. To investigate the sensitivity of the definition of the case–control status, we presented an analysis including only patients for whom the clinicians at the Spine Centre were confident about their decision and an analysis with the case–control status based on the results of the MRI for patients recruited at the medical department.

### Sample Size

The intention was to include a minimum of 100 cases and 100 controls in the case–control study. One-hundred cases would allow an estimate of a sensitivity of 0.95 with a

standard error of 0.022. One-hundred controls would allow an estimate of a specificity of 0.7 with a standard error of 0.046. The sample size estimation of the cohort study was planned to be based on the results of the case–control study.

## Results

### Development of the Screening Questionnaire

Five papers reporting on diagnostic criteria or screening questions for LSS were identified[2,10,13,15,16] with a total of 41 general questions or the presence of symptoms. After merging overlapping questions and removing symptoms related to clinical examinations, 13 questions remained (Table 1). The process of identifying the 13 questions is shown in Additional file 1.

It was decided to include an entrance question about the presence of leg symptoms as patients without leg symptoms would otherwise have to answer irrelevant questions which we considered inconvenient for patients.

Four patients participated in individual semi-structured interviews. All patients were able to describe the meaning of the questions in their own words. One did not find question 7c relevant as he did not ride a bike. All other questions were found relevant and there were no additional suggestions for questions. No new items were developed after feedback from the semi-structured interviews. One patient noticed that the yes/no categories were forcing an answer and suggested the extra categories of "sometimes" and "not applicable". We decided against adding these response options as it would complicate the questionnaire.

### Evaluation of the Screening Questionnaire
#### Patient Population

The recruitment at the medical department started on February 19th, 2019. Due to some technical problems with SpineData and its use by the participating clinicians during the first months, it was not possible to include all eligible patients. Those left out were from the control

**Table 1** Final Screening Questionnaire (Index Test)

| | Yes | No |
|---|---|---|
| For each question, please tick 'yes' or 'no' to indicate whether you have experienced the symptom within the last month. We ask about both pain and numbness. By numbness we mean pins-and-needles feeling, tingling or numbness. | | |
| 1.   How old are you? _____years | | |
| 2.   Do you have low back pain? | ☐ | ☐ |
| 3.   Do you sometimes feel pain or numbness in one/both legs or buttocks? | ☐ | ☐ |
| **(If you answered no in question 3 you have completed the questionnaire)** | | |
| The next question is about where you feel the symptoms i.e. pain or numbness. | | |
| 4.   Do you feel pain or numbness in **both** legs or buttocks? | ☐ | ☐ |
| 5.   Do you feel numbness in the soles of **both** feet? | ☐ | ☐ |
| The next questions are about what increases your pain or numbness. | | |
| 6.   Does your pain or numbness in one/both legs or buttocks **increase** when you: | | |
|       a)    walk? | ☐ | ☐ |
|       b)    stand for a while? | ☐ | ☐ |
| The next question is about what decreases your pain or numbness. | | |
| 7.   Does your pain or numbness in one/both legs or buttocks **decrease** when you: | | |
|       a)    bend forward? | ☐ | ☐ |
|       b)    sit? | ☐ | ☐ |
|       c)    bike? | ☐ | ☐ |
|       d)    use a shopping cart? | ☐ | ☐ |
| The next questions are about what happens when you walk. | | |
| 8.   Do you bend forward while walking? | ☐ | ☐ |
| 9.   Do you feel weakness in the legs while walking? | ☐ | ☐ |
| **Thank you!** | | |

**Note:** The questionnaire was originally in Danish and has been informally translated for the purpose of this paper.

group in particular. From April 8th, 2019 and onwards, the system was working properly. The surgical departments started recruitment on January 22nd, 2019.

Between February 19th, 2019 and October 23rd, 2019, 2,436 patients gave consent to participate, 2,427 answered the entrance question (item 3) of which 87% were positive on leg symptoms. A total of 352 (17%) patients were examined by a project clinician (reference test) and were included in the cohort study. This is lower than the rate we should expect according to the workload covered by the clinicians. However, in the period from April 8th to June 28th, we included 20% of all patients answering item 3, corresponding to the expected percentage. During this period, the prevalence of LSS in the cohort study was 31%.

After including 107 cases and 138 controls, the recruitment from the surgical departments was stopped. Data were extracted from SpineData on June 28th, 2019, and the planned interim analysis was conducted. The analysis indicated an insufficient accuracy of the questionnaire, and recruitment for the cohort study at the SpineCentre was stopped on October 23rd, 2019. At the time of terminating data collection, 153 cases and 230 controls had been

recruited, and our results are based on this population. Of these, 31 cases were additional cases recruited by surgical departments. Figure 2 displays the flow of the inclusion.

Table 2 presents an overview of the basic patient characteristics as reported by the patients or clinicians. Patients included in the study were similar to those not included ("Patient population"), except for an underrepresentation of unemployed patients and patients receiving a disability pension. However, the differences were small. As expected, cases tended to be older than controls and consequently were more often retired, and they suffered from more severe leg pain and higher disability levels.

## Presence and Handling of Missing Values in Single Items

All patients answered the entry questions on back pain and leg pain. Among the 345 patients confirming the presence of leg pain, 88% answered all the remaining 10 items, and the maximal number of missing items was four. The maximal number of missing values per item was 29 (8.4%) for item 7c and 8 (2.3%) for item 7d. All other items had a maximum of five missing values. We decided to regard all missing values
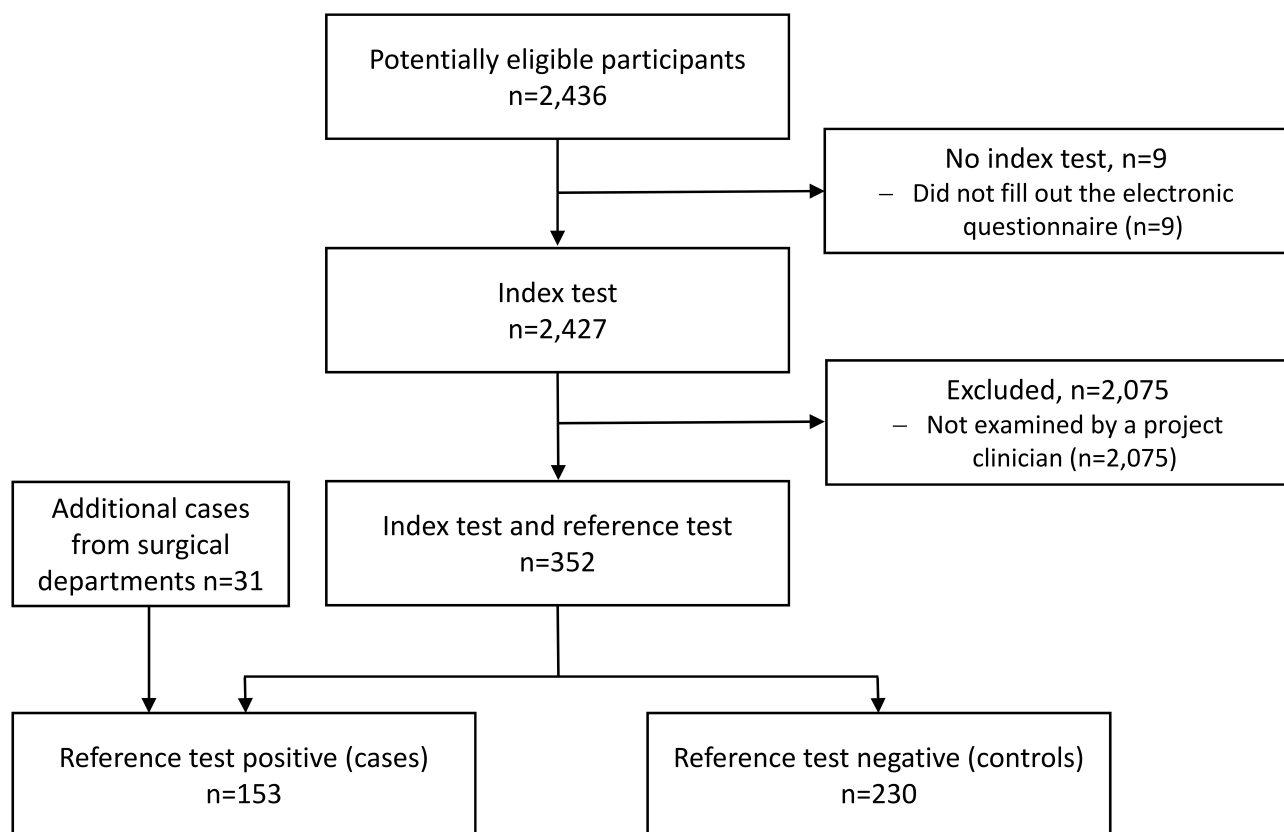


**Figure 2** Flowchart showing the number of participants and reasons for exclusion.

**Table 2** Characteristics of Patients Included from Medical Department (SpineData)

| | Patient Population | | Sample | |
|---|---|---|---|---|
| | Index Test and Reference Test (n=352) | Index Test, No Reference Test (n=2,075) | Cases (n=122) | Controls (n=230) |
| **Patient reported information:** | | | | |
|   Age in years, mean (SD) | 66 (9.9) | 65 (9.4) | 69 (9.8) | 64 (9.6) |
|   Age range in years | 50–89 | 50–91 | 50–89 | 50–88 |
|   Sex (female) | 50.3% | 52.5% | 48.4% | 51.3% |
| **ODI score (0–100 scale)** | | | | |
|   n (<3 missing ODI items) | 348 | 2,041 | 121 | 227 |
|   Mean (SD) | 33.8 (15.8) | 36.2 (16.7) | 36.5 (16.0) | 32.4 (15.6) |
|   p10-p90 | 16–54 | 16–60 | 18–56 | 16–52 |
| **NRS pain intensity (0–10 scale)** | | | | |
|   Current LBP | | | | |
|   n | 346 | 2,054 | 118 | 228 |
|   Mean (SD) | 5.4 (2.6) | 5.4 (2.6) | 5.4 (2.8) | 5.4 (2.5) |
|   p10-p90 | 2–9 | 2–8 | 1–9 | 2–8 |
|   Typical LBP last 14 days | | | | |
|   n | 347 | 2,047 | 119 | 228 |
|   Mean (SD) | 6.2 (2.4) | 6.2 (2.5) | 6.3 (2.7) | 6.2 (2.3) |
|   p10-p90 | 3–9 | 2–9 | 2–9 | 3–9 |
|   Worst LBP last 14 days | | | | |
|   n | 347 | 2,051 | 119 | 228 |
|   Mean (SD) | 7.4 (2.5) | 7.3 (2.6) | 7.4 (2.8) | 7.4 (2.3) |
|   p10-p90 | 4–10 | 3–10 | 2–10 | 4–10 |
|   Current leg pain | | | | |
|   n | 345 | 2,034 | 121 | 224 |
|   Mean (SD) | 4.4 (2.9) | 4.7 (2.9) | 5.0 (2.8) | 4.1 (2.9) |
|   p10-p90 | 0–8 | 0–8 | 1–9 | 0–8 |
|   Typical leg pain last 14 days | | | | |
|   n | 346 | 2,014 | 121 | 225 |
|   Mean (SD) | 5.6 (2.9) | 5.6 (3.0) | 6.6 (2.2) | 5.0 (3.0) |
|   p10-p90 | 1–9 | 0–9 | 4–9 | 0–9 |
|   Worst leg pain last 14 days | | | | |
|   n | 348 | 2,019 | 122 | 226 |
|   Mean (SD) | 6.4 (3.0) | 6.4 (3.1) | 7.5 (2.2) | 5.9 (3.2) |
|   p10–p90 | 1–10 | 0–10 | 5–10 | 0–10 |
| **Present work situation** | | | | |
|   n | 351 | 2, 069 | 122 | 229 |
|   Regular job (full or part time) | 29.1% | 29.6% | 18.0% | 34.9% |
|   Subsidised job | 4.6% | 4.8% | 5.7% | 3.9% |
|   Studying | 0.0% | 0.3% | 0.0% | 0.0% |
|   Current work rehabilitation | 0.0% | 0.9% | 0.0% | 0.0% |
|   Unemployed | 1.7% | 4.0% | 0.0% | 2.6% |
|   Receiving disability pension | 5.7% | 7.9% | 4.1% | 6.6% |
|   Retirement pension | 55.8% | 48.0% | 67.2% | 49.8% |
|   Housewife or househusband | 0.3% | 0.8% | 0.8% | 0.0% |
|   Other | 2.9% | 3.8% | 4.1% | 2.2% |

*(Continued)*

**Table 2** (Continued).

| | Patient Population | | Sample | |
|---|---|---|---|---|
| | Index Test and Reference Test (n=352) | Index Test, No Reference Test (n=2,075) | Cases (n=122) | Controls (n=230) |
| **Intensity of physical activity in recreational activities*** | | | | |
| n | 346 | 2,036 | 119 | 227 |
| - category 1 | 24.0% | 28.6% | 29.4% | 21.1% |
| - category 2 | 59.0% | 54.3% | 58.0% | 59.5% |
| - category 3 | 16.8% | 16.6% | 11.8% | 19.4% |
| - category 4 | 0.3% | 0.4% | 0.8% | 0.0% |
| **Clinician reported information:** | | | | |
| n | 351 | | 121 | 230 |
| Low back pain | 87.2% | | 88.4% | 86.5% |
| n[§] | 314 | | 104 | 210 |
| Unilateral leg symptoms | 53.8% | | 55.8% | 52.9% |
| n[§] | 333 | | 116 | 217 |
| Bilateral leg symptoms | 30.6% | | 47.4% | 21.7% |
| **Certainty of diagnosis (0–100%)** | | | | |
| n | 343 | | 117 | 226 |
| Mean (SD) | 81.4 (23.5) | | 82.6 (18.3) | 80.7 (25.9) |
| p10-p90 | 50–100 | | 51–100 | 45–100 |
| Certainty of diagnosis ≥75% | 75.2% | | 73.5% | 76.1% |
| **MRI** | | | | |
| n | 349 | | 122 | 227 |
| MRI of lumbar spine | 45.3% | | 46.7% | 44.5% |
| n | 156 | | 56 | 100 |
| MRI (if performed) confirmed the clinical diagnosis | 76.3% | | 91.1% | 68.0% |

**Notes:** *Activity category 1: "I normally sit and read, watch television, go to the movies, or spend my leisure time with quieter activity", 2: "I walk, take short trips on the bicycle or do other kinds of physical activities at least 4 hours a week", 3: "I am an active sportsperson, I run, swim, play tennis/badminton at least 3 hours a week. If I do not participate in sports, I often do heavy gardening or other demanding leisure time activities, so I belong to this group too.", or 4: "I participate in competition sports or swim, bike, or run long distances several times a week"; [§]13 clinicians answered "Yes" in both unilateral and bilateral leg pain and therefore excluded.
**Abbreviations:** SD, standard deviation; ODI, Oswestry Disability Index; NRS, numerical rating scale.

in these 10 items in this group as negative answers, representing the absence of the condition asked for.

The ODI was filled out by all patients from the medical department but was not distributed to the patients from the surgical departments. Of these 352 patients, 283 answered all 10 items and 59 answered 9 of the 10 items, the remaining 10 participants answered at least five items. We replaced missing values in single items by a prediction from all other items using an ordinal logistic regression model.

### Diagnostic Value of Single Items
Table 3 depicts the diagnostic value of each single item. Except for back pain, we observed a positive association for all items. However, the associations remain rather

moderate with odds ratios of a magnitude between 1.4 and 3.3. Overall, adjustment for age reduced the odds ratio slightly. For nine items, sensitivity was higher than specificity, and for four of them (2, 3, 6a, 6b) a sensitivity above 80% was reached. Among the four items where specificity was higher than sensitivity, one item (5) reached a specificity above 80%. Figure 3 shows the sensitivity and specificity of each item and illustrates the limited diagnostic accuracy.

### Overall Diagnostic Value
When combing all 13 items investigated in this study into a single score using logistic regression, we were able to reach an AUC of 0.72 with a specificity of 20% when aiming at a sensitivity of 95% meaning that the questionnaire would capture 95% of those with LSS but only single

**Table 3** Diagnostic Value of Single Items in the Questionnaire

| Item | Prevalence | Sensitivity | Specificity | OR | p-value | Adj. OR | Adj. p-value | N |
|------|-----------|-------------|-------------|------|---------|---------|--------------|-----|
| 1* | 0.56 | 0.68 | 0.52 | 2.28 | <0.001 | 0.92 | 0.843 | 383 |
| 2 | 0.90 | 0.88 | 0.09 | 0.67 | 0.238 | 0.81 | 0.557 | 383 |
| 3 | 0.90 | 0.95 | 0.13 | 3.25 | 0.004 | 3.60 | 0.004 | 383 |
| 4 | 0.52 | 0.61 | 0.55 | 1.93 | 0.003 | 1.76 | 0.013 | 345 |
| 5 | 0.23 | 0.32 | 0.82 | 2.16 | 0.003 | 1.80 | 0.027 | 345 |
| 6a | 0.77 | 0.87 | 0.31 | 2.95 | <0.001 | 2.81 | <0.001 | 345 |
| 6b | 0.86 | 0.88 | 0.16 | 1.40 | 0.297 | 1.37 | 0.343 | 345 |
| 7a | 0.40 | 0.53 | 0.70 | 2.66 | <0.001 | 2.59 | <0.001 | 345 |
| 7b | 0.56 | 0.70 | 0.55 | 2.81 | <0.001 | 2.38 | <0.001 | 345 |
| 7c | 0.31 | 0.36 | 0.73 | 1.53 | 0.069 | 1.55 | 0.072 | 345 |
| 7d | 0.54 | 0.68 | 0.56 | 2.71 | <0.001 | 2.49 | <0.001 | 345 |
| 8 | 0.41 | 0.48 | 0.64 | 1.62 | 0.028 | 1.29 | 0.270 | 345 |
| 9 | 0.65 | 0.76 | 0.42 | 2.36 | <0.001 | 2.01 | 0.005 | 345 |

**Note:** *Age was dichotomized at 65 years.
**Abbreviation:** OR, odds ratio.

out 20% of those without LSS. With a lower sensitivity of 85%, the specificity increased to 44.3%. Considering only patients for whom the clinicians were confident about the diagnosis did not influence the AUC. Replacing the clinical reference standard with the MRI results, we observed a decrease of the AUC to 0.64. When using the ODI items as alternative input, we were only able to reach an AUC of 0.68 with specificities of 9.1% and 30.9%, respectively. Details are shown in Table 4.
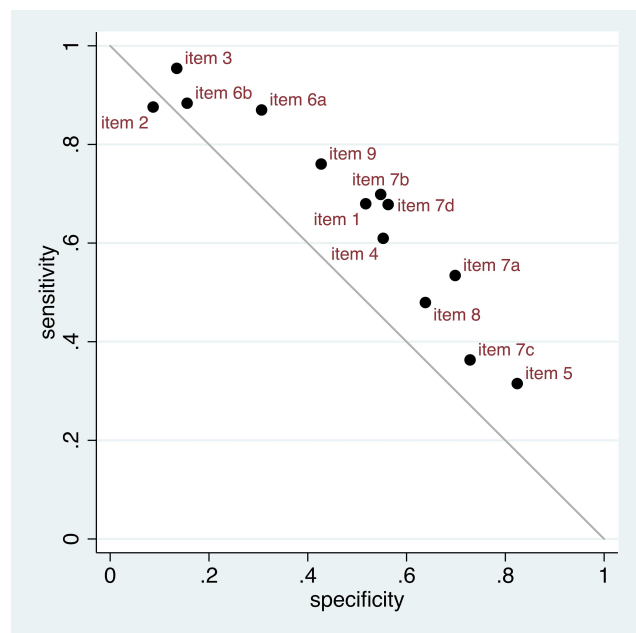


**Figure 3** Sensitivity and specificity for each of the 13 items. Items that are closer to the line have less predictive value.

## Results from Multivariate Analysis

When trying to assess the independent value of the single items in a multiple logistic regression model (Table 5), we observed the odds ratios decrease towards 1 for nearly all items. Only for the items 1, 3, 6a, 7a, 7d, and 9, an estimated OR above 1.6 was observed, and for four of them, we were able to establish an independent value at the 5% significance level. Tibshiarani's Lasso suggested to remove the items 6b, 7c, and 8 in order to arrive at a parsimonious model and to assign the highest weight (OR above 1.6) to items 1, 3, 6a, and 7a.

## Factor Analyses

The results of a factor analysis applied to the 10 additional items in the 345 subjects reporting leg pain suggested that a 2-factor solution could explain nearly all variance. The loading of the items on the two factors are shown in Table 6. We observed that all the items related to "pain relief" load on the first factor, whereas items related to the "presence or increase of symptoms" load on the second factor. Also, the items describing that pain increases while walking load on the second factor, but the "Do you bend forward while walking?" item also loads on the first factor (cross-loading), which may reflect that this item also asks for an action to ease pain. When applying the two corresponding factor regression scores in a logistic regression model controlling for age, both were significantly associated with the case–control status, and we reached an AUC of 0.72 in this subpopulation.

## Discussion

Our study was able to confirm that most of the 13 selected items are indeed useful to distinguish between LSS and non-

**Table 4** Overall Diagnostic Value of the Questionnaire

| Items | Population | Number of | | AUC | Specificity Using a Cut-Point with a Sensitivity of | | |
|---|---|---|---|---|---|---|---|
| | | Cases | Controls | | 95% | 90% | 85% |
| Screening questionnaire items | All patients | 153 | 230 | 0.72 | 20.0% | 28.3% | 44.3% |
| | Diagnostic certainty ≥75% | 122 | 176 | 0.71 | 10.2% | 31.3% | 44.3% |
| | Case–control status according to MRI or additional case | 73 | 114 | 0.64 | 13.7% | 24.7% | 30.1% |
| ODI items | All patients with information on ODI | 122 | 230 | 0.68 | 9.1% | 20.0% | 30.9% |

**Table 5** Diagnostic Odds Ratio Based on a Multiple Logistic Regression Model and on Tibshirani's Lasso

| | | Logistic Regression | | Tbishirani's Lasso |
|---|---|---|---|---|
| | | OR | p-value | OR |
| 1 | Age (dichotomized at 65 years) | 1.84 | 0.013 | 1.63 |
| 2 | Back pain | 0.57 | 0.143 | 0.75 |
| 3 | Leg symptoms | 2.99 | 0.022 | 2.05 |
| 4 | Bilateral leg symptoms | 1.35 | 0.265 | 1.23 |
| 5 | Symptoms in both feet | 1.46 | 0.223 | 1.32 |
| 6a | Symptoms increase when walking | 2.67 | 0.005 | 2.02 |
| 6b | Symptoms increase when standing | 0.65 | 0.262 | 1.00 |
| 7a | Symptoms decrease when bending | 2.00 | 0.016 | 1.70 |
| 7b | Symptoms decrease when sitting | 1.49 | 0.133 | 1.46 |
| 7c | Symptoms decrease when biking | 1.13 | 0.654 | 1.00 |
| 7d | Symptoms decrease when using shopping cart | 1.71 | 0.056 | 1.57 |
| 8 | Bending forward while walking | 0.86 | 0.590 | 1.00 |
| 9 | Heavy legs while walking | 1.67 | 0.071 | 1.43 |

LSS in a general clinical population. However, none of them reached an association above an OR of 3.3, indicating limited usefulness to diagnose LSS. Also, the 13 items together could not provide a diagnostic score with a sufficient association deemed reasonable when using the screening tool in a clinical setting. We only reached an AUC of 0.72 and a specificity of 20% for a fixed sensitivity of 95%. Even accepting a sensitivity of 85%, ie overlooking 15% of the patients with LSS, would still only be able to single out 44% of the patients without LSS.

This somewhat disappointing result cannot be explained by having selected the wrong items. In the multivariate model, several of the items remained significant, demonstrating an independent diagnostic value. Also, many items were selected by Tibshirani's Lasso in an attempt to construct a parsimonious diagnostic model. A factor analysis revealed that the 10

items applicable in patients with leg pain represent two latent factors both related to the presence of LSS and predicting the case–control status as well as the 13 items together. There is also no indication that the results are due to imperfectness of the reference standard. Neither focusing on patients for whom the clinicians were certain about the diagnosis nor the use of MRI as reference standard was able to improve the results in a relevant manner.

In summary, the findings indicate that the 13 items simply perform poorly when distinguishing between LSS and non-LSS in a general clinical population. Some of our items have also been used in previous studies. For example, for the item "Pain/symptoms reduced by bending forward," Konno et al[9] reported an OR of 12 and Aizawa et al[10] an OR of almost 8, compared to 2.7 in our study. The higher ORs observed in these studies

**Table 6** Results of Factor Analysis (Varimax Rotation)

| No. | Question | Factor 1 - Symptom Relieve | pcFactor 2 – Presence or Increase of Symptoms |
|---|---|---|---|
| 4 | Do you feel pain or numbness in both legs or buttocks? | 0.100 | 0.373 |
| 5 | Do you feel numbness in the soles of both feet? | −0.041 | 0.362 |
| 6a | Does your pain or numbness in one or both legs or buttocks increase when you walk? | 0.069 | 0.515 |
| 6b | Does your pain or numbness in one or both legs or buttocks increase when you stand for a while? | 0.163 | 0.339 |
| 7a | Does your pain or numbness in one or both legs or buttocks decrease when you bend forward? | 0.627 | 0.039 |
| 7b | Does your pain or numbness in one or both legs or buttocks decrease when you sit? | 0.397 | 0.265 |
| 7c | Does your pain or numbness in one or both legs or buttocks decrease when you bike? | 0.400 | −0.070 |
| 7d | Does your pain or numbness in one or both legs or buttocks decrease when you use a shopping cart? | 0.609 | 0.086 |
| 8 | Do you bend forward while walking? | 0.281 | 0.324 |
| 9 | Do you feel weakness in the legs while walking? | 0.077 | 0.492 |

illustrate that undoubtedly the item has a value in screening for LSS, but it is just not good enough for screening in a general clinical population.

Diagnostic support tools for screening for LSS have previously been developed by Konno et al[9] who suggested a 10-item questionnaire (SSHQ) designed to distinguish LSS from other specific diagnoses of leg pain and to differentiate subtypes of LSS. The questionnaire was developed in a pre-surgical LSS population and validated in a secondary care population with primary symptoms of pain or numbness in the legs due to either LSS, lumbar disc herniation, diabetic neuropathy or peripheral vascular disease. The questionnaire performed well in the validation cohort with sensitivity and specificity for LSS of 84% and 78% and an AUC of 0.78 (95% CI 0.69–0.87).[9] Also, Aizawa et al[10] developed a 15-item questionnaire to distinguish between LSS and lumbar disc herniation in patients with leg symptoms. This questionnaire also performed well with sensitivity and specificity for LSS calculated to be 93% and 85% and an AUC of 0.95 (95% CI 0.93–0.97).[10] The performance of our screening tool is not in line with these results. This is most likely due to the different purposes of the questionnaires and thereby the choice of population in which the questionnaires are developed and tested. The two questionnaires mentioned above were developed to distinguish between LSS and other specific diagnoses of leg symptoms, whereas our screening questionnaire was developed to distinguish between LSS

and all other causes of leg symptoms including the non-specific causes which are probably the largest group.[21] Therefore, our population was likely more heterogenic and consequently had a large proportion of patients without LSS but with overlapping or similar symptoms.

There seems to be a need for a diagnostic questionnaire to determine the presence of LSS symptoms among a general clinical population. Even though the SSHQ questionnaire was developed to differentiate specific causes of leg symptoms in a secondary care setting, it has also been used to estimate the prevalence of LSS in general populations.[22–24] In our opinion, such usage of a questionnaire not developed or validated in a relevant population is questionable and should be avoided.

There are some limitations to this study. Based on clinical experience from the Spine Centre, it was estimated that the prevalence of LSS would be about 10%. Consequently, recruitment of a very large population to sufficiently power the cohort study was necessary. We, therefore, chose to include a case–control study in the design as not to inconvenience patients and clinicians unnecessarily. However, during the period with no technical system error, the prevalence of LSS was 31%. Therefore, in hindsight, including a case–control study may not have been necessary.

A prevalence of 31% may seem high for a population from secondary care. However, a recently published systematic review[8] of the prevalence of LSS in different

populations reported a pooled prevalence estimate of 29% (95% CI: 22–36%) from nine study samples investigating secondary care populations. The largest of the nine studies (n=125,796) measured LSS in registry data by means of ICD-10 codes whereas eight studies used "Expert opinion" as a case definition. Although eight out of nine studies had a high risk of bias the results are somewhat comparable to our findings, suggesting that our prevalence estimate could be reliable.

As there is no gold-standard diagnostic tool for LSS, the reference test was dependent on the individual clinical assessment. The use of an imperfect reference standard may have introduced bias.[25] We, therefore, introduced a scale measuring the clinician's level of diagnostic certainty enabling us to only include those with high levels of certainty. Also, even though MRI findings are not strongly correlated with clinical symptoms of LSS, it may increase the diagnostic certainty by confirming or rejecting the diagnosis. Finally, cases recruited at the surgical departments were considered to have a higher diagnostic accuracy as they were awaiting surgery for LSS. However, none of the above-mentioned initiatives made substantial changes to the performance of the questionnaire indicating that imperfectness of the reference standard could not explain the poor performance of the screening questionnaire.

LSS is often divided into either central canal stenosis which causes bilateral leg symptoms and recess or foraminal stenosis that causes unilateral leg symptoms and sometimes resembles radiculopathy. The questionnaire was designed to include both types of LSS although item 4 "Do you feel pain or numbness in both legs or buttocks?" and item 5 "Do you feel numbness in the soles of both feet?" aimed at the central canal stenosis. The radicular type would typically not benefit from bending forward or sitting, although there is a large variation in clinical symptoms that needs to be acknowledged.

LSS may often co-exist with other conditions, and although clinicians are trained to look for and identify competing diagnoses, we did not register the occurrence of co-existing disease which could be a limitation to the study. However, the presence of differential diagnoses should have led to a lower level of diagnostic certainty and this would not have affected the analysis including only cases with high levels of diagnostic certainty.

The selection of the questions included in the screening tool was based on questions and symptoms from previously published literature regarding diagnostic criteria for LSS. The selection of the questions included in the screening tool was based on questions and symptoms from previously published literature regarding diagnostic criteria for LSS. These also included the presence of LBP although this symptom is not expected to be specific for LSS and is not part of the description of the disease guidelines.[26,27] Rather than choosing a limited number of questions, we included a broad spectrum excluding only overlapping questions, questions not related to LSS or clinical examination findings. The questionable role of LBP as an indicator is actually corroborated by our study, with LBP being the only item with an OR less than one. Alternatively, we could have chosen to include a Delphi process based on expert opinion to select the most relevant questions. However, three[2,15,16] of the five identified papers were already based on large expert surveys from 2013, 2016 and 2018, respectively, and it is therefore unlikely that redoing the process would have brought about new questions. On the contrary, this study may inform future research focusing on identifying new dimensions to identify LSS patients rather than confirming questions already included. This may be accomplished by qualitative research initiatives based on patient interviews to help develop the questions. The present study was conducted by a team of clinicians, researchers and a statistician, but future research with a focus on identifying new dimensions and development of questions would benefit from including team members with a professional background, eg biomedical linguists, anthropologists and in particular patients. Ideally, patients or patient representatives should be directly involved in both designing and conducting such studies. Also, further development of a screening questionnaire could explore questions ruling out known differential diagnosis. However, it is also possible that identifying patients with LSS in a heterogeneous population of patients with LBP with and without leg pain may not actually be possible by means of a questionnaire.

## Conclusions

The results of this study could confirm that the items selected for the score are associated with LSS and hence have some potential to identify LSS patients. However, the degree of association was insufficient to define a score with acceptable accuracy. Some items demonstrated a high sensitivity to detect LSS patients, but were too frequent in patients without LSS, as indicated by low specificities. Factor analyses suggest that our items cover two relevant constructs in detecting LSS, namely "Symptom relieve" and "Presence or increase

of symptoms" of symptoms. However, there is a need to identify additional dimensions of pain and symptoms in LSS patients or to improve the assessment of these dimensions reflected by our items in order to obtain a reliable instrument for screening purposes.

## Abbreviations

AUC, area under the curve; CI, confidence Interval; LSS, lumbar spinal stenosis; MRI, magnetic resonance imaging; NRS, numerical rating scale; ODI, Oswestry Disability Index.

## Data Sharing Statement

The data supporting the findings of this study are retained at a central repository at the Spine Centre of Southern Denmark, Middelfart, Denmark and restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available. However, data used in the current study can be applied for by contacting the Spine Centre of Southern Denmark.

Additional file 1 provides an overview of the 41 relevant questions extracted from five papers and how they were merged into the final 13 questions.

## Ethics Approval and Informed Consent

The Regional Scientific Ethics Committee for Southern Denmark states on their webpage that this type of study did not require ethical approval.[28] The study conformed to the standards set by the Declaration of Helsinki. The project included only legally competent, consenting patients, and participation was voluntary. All participants were informed about the purpose of the study and written consent was collected. The project was registered at Danish Data Protection Agency [no: 10.180].

## Acknowledgments

The authors would like to acknowledge the clinicians who took part in the data collection at the medical department and the surgical department at the Spine Centre of Southern Denmark and at the department of neurosurgery, Odense University Hospital.

## Author Contributions

RKJ: Study design, data acquisition, drafting the article, interpretation of data. HHL: Study design, statistical analysis, interpretation of data, revision of the article. ADKA: study design, data acquisition, revision of the article. RMM: Data acquisition, revision of the article. BSC: Data acquisition, revision of the article. WV: Study design, drafting the article, statistical analysis, interpretation of data. In addition, all authors contributed to data analysis, drafting or revising the article, have agreed on the journal to which the article will be submitted, gave final approval of the version to be published, and agree to be accountable for all aspects of the work.

## Funding

There is no funding to report.

## Disclosure

The authors declare that they have no competing interests for this work.

## References

1. Porter RW. Spinal stenosis and neurogenic claudication. *Spine*. 1996;21(17):2046–2052. doi:10.1097/00007632-199609010-00024
2. Tomkins-Lane C, Melloh M, Lurie J, et al. Consensus on the clinical diagnosis of lumbar spinal stenosis: results of an international Delphi study. *Spine*. 2016;41(15):1239–1246. doi:10.1097/BRS.000000000 0001476
3. Lurie J, Tomkins-Lane C. Management of lumbar spinal stenosis. *BMJ*. 2016;352:h6234. doi:10.1136/bmj.h6234
4. Malfair D, Beall DP. Imaging the degenerative diseases of the lumbar spine. *Magn Reson Imaging Clin N Am*. 2007;15(2):221–238. doi:10.1016/j.mric.2007.04.001
5. Boden SD, Davis DO, Dina TS, Patronas NJ, Wiesel SW. Abnormal magnetic-resonance scans of the lumbar spine in asymptomatic subjects. A prospective investigation. *J Bone Joint Surg Am*. 1990;72 (3):403–408. doi:10.2106/00004623-199072030-00013
6. Yamada K, Satoh S, Hashizume H, et al. Diffuse idiopathic skeletal hyperostosis is associated with lumbar spinal stenosis requiring surgery. *J Bone Miner Metab*. 2019;37:118–124.
7. Kalichman L, Cole R, Kim DH, et al. Spinal stenosis prevalence and association with symptoms: the Framingham Study. *Spine J*. 2009;9 (7):545–550. doi:10.1016/j.spinee.2009.03.005
8. Jensen RK, Jensen TS, Koes B, Hartvigsen J. Prevalence of lumbar spinal stenosis in general and clinical populations: a systematic review and meta-analysis. *Eur Spine J*. 2020. doi:10.1007/s00586-020-06339-1
9. Konno S, Kikuchi S, Tanaka Y, et al. A diagnostic support tool for lumbar spinal stenosis: a self-administered, self-reported history questionnaire. *BMC Musculoskelet Disord*. 2007;8:102. doi:10.1186/ 1471-2474-8-102
10. Aizawa T, Tanaka Y, Yokoyama T, et al. New diagnostic support tool for patients with leg symptoms caused by lumbar spinal stenosis and lumbar intervertebral disc herniation: a self-administered, self-reported history questionnaire. *J Orthop Sci*. 2016;21(5):579–585. doi:10.1016/j.jos.2016.07.012
11. Bossuyt PM, Reitsma JB, Bruns DE, et al. STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. *BMJ*. 2015;351:h5527. doi:10.1136/bmj.h5527
12. Cohen JF, Korevaar DA, Altman DG, et al. STARD 2015 guidelines for reporting diagnostic accuracy studies: explanation and elaboration. *BMJ Open*. 2016;6(11):e012799. doi:10.1136/bmjopen-2016-012799

13. Konno S, Hayashino Y, Fukuhara S, et al. Development of a clinical diagnosis support tool to identify patients with lumbar spinal stenosis. *Eur Spine J*. 2007;16(11):1951–1957. doi:10.1007/s00586-007-0402-2

14. Kent P, Kongsted A, Jensen TS, Albert HB, Manniche C, Schiøttz-Christensen B. SpineData – a Danish clinical registry of people with chronic back pain. *Clin Epidemiol*. 2015;7:369–380. doi:10.2147/CLEP.S83830

15. Sandella DE, Haig AJ, Tomkins-Lane C, Yamakawa KS. Defining the clinical syndrome of lumbar spinal stenosis: a recursive specialist survey process. *PM R*. 2013;5(6):491–495. doi:10.1016/j.pmrj.2012.10.001

16. Genevay S, Courvoisier DS, Konstantinou K, et al. Clinical classification criteria for neurogenic claudication caused by lumbar spinal stenosis. The N-CLASS criteria. *Spine J*. 2018;18(6):941–947. doi:10.1016/j.spinee.2017.10.003

17. Fairbank JC, Pynsent PB. The Oswestry disability index. *Spine*. 2000;25(22):2940–2952. doi:10.1097/00007632-200011150-00017

18. Fairbank JC, Couper J, Davies JB, O'Brien JP. The Oswestry low back pain disability questionnaire. *Physiotherapy*. 1980;66(8):271–273.

19. Childs JD, Piva SR, Fritz JM. Responsiveness of the numeric pain rating scale in patients with low back pain. *Spine*. 2005;30(11):1331–1334. doi:10.1097/01.brs.0000164099.92112.29

20. Tibshirani R. Regression shrinkage and selection via the lasso. *J R Stat Soc Series B Stat Methodol*. 1996;58(1):267–288.

21. Kongsted A, Kent P, Albert H, Jensen TS, Manniche C. Patients with low back pain differ from those who also have leg pain or signs of nerve root involvement - a cross-sectional study. *BMC Musculoskelet Disord*. 2012;13:236. doi:10.1186/1471-2474-13-236

22. Yabuki S, Fukumori N, Takegami M, et al. Prevalence of lumbar spinal stenosis, using the diagnostic support tool, and correlated factors in Japan: a population-based study. *J Orthop Sci*. 2013;18(6):893–900. doi:10.1007/s00776-013-0455-5

23. Chiba D, Tsuda E, Wada K, et al. Lumbar spondylosis, lumbar spinal stenosis, knee pain, back muscle strength are associated with the locomotive syndrome: rural population study in Japan. *J Orthop Sci*. 2016;21(3):366–372. doi:10.1016/j.jos.2016.02.006

24. Otani K, Kikuchi S, Yabuki S, et al. Lumbar spinal stenosis has a negative impact on quality of life compared with other comorbidities: an epidemiological cross-sectional study of 1862 community-dwelling individuals. *ScientificWorldJournal*. 2013;2013:590652. doi:10.1155/2013/590652

25. Pepe MS. Incomplete data and imperfect reference tests. In: Pepe MS, editor. *The Statistical Evaluation of Medical Tests for Classification and Prediction*. Oxford, UK: Oxford University Press; 2004:168e213.

26. Rousing R, Jensen RK, Fruensgaard S, et al. Danish national clinical guidelines for surgical and nonsurgical treatment of patients with lumbar spinal stenosis. *Eur Spine J*. 2019;28(6):1386–1396. doi:10.1007/s00586-019-05987-2

27. Kreiner DS, Shaffer WO, Baisden JL, et al. An evidence-based clinical guideline for the diagnosis and treatment of degenerative lumbar spinal stenosis (update). *Spine J*. 2013;13(7):734–743. doi:10.1016/j.spinee.2012.11.059

28. The regional committees on health research ethics for Southern Denmark. 2018. Available from: https://komite.regionsyddanmark.dk/wm428123. Accessed March 8, 2018.