Spine

SPINE Volume 46, Number 12, pp 788–795

RANDOMIZED TRIAL

OPEN

Responsiveness of Outcome Measures in Nonsurgical Patients with Lumbar Spinal Stenosis

A Secondary Analysis From a Randomized Controlled Trial

Cristiane Carlesso, PT, MS,[a] Sara R. Piva, PT, PhD,[a] Clair Smith, MS,[a] Carlo Ammendolia, DC, PhD,[b] and Michael J. Schneider, DC, PhD[a]

Study Design. Secondary analysis from a randomized controlled trial on nonsurgical interventions for patients with lumbar spinal stenosis (LSS).

Objective. The aim of this study was to assess the responsiveness of the Self-Paced Walking Test (SPWT), Swiss Spinal Stenosis Questionnaire (SSS), and Oswestry Disability Index (ODI) and determine their minimal clinically important differences (MCID) in nonsurgical LSS patients.

Summary of Background Data. Limited information is available about the responsiveness of these tests in nonsurgical LSS population.

Methods. A total of 180 participants completed the SPWT, SSS, and ODI at baseline, 2, and 6 months. Responsiveness was assessed by distribution-based method, including effect size and standardized response mean, and anchor-based method, using the patient global index of change (PGIC) as the external anchor to distinguish responders and non-responders. Areas under the curve (AUC) were calculated along with MCIDs for ''minimal'' and ''moderate improvement'' subgroups.

Results. The following values represent 2- and 6-month analyses of each outcome measure, respectively. Standard effect sizes: 0.48 and 0.50 for SPWT, −0.42 and −0.36 for SSS, and −0.29 and −0.25 for ODI. Spearman correlation coefficients between PGIC and outcomes were: 0.44 and 0.39 for SPWT, −0.53 and −0.55 for SSS, and −0.46 and −0.54 for ODI. MCIDs for the ''minimal improvement'' subgroup were: 375.9 and 319.3 ms for SPWT, −5.3 and −5.8 points for SSS, and −9.3 and −10.8 points for ODI. AUCs was 0.68 to 0.76. MCIDs for the ''moderate improvement'' subgroup were: 344.2 and 538.2 m for SPWT, −5.5 and −7.5 points for SSS, and −9.1 and −13.6 points for ODI. AUCs ranged from 0.68 to 0.76.

Conclusion. The SPWT, SSS, and ODI are responsive outcome measures to assess nonsurgical patients with LSS. This finding, along with the reported MCIDs, can help clinicians to monitor changes in their patients' walking and physical function over time and make clinical decisions. They also provide researchers with reference for future studies in LSS.

Key words: anchor-based, distribution-based, lumbar spinal stenosis, minimal clinical important difference, non-surgical patients, outcome measures, responsiveness.

Level of Evidence: 2

Spine 2021;46:788–795

Lumbar spinal stenosis (LSS) is a condition that is highly associated with disability due to the narrowing of the lumbar spinal canal and compression of neurovascular structures.[1] It occurs mostly as a result of degenerative changes, with prevalence between 11% and 39% in adults presenting clinical symptoms and/or diagnostic imaging findings.[2,3] LSS is associated with limited walking capacity and physical function.[4–6] Therefore, evaluating these patients' progress during treatment of LSS requires the administration of outcome measures that are sensitive to detecting changes over time (responsive) in these domains.

Walking capacity and physical function in LSS can be measured by patient-reported outcomes (PROs) and performance-based tests. PROs address relevant aspects of patients' lives through individual items combined in a summary score that reflects their severity or disability level. Advantages of PROs include direct response from patients, low response burden, and ability to compare values across

From the [a]Department of Physical Therapy, University of Pittsburgh, Pittsburgh, PA; and [b]DC, PhD. Institute for Health Policy, Management and Evolution, University of Toronto, Mount Sinai Hospital, Toronto, Ontario, Canada.

Acknowledgment date: June 23, 2020. First revision date: September 7, 2020. Acceptance date: November 2, 2020.

The manuscript submitted does not contain information about medical device(s)/drug(s).

Patient-Centered Outcomes Research Institute award (CER-1410-25056) funds were received in support of this work.

Relevant financial activities outside the submitted work: consultancy, grants, payment for lecture, travel/accommodations/meeting expenses.

Address correspondence and reprint requests to Cristiane Carlesso, PT, MS, Department of Physical Therapy, University of Pittsburgh, Bridgeside Point 1. 100 Technology Drive, Pittsburgh, PA 15219; E-mail: crc120@pitt.edu

DOI: 10.1097/BRS.0000000000003920

studies. The Oswestry Disability Index (ODI) and Swiss Spinal Stenosis Questionnaire (SSS) are often chosen to measure symptoms and physical limitations associated with LSS and have shown adequate validity and reliability.[7–13] However, most of the studies on LSS have used these PROs in a population that included surgical patients,[14,15] those who have higher levels of disability.[16] As a result, information about the responsiveness of these outcome measures in nonsurgical LSS patients is limited.

Performance-based tests are also commonly used in the LSS population. These measures provide data directly from observation of patients' activities and capture specific features of functional skills that are highly relevant to these patients. The Self-Paced Walking Test (SPWT) is an example of a performance-based test that has shown good reliability and validity.[17] It has also been considered more accurate in measuring walking capacity in patients with LSS when compared to treadmill testing because patients walk at their individual pace mimicking real-life conditions.[18–20] Additionally, the SPWT is a simple test that does not require any complex equipment. Nevertheless, the responsiveness of this test with LSS patients has been presented in only two previous studies with small sample sizes that included surgical participants.[19,21] Thus, the responsiveness of the SPWT in LSS patients receiving nonoperative interventions remains unexplored.

Given that the majority of LSS patients have mild or moderate levels of disability and potentially benefit from non-surgical treatments,[22,23] there is an urgency for evidence to guide the selection of responsive outcome measures for use in this population. Analysis of how well PROs and performance-based tests detect changes over time will help clinicians and researchers in the field.[24] The aims of this project are to assess the responsiveness of the SPWT, SSS, and ODI in patients undergoing nonsurgical interventions for LSS and to provide minimal clinically important difference (MCID) values for each of these outcome measures.

METHODS

Study Design

This is a secondary analysis of data derived from a parent randomized controlled trial (RCT) comparing three different nonsurgical interventions for LSS patients, which has been published.[22,25] In this trial, participants were recruited from November 2013 to June 2016 and treated at the Physical Therapy—Clinical and Translation Research Center at the University of Pittsburgh. Informed consent was obtained from all participants, and the study was approved by the University of Pittsburgh Institutional Review Board (PRO12120422) and registered at ClinicalTrials.gov (NCT01943435).

Subjects were randomized to one of three nonsurgical interventions for their LSS delivered over the course of 6 weeks. In one group, patients were followed by a medical physician to manage their condition with prescription medication, advice to stay active and epidural steroid injection if warranted. Another group participated in community-based exercise classes for older adults supervised by fitness instructors. The third group had clinic-based manual therapy and individualized exercises provided by either a chiropractor or physical therapist. The outcomes were assessed at baseline (before interventions), 2, and 6 months after enrollment. In this secondary analysis, data from the three groups were combined, resulting in a wide variability of change over time among the outcome measures as recommended for responsiveness assessment.[26]

Participants

Inclusion criteria for the parent RCT were age $\geq$60 years, clinical history and diagnostic imaging evidence of LSS, ability to read and write English, neurogenic claudication, ability to engage in mild exercise, availability to participate, and willingness to be randomized. Exclusion criteria were history of metastatic cancer, cauda equina symptoms, previous lumbar decompressive surgery, history of severe peripheral artery disease, contraindication to exercise, history of neurologic condition other than LSS that affects the subject's ability to walk, inability to complete the SPWT without an assistive device or for any reason other than symptoms related to LSS.[25]

Outcome Measures

Four outcome measures were included in this responsiveness analysis: SPWT, SSS, ODI, and Patient Global Index of Change (PGIC). All outcome measures were completed at baseline, 2, and 6 months, except the PGIC which was not collected at baseline.

The SPWT is a reliable and valid performance-based test that measures walking capacity and has been suggested as the test of choice when measuring this domain in patients with LSS.[18] Participants walk at their own pace on a level surface without support until they need to stop because of LSS symptoms, or until 30 minutes have passed.[17,18] The distance walked is recorded in meters.

The SSS is a validated 18-item questionnaire measuring disability in patients with LSS using three subscales: seven-item symptom severity (SS), five-item physical functional (PF), and six-item patient satisfaction with surgery.[27,28] This current analysis used only the first two subscales since the participants were nonsurgical candidates. The total score of the combined SS and PF subscales ranges from 12 to 55 points, with higher scores representing worse symptoms and greater disability.

The ODI is a validated and reliable 10-item questionnaire evaluating limitations of daily activities caused by low back pain and has been widely used in LSS studies.[7,12,13,20] Each item is scored on a 6-point Likert scale (0–5 points). The score is transformed to a 0- to100-point percentage scale, with higher scores indicating more severe disability.

The PGIC is a self-reported measure of health status often used in chronic pain research. It is designed to quantify patients' change over time to analyze the effect of a particular intervention.[29,30] A 7-point PGIC scale was used to

quantify the amount of change since the start of treatment. Patients rated their overall status as "very much worse," "much worse," "minimally worse," "no change," "minimally improved," "much improved," or "very much improved." The descriptors were given numerical values from −3 (very much worse) to 0 (no change) to +3 (very much improved).

Statistical Analysis

Paired *t* tests were used to identify whether participants changed over time for the ODI, SSS, and SPWT at both 2- and 6-month follow-ups. Responsiveness of the SPWT, SSS, and ODI was investigated by distribution and anchor-based methods.

Distribution-based responsiveness was determined by first obtaining the mean change scores for each outcome measure (follow-up score minus baseline score) from the entire sample. Standardized effect sizes and the standardized response means were calculated for each outcome measure at both timepoints. Standard effect size is the mean change score divided by the standard deviation of the baseline score and standard response mean is the mean change score divided by the standard deviation of the change score.[26,31] According to Cohen, effect sizes of 0.2 are considered small, 0.5 medium, and 0.8 large.[32]

Anchor-based analysis selected the PGIC as the external anchor.[33] Spearman correlation coefficients (Rho) were then calculated between the mean change in each outcome measure and the PGIC scores at both follow-ups. Rho values can be interpreted as low for values below 0.3, moderate for values between 0.3 and 0.6, and strong for values >0.6.[34] Based on the PGIC, we defined two subgroups of responders. The "minimal improvement" subgroup included patients who responded at least "minimally improved" when asked about their overall status (PGIC ≥1). The "moderate improvement" subgroup included patients who responded at least "much improved" to the same question (PGIC ≥2).

Mean changes of the outcome measures were calculated at 2 and 6 months for responders and nonresponders in each previously described subgroup.[35] MCID is the smallest amount of change that represents a clinically meaningful improvement.[31] In this analysis, we selected two methods to derive each MCID to examine their consistency. The first method ($MCID^1$) corresponded to the mean improvement of the responders. The second method ($MCID^2$) refers to the difference between the mean changes of the responders and non-responders.[36–38]

Receiver-operating characteristic (ROC) curves, their respective Area Under the ROC Curves (AUC), and their 95% confidence intervals were calculated[26,39] to quantify the ability of each outcome measures to distinguish patients who responded to an intervention over time from those who did not, based on the PGIC.[39] AUC values can be interpreted as acceptable discrimination between 0.7 and 0.8, excellent discrimination between 0.8 and 0.9, and outstanding discrimination when ≥0.9.[35] All analyses were conducted in SAS version 9.4 (SAS Institute, Cary, NC).

RESULTS

Two hundred fifty-nine subjects were enrolled in the parent clinical trial. Of these, we analyzed the data from the subset of 180 subjects who had completed the SPWT, SSS, ODI, and PGIC at all timepoints as described previously. Participants' characteristics are presented in Table 1. Their mean changes from baseline to 2 and 6 months were an increase of 214 and 223 m for the SPWT, a decrease of 2.5 and 2.1 points for the SSS, and a decrease of 3.7 and 3.2 points for the ODI, respectively (Table 2). The changes over time in all outcome measures were significant ($P < 0.01$). The magnitude of effect sizes ranged from small to medium, being 0.48 and 0.50 for the SPWT, −0.42 and −0.36 for the SSS, and −0.29 and −0.25 for the ODI, at 2 and 6 months, respectively (Table 3).

The correlations between the PGIC and each outcome measure were moderate (Rho: 0.39–0.54). For the "minimal improvement" subgroup at 2 months, the $MCIDs^1$ and $MCIDs^2$ were 331 and 376 m for the SPWT, −4.2 and −5.3 points for SSS, and −6.6 and −9.3 points for ODI, respectively. At 6 months, the $MCIDs^1$ and $MCIDs^2$ were 346 and 319 m for the SPWT, −4.4 and −5.8 points for SSS, and −7.4 and −10.8 points for ODI, respectively. For the "moderate improvement" subgroup at 2 months, the $MCIDs^1$ and $MCIDs^2$ were 436 and 344 m for the SPWT, −6.1 and −5.5 points for SSS, and −9.6 and −9.1 points for ODI, respectively. At 6 months, the $MCIDs^1$ and $MCIDs^2$ were 621 and 538 m for the SPWT, −7.1 and −7.5 points for SSS, and −13.3 and −13.6 points for ODI, respectively (Table 4). The AUCs are presented in Figure 1.

DISCUSSION

Walking capacity and physical function are the most fundamental parameters for determining the clinical progress and treatment effectiveness in nonsurgical LSS, highlighting the importance of having responsive instruments to measure these outcomes. To our knowledge, this is the first study to assess the responsiveness of the SPWT, SSS, and ODI in nonsurgical LSS patients with moderate disability. The results suggest that all outcome measures analyzed exhibit an adequate level of responsiveness in this population.

This study is novel because previous studies on LSS have investigated outcome measurements responsiveness in patients with greater level of disability, having surgery as a reasonable intervention. Based on this, it becomes difficult to extrapolate responsiveness findings from these studies to nonsurgical LSS patients. For example, our mean baseline ODI score of 37.8 points is considerably lower than the corresponded value of 45.3 points obtained in the as-treated analysis of patients with LSS undergoing surgery in the Spine Patient Outcomes Research Trial.[15,40]

The ability of the SPWT and PROs to monitor clinically important changes over time is consistent with results from

TABLE 1. Baseline Characteristics

Characteristic (n = 180)	Value
Age, y, Mean ± SD	73.1 ± 7.6
Female, n (%)	97 (54)
BMI, kg/m^2, Mean ± SD	30.6 ± 6.3
Smoking status, n (%)	
Never	75 (42)
Former	90 (50)
Current	11 (6)
Race, n (%)	
White	142 (79)
Black	37 (21)
Other	1 (1)
Married, n (%)	99 (55)
Household income >$40,000/y, n (%)	89 (49)
Education, n (%)	
High school	27 (15)
Any college or technical training	145 (81)
No. of comorbidities, Mean ± SD	4.5 ± 2.2
Duration of back symptoms, n (%)	
≤6 mo	19 (11)
>6 mo	161 (89)
Duration of leg symptoms, n (%)	
≤6 mo	47 (26)
>6 mo	133 (74)
Diagnostic imaging results*, n (%)	
Central canal stenosis	97 (54)
Lateral recess stenosis	144 (80)
Foraminal stenosis	150 (83)
Spondylolisthesis present	109 (62)

BMI indicates body mass index.

**Percentages do not add to 100 because participants could have more than one diagnostic imaging result.*

some related clinical trials. The SPWT in this study presented MCIDs ranging from 319 to 376 m, which are comparable to one study on surgical LSS patients reporting 387 m[21] and to another study on surgical and non-surgical patients showing 363 m[19] as their MCID. The SSS was able to differentiate responders and nonresponders in our analysis (AUCs: 0.76–0.83) by using the first two subscales, which was also reported in a study of surgical LSS patients (AUC: 0.83) using a similar anchor.[41] This fact lends credibility to the MCIDs we derived for the SSS to be used with nonsurgical patients. The ODI in our study presented MCIDs ranging from a reduction of 6.6 to 13.3 points, which are values comparable to a reduction of 5.3 and 9.5 points found in two studies of patients with nonspecific chronic low back pain undergoing conservative therapy.[42,43]

Although these results indicate that the SPWT and PROs are similarly responsive, there were some slight variations between them. The SPWT demonstrated larger effect sizes in the distribution-based method, whereas the PROs presented a slightly higher correlation with the PGIC and larger AUCs. These minor differences can be explained by the distinct methods used to assess responsiveness. The larger association between the PROs and the external anchor might be related to the fact that they are all self-reported measures

TABLE 3. Distribution-based Responsiveness (n = 180)

	SPWT*	SSS†	ODI†
Standardized effect size‡			
Baseline to 2 mo	0.48	−0.42	−0.29
Baseline to 6 mo	0.50	−0.36	−0.25
Standardized response mean§			
Baseline to 2 mo	0.44	−0.45	−0.33
Baseline to 6 mo	0.37	−0.34	−0.27

ODI indicates Oswestry Disability Index; SD, standard deviation; SPWT, Self-Paced Walking Test; SSS, Swiss Spinal Stenosis Questionnaire.

**Positive values represent improved physical function.*

†Negative values represent improved physical function.

‡Standardized Effect Size was calculated as: mean change/SD baseline.

§Standardized Response Mean was calculated as: mean change/SD change.

TABLE 2. Outcomes Over Time (n = 180)

Timepoints	SPWT	SSS	ODI
Baseline, mean ± SD	446.3 ± 449.4	31.1 ± 6.0	37.8 ± 12.9
2 mo, mean ± SD	660.4 ± 639.7	28.6 ± 6.5	34.1 ± 14.8
6 mo, mean ± SD	669.7 ± 700.9	29.0 ± 6.7	34.6 ± 14.5
Change Δ	**SPWT***	**SSS†**	**ODI†**
Baseline to 2 mo ± SD (95% CI)	214.2 ± 489.3 (142.2 to 286.1)	−2.5 ± 5.6 (−3.3 to −1.7)	−3.7 ± 11.3 (−5.4 to −2.1)
Baseline to 6 mo ± SD (95% CI)	223.4 ± 598.6 (135.4 to 311.4)	−2.1 ± 6.3 (−3.1 to −1.2)	−3.2 ± 12.1 (−5.0 to −1.5)

95% CI indicates 95% confidence interval derived from paired t-test; ODI, Oswestry Disability Index (Score ranges from 0 to 100. Higher scores indicate grater disability/severity); SD, standard deviation; SPWT, Self-Paced Walking Test (distance in meters walked up to 30 minutes); SSS, Swiss Spinal Stenosis Questionnaire (score ranges from 12 to 55. Higher scores indicate grater disability/severity).

**Positive values represent improved physical function.*

†Negative values represent improved physical function.

TABLE 4. Anchor-based Responsiveness

	SPWT*				SSS†				ODI†			
	n	0–2 mo	N	0–6 mo	n	0–2 mo	n	0–6 mo	n	0–2 mo	n	0–6 mo
Spearman rho‡ (*P*)	180	0.44 (<0.0001)	180	0.39 (<0.0001)	180	−0.53 (<0.0001)	180	−0.55 (<0.0001)	180	−0.46 (<0.0001)	180	−0.54 (<0.0001)
Minimal improvement§												
Responders, Mean ± SD	124	331.1 ± 490.2	111	345.8 ± 576.9	124	−4.2 ± 5.3	111	−4.4 ± 6.2	124	−6.6 ± 11.1	111	−7.4 ± 12.1
Nonresponders, Mean ± SD	56	−44.8 ± 377.9	69	26.5 ± 583.9	56	1.1 ± 4.2	69	1.4 ± 4.7	56	2.7 ± 8.7	69	3.4 ± 8.8
MCID[1]	180	331.1	180	345.8	180	−4.2	180	−4.4	180	−6.6	180	−7.4
MCID[2]	180	375.9	180	319.3	180	−5.3	180	−5.8	180	−9.3	180	−10.8
AUC (95% CI)	180	0.76 (0.69–0.84)	180	0.68 (0.60–0.76)	180	0.78 (0.71–0.85)	180	0.76 (0.69–0.83)	180	0.76 (0.68–0.83)	180	0.76 (0.69–0.83)
Moderate improvement‖												
Responders, Mean ± SD	64	436.0 ± 519.2	47	621.1 ± 666.5	64	−6.1 ± 5.4	47	−7.7 ± 6.4	64	−9.6 ± 12.0	47	−13.3 ± 12.9
Nonresponders, Mean ± SD	116	91.8 ± 427.2	133	82.9 ± 504.8	116	−0.6 ± 4.6	133	−0.2 ± 5.0	116	−0.5 ± 9.4	133	0.3 ± 9.6
MCID[1]	180	436.0	180	621.1	180	−6.1	180	−7.7	180	−9.6	180	−13.3
MCID[2]	180	344.2	180	538.2	180	−5.5	180	−7.5	180	−9.1	180	−13.6
AUC (95% CI)^	180	0.71 (0.63–0.79)	180	0.74 (0.65–0.83)	180	0.77 (0.70–0.85)	180	0.82 (0.75–0.90)	180	0.73 (0.65–0.81)	180	0.81 (0.73–0.88)

95% CI indicates 95% confidence interval; AUC, area under the receiver operating characteristic curve; MCID, minimal clinically important difference; MCID[1], method 1 (defined as the mean change of the responders from baseline); MCID[2], method 2 (defined as the difference between responders and nonresponders); ODI, Oswestry Disability Index (total score ranges from 0 to 100. Higher scores indicate grater disability/severity); SD, standard deviation; SPWT, Self-Paced Walking test (total distance in meters walked up to 30 minutes); SSS, Swiss Spinal Stenosis Questionnaire (total score ranges from 12 to 55. Higher scores indicate grater disability/severity).

**Positive values represent improved physical function.*

†Negative values represent improved physical function.

‡Spearman correlation coefficient between outcome and patient global index of change.

§Patients who reported at least "minimally improved" on the Patient Global Index of Change (PGIC ≥1).

‖Patients who reported at least "much improved" on the Patient Global Index of Change (PGIC ≥2).

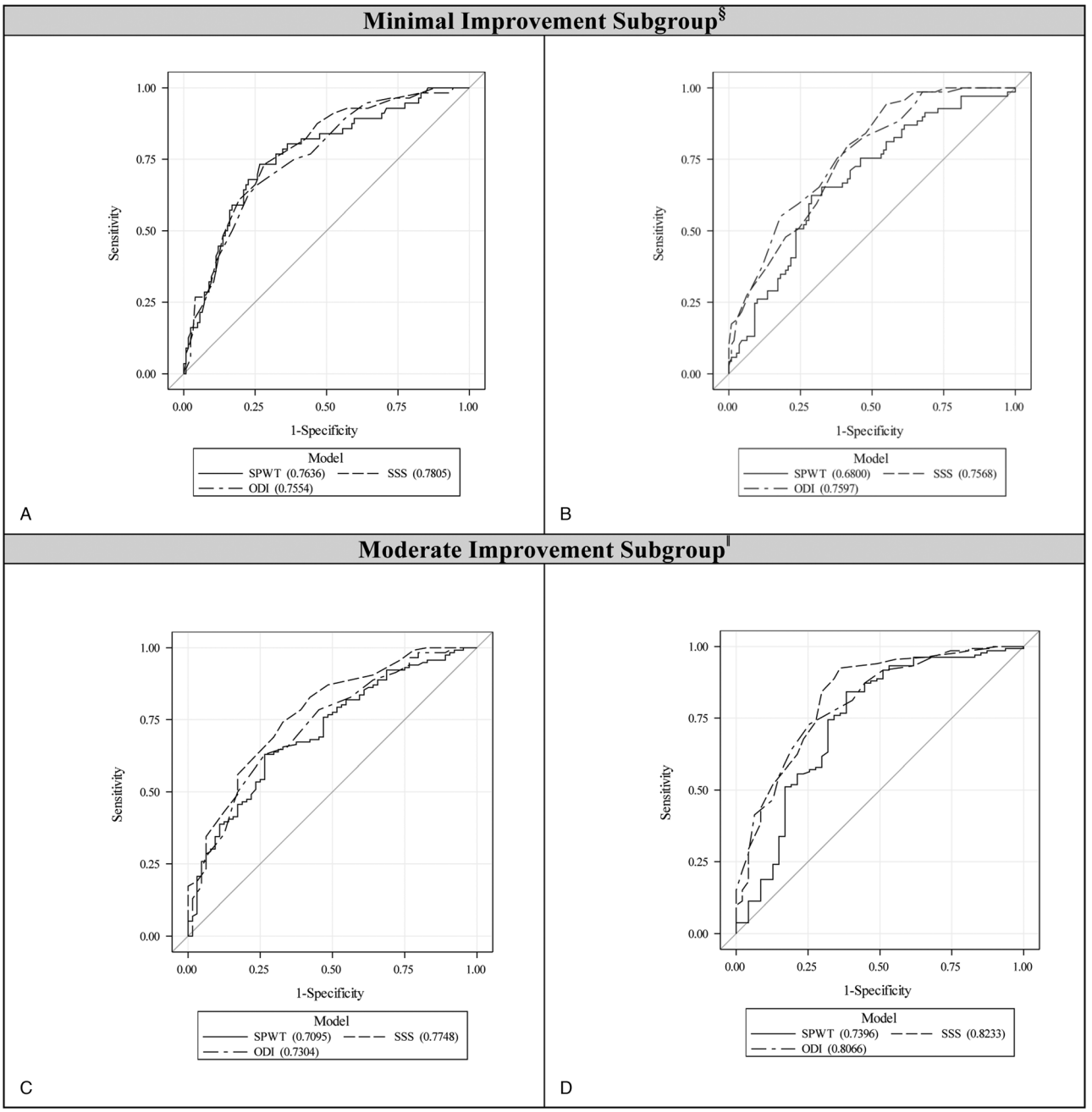

Figure 1. Receiver-operating characteristics curves (ROC) of outcome measures at 2 and 6 months. ODI indicates Oswestry Disability Index; SPWT, Self-Paced Walking Test; SSS, Swiss Spinal Stenosis Questionnaire. §Patients who responded at least ''minimally improved'' on the Patient Global Index of Change (PGIC $\geq$1). ‖Patients who reported at least ''much improved'' on the Patient Global Index of Change (PGIC $\geq$2). (**A**) Minimal Improvement Subgroup assessed at 2 months (responders = 124). (**B**) Minimal Improvement Subgroup assessed at 6 months (responders = 111). (**C**) Moderate Improvement Subgroup assessed at 2 months (responders = 64). (**D**) Moderate Improvement Subgroup assessed at 6 months (responders = 47).

and represent patients' perception of their own health status. Therefore, how patients perceive their changes affects the PROs and PGIC evenly, whereas the SPWT represents a direct observation of the patient's walking performance.

The AUCs of all outcome measures show acceptable to excellent discrimination between responders and non-responders at both follow-ups. Two methods of deriving MCIDs were included to check whether their values were consistent. The first method ($MCID^1$) was considered the mean change of the responders, whereas the second method ($MCID^2$) was calculated by taking the difference between the mean changes of the responders and nonresponders. We expected both methods to provide consistently larger MCID values for the ''moderate'' *versus* ''minimal improvement''

subgroups. However, the MCID[2] provided a smaller value of 344 m for the "moderate improvement" subgroup compared to the value of 376 m for the "minimal improvement" subgroup. This happened because the nonresponders in the "minimal improvement" subgroup walked an average of 45 m less at 2 months than at baseline (due to some extreme values). Therefore, we recommend using the MCID[1] values as more consistent estimates of MCIDs for all outcome measures.

Limitations of this study include the inability to use the entire sample from the parent RCT (n = 259) because of dropouts, which led to incomplete data collection required for the analysis. We cannot rule out the possibility that the 180 participants included in this analysis may represent those who experienced better outcomes and were more cooperative with returning for their follow-ups. Another limitation is the inability to derive MCIDs based on the AUC (*e.g.*, Youden index) because the ROC curves did not provide data points with an adequate level of sensitivity and specificity.

Despite limitations, this analysis provides corroboration for responsiveness of three commonly used outcome measures in a large sample of patients with LSS undergoing different non-surgical interventions. Having different conservative approaches combined in the analysis enhances the generalizability of our results, which may beneficially affect both clinical and research settings. Our findings provide scientific evidence for clinicians to use the SSS or ODI as measures of self-reported disability, and the SPWT as an objective measure of walking performance to monitor the clinical progress of their LSS patients. Clinicians can use the reported MCIDs as reasonable estimates of clinical progress to support modifications of their interventions and in the decision-making process for surgical consultation when a patient has not achieved the MCID within a reasonable period. Researchers may also find these MCIDs and effect sizes to be useful as reference points for sample size and power calculations for future studies involving patients with LSS.

CONCLUSION

The SPWT, SSS, and ODI exhibit an adequate level of responsiveness as outcome measures to assess nonsurgical patients with LSS. We presented MCIDs for each of these outcome measures derived from both distributional and anchor-based methods, which may be of benefit in both clinical and research settings.

➢ Key Points

- ❑ This study aimed at providing evidence about the responsiveness of outcome measures in a nonsurgical LSS population to fill this gap of knowledge existing in the current literature.
- ❑ The SPWT, SSS, and ODI are responsive outcome measures to assess nonsurgical patients with LSS.
- ❑ Clinicians and researchers can use these outcome measures with their nonsurgical LSS patients.
- ❑ Clinicians may use the MCIDs provided as estimates of clinical progress to monitor changes over time and support decision-making related to their patients.
- ❑ Researchers may use these MCIDs, along with the effect sizes, as reference for empowering future studies in this population.

References

1. Fritz JM, Delitto A, Welch WC, et al. Lumbar spinal stenosis: a review of current concepts in evaluation, management, and outcome measurements. *Arch Phys Med Rehabil* 1998;79:700–8.
2. Kalichman L, Cole R, Kim DH, et al. Spinal stenosis prevalence and association with symptoms: the Framingham Study. *Spine J* 2009;9:545–50.
3. Jensen RK, Jensen TS, Koes B, et al. Prevalence of lumbar spinal stenosis in general and clinical populations: a systematic review and meta-analysis. *Eur Spine J* 2020;29:2143–63.
4. Winter CC, Brandes M, Muller C, et al. Walking ability during daily life in patients with osteoarthritis of the knee or the hip and lumbar spinal stenosis: a cross sectional study. *BMC Musculoskelet Disord* 2010;11:233.
5. Kim HJ, Chun HJ, Han CD, et al. The risk assessment of a fall in patients with lumbar spinal stenosis. *Spine (Phila Pa 1976)* 2011;36:E588–92.
6. Tomkins-Lane C, Melloh M, Lurie J, et al. ISSLS prize winner: consensus on the clinical diagnosis of lumbar spinal stenosis: results of an International Delphi Study. *Spine (Phila Pa 1976)* 2016;41:1239–46.
7. Fairbank JC, Couper J, Davies JB, et al. The Oswestry low back pain disability questionnaire. *Physiotherapy* 1980;66:271–3.
8. Fairbank JC, Pynsent PB. The Oswestry Disability Index. *Spine (Phila Pa 1976)* 2000;25:2940–52; discussion 52.
9. Heshmati AA, Mirzaee M. Reliability and validity of the Swiss Spinal Stenosis Questionnaire for Iranian patients with lumbar spinal stenosis. *Arch Bone Jt Surg* 2018;6:119–23.
10. Marchand AA, Tetreau C, O'Shaughnessy J, et al. French-Canadian adaptation and validation of the Swiss Spinal Stenosis Questionnaire for patients with lumbar spinal stenosis. *Spine (Phila Pa 1976)* 2019;44:E487–93.
11. Tomkins CC, Battie MC, Hu R. Construct validity of the physical function scale of the Swiss Spinal Stenosis Questionnaire for the measurement of walking capacity. *Spine (Phila Pa 1976)* 2007;32:1896–901.
12. Fokter SK, Yerby SA. Patient-based outcomes for the operative treatment of degenerative lumbar spinal stenosis. *Eur Spine J* 2006;15:1661–9.
13. Roland M, Fairbank J. The Roland-Morris Disability Questionnaire and the Oswestry Disability Questionnaire. *Spine (Phila Pa 1976)* 2010;25:3115–24.
14. Patel AA, Dodwad SM, Boody BS, et al. Validation of Patient Reported Outcomes Measurement Information System (PROMIS) Computer Adaptive Tests (CATs) in the surgical treatment of lumbar spinal stenosis. *Spine (Phila Pa 1976)* 2018;43:1521–8.
15. Weinstein JN, Tosteson TD, Lurie JD, et al. Surgical versus nonsurgical therapy for lumbar spinal stenosis. *N Engl J Med* 2008;358:794–810.
16. Neuman BJ, Baldus C, Zebala LP, et al. Patient factors that influence decision making: randomization versus observational nonoperative versus observational operative treatment for adult symptomatic lumbar scoliosis. *Spine (Phila Pa 1976)* 2016; 41:E349–58.
17. Tomkins-Lane CC, Battie MC. Validity and reproducibility of self-report measures of walking capacity in lumbar spinal stenosis. *Spine (Phila Pa 1976)* 2010;35:2097–102.

18. Tomkins CC, Battie MC, Rogers T, et al. A criterion measure of walking capacity in lumbar spinal stenosis and its comparison with a treadmill protocol. *Spine (Phila Pa 1976)* 2009;34:2444–9.
19. Tomkins-Lane CC, Battie MC, Macedo LG. Longitudinal construct validity and responsiveness of measures of walking capacity in individuals with lumbar spinal stenosis. *Spine J* 2014;14:1936–43.
20. Conway J, Tomkins CC, Haig AJ. Walking assessment in people with lumbar spinal stenosis: capacity, performance, and self-report measures. *Spine J* 2011;11:816–23.
21. Rainville J, Childs LA, Pena EB, et al. Quantification of walking ability in subjects with neurogenic claudication from lumbar spinal stenosis—a comparative study. *Spine J* 2012;12:101–9.
22. Schneider MJ, Ammendolia C, Murphy DR, et al. Comparative clinical effectiveness of nonsurgical treatment methods in patients with lumbar spinal stenosis: a randomized clinical trial. *JAMA Netw Open* 2019;2:e186828.
23. Delitto A, Piva SR, Moore CG, et al. Surgery versus nonsurgical treatment of lumbar spinal stenosis: a randomized trial. *Ann Intern Med* 2015;162:465–73.
24. Anderson DB, Mathieson S, Eyles J, et al. Measurement properties of walking outcome measures for neurogenic claudication: a systematic review and meta analysis. *Spine J* 2019;19:1378–96.
25. Schneider M, Ammendolia C, Murphy D, et al. Comparison of non-surgical treatment methods for patients with lumbar spinal stenosis: protocol for a randomized controlled trial. *Chiropr Man Ther* 2014;22:19.
26. Husted JA, Cook RJ, Farewell VT, et al. Methods for assessing responsiveness: a critical review and recommendations. *J Clin Epidemiol* 2000;53:459–68.
27. Comer CM, Conaghan PG, Tennant A. Internal construct validity of the Swiss Spinal Stenosis questionnaire: Rasch analysis of a disease-specific outcome measure for lumbar spinal stenosis. *Spine (Phila Pa 1976)* 2011;36:1969–76.
28. Stucki G, Daltroy L, Liang MH, et al. Measurement properties of a self-administered outcome measure in lumbar spinal stenosis. *Spine (Phila Pa 1976)* 1996;21:796–803.
29. Ferguson L, Scheman J. Patient global impression of change scores within the context of a chronic pain rehabilitation program. *J Pain* 2009;10:S73.
30. Marchand AA, Suitner M, O'Shaughnessy J, et al. Effects of a prehabilitation program on patients' recovery following spinal stenosis surgery: study protocol for a randomized controlled trial. *Trials* 2015;16:483.
31. Copay AG, Subach BR, Glassman SD, et al. Understanding the minimum clinically important difference: a review of concepts and methods. *Spine J* 2007;7:541–6.
32. Cohen J. A power primer. *Psychol Bull* 1992;112:155–9.
33. Parai C, Hagg O, Lind B, et al. The value of patient global assessment in lumbar spine surgery: an evaluation based on more than 90,000 patients. *Eur Spine J* 2018;27:554–63.
34. Campbell MJ, Swinscow TDV. Correlation and regression. In: Statistics at Square One. 11th ed. New York, NY2009.
35. Dworkin RH, Turk DC, McDermott MP, et al. Interpreting the clinical importance of group differences in chronic pain clinical trials: IMMPACT recommendations. *Pain* 2009;146:238–44.
36. Hagg O, Fritzell P, Nordwall A. The clinical importance of changes in outcome scores after treatment for chronic low back pain. *Eur Spine J* 2003;12:12–20.
37. Stucki G, Liang MH, Fossel AH, et al. Relative responsiveness of condition-specific and generic health status measures in degenerative lumbar spinal stenosis. *J Clin Epidemiol* 1995;48:1369–78.
38. Copay AG, Glassman SD, Subach BR, et al. Minimum clinically important difference in lumbar spine surgery patients: a choice of methods using the Oswestry Disability Index, Medical Outcomes Study questionnaire Short Form 36, and pain scales. *Spine J* 2008;8:968–74.
39. Deyo RA, Centor RM. Assessing the responsiveness of functional scales to clinical change: an analogy to diagnostic test performance. *J Chronic Dis* 1986;39:897–906.
40. Lurie JD, Tosteson TD, Tosteson A, et al. Long-term outcomes of lumbar spinal stenosis: eight-year results of the Spine Patient Outcomes Research Trial (SPORT). *Spine (Phila Pa 1976)* 2015;40:63–76.
41. Mannion AF, Fekete TF, Wertli MM, et al. Could less be more when assessing patient-rated outcome in spinal stenosis?. *Spine (Phila Pa 1976)* 2015;40:710–8.
42. Ma C, Wu S, Xiao L, et al. Responsiveness of the Chinese version of the Oswestry disability index in patients with chronic low back pain. *Eur Spine J* 2011;20:475–81.
43. Monticone M, Baiardi P, Vanti C, et al. Responsiveness of the Oswestry Disability Index and the Roland Morris Disability Questionnaire in Italian subjects with sub-acute and chronic low back pain. *Eur Spine J* 2012;21:122–9.