Validation of the Oswestry Disability Index in Adult Spinal Deformity

Omid Jalali, MD, Justin S. Smith, MD, PhD, Shay Bess MD, Richard Hostin MD, Renaud Lafage, MS, Virginie Lafage, PhD, Christopher I. Shaffrey, MD, Christopher P. Ames, MD, Lawrence G. Lenke, MD, Michael P. Kelly, MD, on behalf of the International Spine Study Group

1 Department of Orthopaedic Surgery, University of California, San Diego, San Diego, CA

2 Department of Neurosurgery, University of Virginia, Charlottesville, VA

3 Presbyterian St Luke’s Medical Center, Denver, CO, USA

© 2024 Wolters Kluwer Health, Inc. All rights reserved. Unauthorized reproduction of the article is prohibited.
Corresponding Author

Michael P. Kelly, MD
University of California, San Diego
Rady Children’s Hospital
Mkelly5@rchsd.org

Conflict of Interest Disclosure

The International Spine Study Group (ISSG) is funded through research grants from Medtronic, Globus, Stryker, SI Bone and individual donations.

Key Words: Oswestry disability index, adult spinal deformity, patient-reported outcomes, validation, scoliosis
Authorship Appendix

International Spine Study Group

Jay Turner, MD¹
Alan Daniels, MD²
Bassel Diebo, MD²
Lawrence G. Lenke, MD³
Christopher I. Shaffrey, MD⁴
Richard Hostin, MD⁵
Han Jo Kim, MD⁶
Khaled Kebaish, MD⁷
Douglas C. Burton, MD⁸
Frank J. Schwab, MD⁹
Virginie Lafage, PhD⁹
Renaud Lafage, BS⁹
Themistocles S. Protopsaltis, MD¹⁰
Peter G. Passias, MD¹⁰
Shay Bess, MD¹¹
Breton Line, BS¹¹
Michael P. Kelly, MD¹²
Gregory M. Mundis, MD¹³
Robert K. Eastlack, MD¹³
Eric O. Klineberg, MD¹⁴
Christopher P. Ames, MD¹⁵
Alex Soroceanu, MD¹⁶
Jeffrey L. Gum, MD¹⁷
Kojo Hamilton, MD¹⁸
Thomas Buell, MD¹⁸
Stephen Lewis, MD¹⁹
Justin S. Smith, MD, PhD
Munish C. Gupta, MD

1 Barrow Neurological Institute, Phoenix, AZ
2 Department of Orthopaedic Surgery, Warren Alpert School of Medicine, Brown University, Providence, RI
3 Och Spine Hospital, Department of Orthopedic Surgery, Columbia University College of Physicians and Surgeons, New York, NY
4 Department of Neurosurgery, Duke University Medical Center, Durham, NC
5 Department of Orthopedic Surgery, Baylor Scoliosis Center, Dallas, TX
6 Department of Orthopedic Surgery, Hospital for Special Surgery, New York, NY
7 Department of Orthopaedic Surgery, Johns Hopkins University School of Medicine, Baltimore, MD
8 Department of Orthopaedic Surgery, University of Kansas Medical Center, Kansas City, KS
9 Department of Orthopaedics, Northwell Health, Lenox Hill Hospital, Northwell Health, New York, NY
10 Departments of Orthopaedic Surgery, NYU Langone Medical Center, New York, NY
11 Department of Spine Surgery, Denver International Spine Clinic, Presbyterian St. Luke's/Rocky Mountain Hospital for Children, Denver, CO
12 Department of Orthopedic Surgery, Rady Children's Hospital, San Diego, CA
13 Department of Orthopaedic Surgery, Scripps Clinic, San Diego
14 Department of Orthopedic Surgery, University of Texas, Houston, TX
15 Department of Neurosurgery, University of California, San Francisco, San Francisco, CA
16 Department of Orthopaedic Surgery, University of Calgary, Calgary, AB T2N 1N4, Canada
17 Department of Orthopaedic Surgery, Norton Leatherman Spine Center, Louisville, KY
18 Department of Neurosurgery, University of Pittsburgh School of Medicine, Pittsburgh, PA
19 Department of Orthopaedic Surgery, Toronto Western Hospital, Schroeder Arthritis Institute, University of Toronto, Toronto, ON, Canada
20 Department of Neurosurgery, University of Virginia Medical Center, Charlottesville, VA
21 Department of Orthopedic Surgery, Washington University in St. Louis, St. Louis, MO
Abstract

Study Design: Retrospective cohort

Objective: To examine the validity of the ODI in ASD patients treated with surgery.

Summary of Background Data: The ODI is a patient-reported outcome measure (PROM) of low back pain and disability. While nearly ubiquitous in adult spinal deformity (ASD) research, the measure has not been validated in this patient population.

Methods: A registry of ASD patients was queried for baseline and 1-year PROM data, including the ODI, the SRS-22r, and the Patient Reported Outcomes Measurement Information System – Pain Interference (PI), and -Physical Function (PF) computer adaptive tests (CAT). Internal reliability was assessed with Cronbach’s alpha, where values ≥ 0.7 are considered reliable. Validity was assessed with Spearman correlation coefficients calculated for the ODI against validated PROMIS-Pain Interference and -Physical Function and legacy measures SRS-Pain, SRS-Activity. Responsiveness to change was measured with the adjusted effect size (aES).

Results: 325 patients were enrolled, with 208 completing baseline and one-year PROMs. The majority (149, 72%) were female and white (193, 93%), median Charlson Comorbidity Index 0 (IQR 0-2). The majority of cases included sagittal plane deformity (mean T1PA 24.2° (13.9). Cronbach’s alpha showed excellent internal reliability (Baseline = 0.89, 1yr = 0.90). ODI was valid, with strong correlations between PROMIS-PI, -PF, SRS-Pain, and SRS-Activity at
baseline and one-year follow-up. All measures were responsive to change, with the ODI showing greater responsiveness than PROMIS-PI, PROMIS-PF and SRS-Activity.

Conclusions: The ODI is a valid measure of disability as measured by pain and function in patients with ASD. It is responsive to change in a manner not different from validated PROMIS-CAT or the SRS-22r legacy measure. It is multidimensional, however, as it assesses both pain and function simultaneously. It does not measure disability related to Self-Image and may not account for all disease-related disability in ASD patients.

Key Points

1. The Oswestry Disability Index is frequently used to measure health-related quality of life in adult spinal deformity, though it has not been validated.
2. The Oswestry Disability Index measures pain and disability, with responsiveness to change, in a manner similar to the SRS-22r and the Patient Reported Outcomes Measurement Information System.
3. Self-Image is a domain affecting spinal deformity patients and it is not measured by the Oswestry Disability Index.
Introduction

Adult spinal deformity (ASD) represents a broad range of pathologies resulting in abnormal deviations in spinal alignment with significant impact on patient’s quality of life. Due to an aging population and increase in life expectancies, the prevalence of ASD grows and it is predicted that over 60 million adults will have some sort of spinal deformity by 2050. In conjunction with increasing utilization of surgery for ASD, the substantial costs of surgery for ASD highlight the potential economic burden of treatment of this condition.

Though surgery has been shown to be superior to the nonoperative treatment of ASD, outcomes after surgery are variable and difficult to predict. As emphasis on value-based care continues to grow, there remains an ever increasing importance of patient-reported outcome measures (PROMs) which reflect the impact of ASD on patients’ health related quality of life (HRQoL) and are used to evaluate the effectiveness of treatments in ASD care. Commonly employed HRQoL measures in clinical practice and research for ASD include generic measures of health, such as the Patient-Reported Outcomes Measurement Information System (PROMIS) and Short Form-36 (SF-36), and disease-specific instruments, including the Scoliosis Research Society-22r (SRS-22r) questionnaire and the Oswestry Disability Index (ODI). Both the PROMIS system and SRS-22r have been shown to be valid and responsive to change in ASD.

The ODI is a disease-specific PROM widely accepted as the ‘gold standard’ for measuring disability and health related quality of life (HRQoL) in a variety of lumbar degenerative diseases. The ODI takes 5 minutes to complete and less than 1 minute to score. It has been validated in several studies for lumbar degenerative diseases and has been adapted and validated across a multitude of languages and cultural groups, and has been shown to be one of...
the most readable PROMs in orthopedics\textsuperscript{17}. Such factors likely contribute to the finding that the ODI is the most commonly used PROM in studies measuring outcomes in ASD surgery\textsuperscript{8}. Despite its nearly ubiquitous use in ASD research, the ODI has not been validated in the ASD population to our knowledge\textsuperscript{11}.

The assessment of a HRQoL measure to comprehensively characterize a specific disease condition requires a step-by-step process. The Consensus-based Standards for the selection of health status Measurement Instrument (COSMIN) checklist is widely accepted as the gold standard for the reporting of studies on measurement properties of PROMs\textsuperscript{18}. They describe nine criteria important for the characterization of the measurement properties of a PROM: content validity, structural validity, internal consistency, cross-cultural validity/measurement invariance, reliability, measurement error, criterion validity, hypotheses testing for construct validity, and responsiveness\textsuperscript{19}. Due to the complexity involved in exploring each of these criteria, multiple studies are typically required to establish the material properties of a given PROM\textsuperscript{20}, which includes three main categories: validity, reliability, and responsiveness\textsuperscript{21}.

The objective of the present study was to examine the measurement properties of the ODI in the context of surgical ASD patients. To achieve this, we sought to establish the criterion validity, internal reliability, and responsiveness to change of ODI scores at 1 year following ASD surgery. We hypothesized that the patients would demonstrate significant improvements in ODI scores and that the ODI would be reliable and correlate strongly with the corresponding domains of the PROMIS-CAT system and SRS-22r legacy scores. Such findings would provide justification for the current utilization of the ODI in the assessment of disability and HRQoL in ODI patients and provide context for the interpretation for the multitude of studies exploring outcomes following surgery for ASD.
Materials and Methods

Study Design, Population, and Data Collection

This study is a retrospective cohort study of patients enrolled in a multicenter adult thoracolumbar spinal deformity database. Institutional review board approval was obtained at all 11 participating centers. Patients were enrolled in the study if they were 18-years or older and meet at least one of the following criteria: 1) complex surgical procedure, defined as > 12 levels fused or at last one three-column osteotomy (3CO) or anterior column release (ACR); 2) geriatric deformity surgery, defined as age > 65 years old and at least 7 levels fused; and 3) severe radiographic deformity, defined as at least one of PI-LL ≥ 25°, TPA ≥ 30°, SVA ≥ 15 cm, thoracic cobb angle ≥ 70° or thoraco-lumbar cobb ≥ 50°. Exclusion criteria were patients with an inflammatory or autoimmune disease, a neuromuscular disorder (ie. Parkinson’s), neoplasm, infection, syndromic scoliosis or post-traumatic deformities. All included patients signed informed consent for inclusion in the multicenter database prior to surgery.

The registry of ASD patients was queried for all patients who underwent surgical treatment of ASD and completed both baseline and 1-year post-operative PROM questionnaires. 1-year was chosen for validation of this PROM as recovery is complete at 1-year.22

Baseline descriptive data collected included age, gender, and preoperative Charlson comorbidity index (CCI). Standard radiographic measures of deformity magnitude were measured at a central site with digital imaging software.
Patients completed the Scoliosis Research Society-22r (SRS-22r), Patient Reported Outcomes Information System (PROMIS) – Computer Adaptive Tests (CAT) for Pain Interference (PI) and Physical Function (PF) and the ODI at baseline and one-year followup.

Statistical Methodology

Internal reliability

The internal consistency of a PROM is a measure of its reliability and describes how reliably survey or test items that are designed to measure the same construct actually do so. The reliability (internal consistency) was assessed using Cronbach Alpha. Cronbach’s alpha values >= 0.7 are considered reliable and considered acceptable for group comparisons.

Validation

The criterion validity evaluates how accurately a test measures the outcome it was designed to measure. To explore the validity of the ODI in the context of ASD patients, the concurrent validity (a form of criterion validity) was assessed by calculating the Spearman correlation coefficients for the ODI against corresponding PROMIS-CAT measures (PROMIS-PI and PROMIS-PF) and legacy measures (SRS-Pain and SRS-Activity). These PROMs were selected for analysis as they have been previously validated in the ASD population. Spearman correlations (with 95% confidence intervals) between the ODI and the PROMIS-PI, PROMIS-PF, SRS-22 activity domain, ad SRS-22 pain were calculated at baseline and at 1 year post-operatively. Evans’ criteria were used to evaluate the strength of obtained correlations: “very
weak” (0.00 to 0.19), “weak” (0.20 to 0.39), “moderate” (0.40 to 0.59), “strong” (0.60 to 0.79), and “very strong” (0.80 to 1.00)

**Responsiveness**

Responsiveness reflects the sensitivity of a PROM to detect a change in the outcome of interest. The responsiveness to change of the ODI in patients treated with surgery for ASD was measured with the adjusted effect size (aES) for continuous variables (PROMIS) and η² for ordinal variables (ODI, SRS-22r). Each η² was converted to a Cohen’s d and adjusted to estimate the aES for comparison with PROMIS aES. The adjustment accounts for relatedness between baseline and 1yr values and is necessary when assessing PROM. Larger aES indicate better responsiveness to change, where aES ≥ 1.0 is considered large.

The minimum clinically important difference (MCID) is a threshold of change to mark a clinically relevant change in PROM. This may be viewed as another measure of responsiveness to change. Responder analyses were performed for ODI (MCID=12.8), SRS-Pain (0.4), SRS-Activity (0.6), PROMIS-PI(-5), and PROMIS-PF(4.2) and the rates of improvement compared across these measures.

All statistical analyses were performed using IBM SPSS Statistics v28.0. Statistical significance was defined as p<.05.

**Funding**
The International Spine Study Group (ISSG) is funded through research grants from Medtronic, Globus, Stryker, SI Bone and individual donations.

Results

There were 325 patients enrolled in the registry and 208 (64%) completed baseline PROM for analysis. The majority of patients were females (72%), white (93%), underwent primary ASD surgery (54%) with a median of 11 (IQR 8-14) instrumented levels. Demographic and radiographic data are found in Table 1.

Internal Reliability

Cronbach’s alpha was >0.7 for both baseline (0.89) and 1-year (0.90) followup ODI measures, indicating a high level of internal reliability.

Validation

Spearman rho correlation coefficients found “strong” to “very strong” correlations between contemporaneous ODI and PROMI-PI, -PF, SRS-Activity, and -Pain measurements (Table 2). 95% confidence intervals suggest adequate sample size.

Responsiveness to Change

The ODI exhibited responsiveness to change, as estimated by an adjusted effect size, not different from the legacy measure SRS-22r Activity and Pain domains nor different from PROMIS-PI and -PF estimates. All five instruments showed high adjusted effect sizes (Table 3). Responder analyses were not substantially different across the PROM instruments.
Discussion

To ensure value in ASD care in the face of high costs\(^3\), it is critical that reliable and accurate tools are used to assess the impact of this condition on patient’s quality of life and to determine the effectiveness of the treatments specific to this patient population. Though used almost ubiquitously in ASD research, no previous study has been performed to establish the measurement properties of the ODI in the ASD population\(^1\). As a result, interpretation of these studies has remained a challenge, ultimately contributing to limitations in the ability to develop evidence-based guidelines and appropriately guide clinicians and policy makers regarding the effectiveness of different treatment strategies for ASD.

The present study demonstrates that the ODI is a valid measure of disability as measured by pain and function in patients undergoing surgery for ASD. In addition, the ODI was found to have excellent internal reliability in the assessment of these patients and to be responsive to change in a manner similar to the corresponding domains of the PROMIS-CAT and the SRS-22r legacy measure in this population. These findings support the use of the ODI in the assessment of patients with ASD.

The SRS-22r and PROMIS-PI and -PF domains of the PROMIS-CAT system are valid in the ASD population\(^1\). In our study, we found strong correlations (all \(r\) \(\geq 0.7\)) between the ODI and PROMIS-PI, PROMIS-PF, and SRS-22r domains of Activity and pain both at baseline and 1 year following surgery. Such findings are in agreement with previous work demonstrating similar correlations with the ODI in studies performed to evaluate the validity of the SRS-22\(^1\) and PROMIS-CAT system\(^1\) in ASD patients, and are similar to the results reported for studies exploring the validity of the ODI in a variety of lumbar degenerative conditions\(^2\). These strong correlations indicate that the ODI may assess pain and physical function similarly to these...
instruments in the ASD population both pre-operatively and post-operatively at a time point in which complete recovery is expected. Furthermore, the internal consistency of the ODI was excellent at baseline as well as at 1 year (Cronbach’s alpha: baseline: 0.89, 1 yr= 0.90), suggesting that the items within the defined domains of the ODI correlate strongly with one another in the assessment of disability related to ASD\textsuperscript{24}. Taken together, these findings support the validity and internal reliability of the ODI as a measurement tool to evaluate disability and HRQoL in ASD patients. From a pragmatic view, the ODI may be an easy to score and appropriate single PROM for practices seeking to minimize question burden and examine both pain and disability.

The ODI demonstrated excellent responsiveness to change following surgery with adjusted effect sizes of 2.8. This was greater than the observed responsiveness of the SRS-Activity and PROMIS-PF questionnaires. The PROMIS-PF\textsuperscript{29} and SRS-Activity\textsuperscript{30} item banks have been shown to be unidimensional in nature, meaning that they examine function as an individual trait. HRQoL in ASD patients is multidimensional and the overall disease state and perceived impact on their quality of life is a combination of the various domains known to be impacted in patients suffering from ASD\textsuperscript{14,31}. Additionally, surgery may have contrasting effects on pain and function and there exists a complex interplay of pain, function, and other factors which contribute to a patient’s overall perception of response to these procedures. In contrast to the PROMIS-PF and SRS-Activity measurement tools, the ODI is multidimensional, as it assesses both pain and function simultaneously\textsuperscript{32}. Such an understanding may account for the increased sensitivity of the ODI to change observed in our study. Despite the multidimensionality, the MCID responder analysis suggests that the ODI will identify clinically relevant improvement at a rate similar to the SRS-22r and PROMIS-CAT. Overall, the results of
our study suggest that the ODI can be used to evaluate the response to surgery in a manner like the validated PROMIS-CAT and SRS-22r legacy measures in ASD patients.

Of importance, however, the ODI does not contain a self-image domain, which is known to be an important driver of patients’ perceived HRQoL in spinal deformity. A previous study by Bridwell et al. found the greatest mean improvements in the SRS-22r composite score when compared to the ODI and SF-12 questionnaire 2-years post-operatively, even though the SRS-Activity domain failed to demonstrate a statistically significant improvement after surgery. The success was driven in large part by changes observed in the SRS-Self Image domain. Previous studies of the construct validity of the ODI and SRS-22r in patients with prolonged degenerative thoracolumbar diseases have similarly found that the ODI appears to correlate better with a patient’s perception of their overall physical functional status, while the SRS appears to better capture psychological and emotional components of disability related to their disease. Likely owing to its ease of use, however, the ODI has been shown to be utilized more commonly in the ASD population than the SRS-22r. Given the lack of information pertaining to disability related to self-image in the ODI, future studies exploring the supplementation of the ODI with a validated self-image PROM are needed. As the emphasis on value-based care and economic burden of ASD surgery continue to grow, such information is of paramount importance to enhance ASD-related care and decision making.

There are several limitations to the present study. Several criteria outlined in the COSMIN guidelines were not examined in the present study, most notably related to the construct validity and test-retest reliability. However, the purpose of the present study was to not to complete the examination of all measurement properties of the ODI in ASD patient, but rather to ensure that the ODI demonstrates sufficient validity, reliability, and responsiveness to justify
its current inclusion in the vast majority of ASD research. Additionally, the construct validity of
the ODI in patients with degenerative thoracolumbar disease, including ASD, has been
established previously\textsuperscript{33} and the ODI has demonstrated excellent test-retest reliability in a
multitude of lumbar degenerative diseases and treatments\textsuperscript{17,28}. An additional limitation relates to
the assumption of a linear correlation of the ODI with disability inherent to our response to
change analysis. The data gathered in the ODI is ordinal and is eventually converted to discrete
quantitative data by summing the results of each category. However, the structure of some
subsections of the ODI are not linear, and therefore it remains difficult to determine the degree of
disability attributable to a given effect size observed following surgery using the ODI. We have
treated the ODI as an ordinal value and the use of the aES in our study attempts to control for the
contribution of differences in baseline severity of disability in the observed responsiveness to
treatment. Nonetheless, this highlights the important limitations of the use of ordinal measures in
PROMs to describe disease-related disability in spinal disorders and evaluate the effectiveness of
treatments. Use of continuous measures, as employed in the PROMIS-CAT system, may provide
an improved means for the aggregation of data and improve ease of comparison across studies.
The lack of self-image and mental health data in the ODI also highlights the need for future
studies exploring the most relevant mental health measures for ASD decision making, and their
adoption in PROMs utilized in ASD research. It is important to note, also, that ASD is a disease
composite of malalignment, disc degeneration, and stenosis with central/lateral recess/foraminal
stenosis, all of which cause different pain complaints. The surgeries to treat ASD are as
heterogeneous as the presenting complaints. This makes analyses of the ODI’s sensitivity to
change for primarily leg pain and/or back pain complaints not possible. We have investigated the
dataset and found that median change in ODI scores for these two groups were not different.
Finally, our patient population is primarily comprised of Caucasian females, a problem frequently observed in large cohorts of ASD patients. There are data to support equivalent performance of the ODI across race and gender, however, and we do not think this limitation precludes validation.\textsuperscript{34-38}

In conclusion, we examined 208 ASD patients with one-year follow-up and found the ODI to be a valid measure of pain and disability in ASD patients. To our knowledge, this is the first effort to validate this commonly employed PROM in this patient population. It is important to note that ASD is a multidimensional health issue, affecting self-image as well. The ODI does not measure self-image and may miss some portion of both disease-related disability and treatment-related improvement in this domain.


25. Middel B, van Sonderen E. Statistical significant change versus relevant or important change in (quasi) experimental design: some conceptual and methodological problems in estimating magnitude of intervention-related change in health services research. Int J Integr Care 2002;2:e15.


Table 1. Demographic and Radiographic Data. Continuous variables are presented with mean (standard deviation). Ordinal data are presented with median (interquartile range).

<table>
<thead>
<tr>
<th>Patient Characteristics</th>
<th>N=208</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>60.5 (14.0)</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>149 (72%)</td>
</tr>
<tr>
<td>Height (cm)</td>
<td>164.5 (10.5)</td>
</tr>
<tr>
<td>Weight (kg)</td>
<td>75.3 (18.8)</td>
</tr>
<tr>
<td>Race</td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>193 (93%)</td>
</tr>
<tr>
<td>Black</td>
<td>7 (4%)</td>
</tr>
<tr>
<td>Other/NA</td>
<td>8 (3%)</td>
</tr>
<tr>
<td>Charlson Comorbidity Index</td>
<td>0 (0-2)</td>
</tr>
<tr>
<td>CSHA Clinical Frailty Scale</td>
<td>3 (2-4)</td>
</tr>
<tr>
<td>Edmonton Frailty Score</td>
<td>3 (1-5)</td>
</tr>
<tr>
<td>Instrumented Levels</td>
<td>11 (8-14)</td>
</tr>
<tr>
<td>Revision Surgery</td>
<td>96 (46%)</td>
</tr>
<tr>
<td>Radiographic Parameters</td>
<td></td>
</tr>
<tr>
<td>Maximum Coronal Cobb Angle</td>
<td>38.3° (23.2)</td>
</tr>
<tr>
<td>T1 Pelvic Angle</td>
<td>24.2° (13.9)</td>
</tr>
<tr>
<td>Pelvic Incidence Minus Lumbar Lordosis</td>
<td>16.4° (22.2)</td>
</tr>
<tr>
<td>C2 Tilt</td>
<td>2.7° (6.3)</td>
</tr>
<tr>
<td>Patient Reported Outcome Measures</td>
<td></td>
</tr>
<tr>
<td>----------------------------------</td>
<td>-------</td>
</tr>
<tr>
<td>Oswestry Disability Index</td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>44 (30-56)</td>
</tr>
<tr>
<td>One Year</td>
<td>24 (12-40)</td>
</tr>
<tr>
<td>SRS-22r</td>
<td></td>
</tr>
<tr>
<td>SRS-Activity</td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>2.9 (2.4-3.6)</td>
</tr>
<tr>
<td>One Year</td>
<td>3.6 (3.0-4.2)</td>
</tr>
<tr>
<td>SRS-Pain</td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>2.4 (1.8-3.2)</td>
</tr>
<tr>
<td>One Year</td>
<td>3.6 (2.8-4.2)</td>
</tr>
<tr>
<td>PROMIS</td>
<td></td>
</tr>
<tr>
<td>Physical Function</td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>34.8 (7.2)</td>
</tr>
<tr>
<td>One Year</td>
<td>40.0 (7.6)</td>
</tr>
<tr>
<td>Pain Interference</td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>63.3 (7.7)</td>
</tr>
<tr>
<td>One Year</td>
<td>55.3 (9.0)</td>
</tr>
</tbody>
</table>

CSHA = Canadian Study for Health and Aging, SRS-22r = Scoliosis Research Society-22r, PROMIS = Patient Reported Outcomes Measurement Information System
<table>
<thead>
<tr>
<th></th>
<th>PROMIS-PI</th>
<th>PROMIS-PF</th>
<th>SRS-Activity</th>
<th>SRS-Pain</th>
</tr>
</thead>
<tbody>
<tr>
<td>ODI</td>
<td>Baseline</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.77 (.72; .81)</td>
<td>-.79 (-.83; -.74)</td>
<td>-.81 (-.84; -.76)</td>
<td>-.70 (-.75; -.64)</td>
</tr>
<tr>
<td>ODI</td>
<td>One Year</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.83 (.78; -.86)</td>
<td>-.80 (-.84; -.74)</td>
<td>-.81 (-.85; -.76)</td>
<td>-.80 (-.84; -.74)</td>
</tr>
</tbody>
</table>

Table 2. Spearman rho correlation coefficients between Oswestry Disability Index (ODI) scores and PROMIS-Pain interference (PI), -Physical Function (PF), and Scoliosis Research Society-22r -Activity and -Pain scores at baseline and one-year follow-up.
Table 3. Responsiveness to change at one-year follow-up.

<table>
<thead>
<tr>
<th>Score</th>
<th>Baseline</th>
<th>One-Year</th>
<th>Effect Size</th>
<th>Adjusted Effect Size (Cohen’s d)</th>
<th>Adjusted Effect Size / η²</th>
<th>Achieved Minimum Clinically Important Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>ODI</td>
<td>42 (19)</td>
<td>24 (18.5)</td>
<td>2.2</td>
<td>2.8 / 0.54</td>
<td>130 (40%)</td>
<td></td>
</tr>
<tr>
<td>SRS-Activity</td>
<td>3.0 (0.9)</td>
<td>3.7 (0.8)</td>
<td>1.9</td>
<td>2.4 / 0.48</td>
<td>121 (37%)</td>
<td></td>
</tr>
<tr>
<td>SRS-Pain</td>
<td>2.6 (0.9)</td>
<td>3.6 (0.9)</td>
<td>2.4</td>
<td>3.5 / 0.58</td>
<td>162 (50%)</td>
<td></td>
</tr>
<tr>
<td>PROMIS-PF</td>
<td>35.6 (7.5)</td>
<td>40.9 (8.2)</td>
<td>-.69</td>
<td>1.0</td>
<td>127 (39%)</td>
<td></td>
</tr>
<tr>
<td>PROMIS-PI</td>
<td>62.9 (7.8)</td>
<td>54.1 (9.8)</td>
<td>.90</td>
<td>1.3</td>
<td>143 (44%)</td>
<td></td>
</tr>
</tbody>
</table>

Cohen’s d calculated from the η² for ordinal outcome measures (ODI, SRS-22r). All Cohen’s were adjusted by dividing by the square root of one minus the correlation coefficient between the baseline and one-year PROM value.