Is the Neck Disability Index an Appropriate Measure for Changes in Physical Function After Surgery for Cervical Spondylotic Myelopathy?

DHRUV K.C. GOYAL, MD, HAMADI A. MURPHY, MD, MS, DOUGLAS A. HOLLERN, MD,
SRIKANTH N. DIVI, MD, KRISTEN NICHOLSON, PhD, CHRISTIE STAWICKI, BA, I. DAVID KAYE, MD,
GREGORY D. SCHROEDER, MD, BARRETT I. WOODS, MD, MARK F. KURD, MD, JEFFREY A. RIHN, MD,
D. GREG ANDERSON, MD, CHRISTOPHER K. KEPLER, MD, MBA, ALAN S. HILIBRAND, MD,
ALEXANDER R. VACCARO, MD, PhD, MBA, KRISTEN E. RADCLIFF, MD

Department of Orthopaedic Surgery, Rothman Institute, Thomas Jefferson University, Philadelphia, Pennsylvania

ABSTRACT

Background: The Neck Disability Index (NDI) is a 10-item questionnaire about symptoms relevant to cervical spine pathology, originally validated in the physical therapy literature. It is unclear if all of the items apply to spine surgery. The purpose of this study was to determine if improvements in the composite NDI score or specific NDI domains are appropriate measures for tracking changes in physical function after surgical intervention for cervical spondylotic myelopathy (CSM).

Methods: A retrospective cohort review of patients treated at a major academic medical center was undertaken. Baseline and postoperative standardized outcome measurement scores, including composite NDI, NDI subdomain, and SF-12 physical component score (PCS), were collected. Wilcoxon signed-rank test was used to determine whether patients exhibited improvement in each of the outcome measures included. Multiple linear regression was performed to determine whether change in NDI composite or subdomain scores predicted change in physical function after surgery for CSM—compared with the well-validated PCS score—controlling for factors such as age, sex, etc.

Results: Baseline data were collected on 118 patients. All outcome measures exhibited significant improvement after surgery based on the Wilcoxon signed-rank test. On linear regression, work ($\beta = -2.419$ [-3.831, -1.006]; $P = .001$) and recreation ($\beta = -1.354$ [-2.640, -0.068]; $P = .039$), as well as the NDI composite score ($\beta = -0.223$ [-0.319, -0.127]; $P < .001$), were significant predictors of change in physical function over time.

Conclusions: Although the NDI composite score did predict change in PCS over time, only 2 of the 10 NDI subdomains were found to be associated with change in physical function over time. Based on these results, the item bank and composite scoring of the NDI are inappropriate for evaluating quality of life in studies of surgically treated cervical spondylotic myelopathy patients.

Clinical Relevance: NDI may not be a valid tool in the determination of physical function changes after surgery for CSM.

Level of Evidence: III.

Cervical Spine

Keywords: Neck Disability Index, cervical spondylotic myelopathy, cervical surgery, patient-reported outcomes, Short Form-12 mental component score, Short Form-12 physical component score

INTRODUCTION

Patient-reported outcome measures (PROs) are tools that allow clinicians and researchers to assess patient-reported health status for physical, mental, and social well-being.[1] With the passing of the affordable care act in 2010, there has been a shift from volume-based to value-based care.[2] As a result, it is becoming increasingly important that the most apt PRO measures are identified for the appropriate patient populations.

The Neck Disability Index (NDI), a PRO measure first developed by Vernon in 1991, is a self-assessment measure used to assess disability in patients with neck pain.[3] It was developed using the Oswestry Disability Index, a PRO measure used in patients with low back pain, as a model.[4] Initial validation for the NDI considered only "whiplash"

injury patients treated in an outpatient setting; however, the NDI is now widely used to evaluate the efficacy of cervical surgery.[5–9] The NDI consists of 10 domains—pain intensity, personal care, lifting, reading, headache, concentration, work, driving, sleep, and recreation—designed to assess the level of disability in patients with neck pain. Patients self-report their level of function/disability in each domain, after which a composite score is calculated.

As the first PRO measure specific to patients with neck pain, the NDI is arguably the most widely used PRO for patients with neck disorders.[10] The use of scales, such as the NDI, presents a unique challenge when evaluating patients undergoing treatment for cervical spondylotic myelopathy (CSM) where function, not pain, is the primary symptom. The NDI is still commonly used for these patients despite a lack of validation in this particular population.[6–9] The Short Form-36 (SF-36) and the Short Form-12 (SF-12) are 2 PRO measures that provide a global assessment of health, have strong internal psychometric properties, and have been used and validated in a wide variety of patient populations, including those experiencing CSM.[11–13] SF-36 has previously been used in studies as an anchor for analysis when looking at the NDI and other PRO measures; therefore, in this study we used the abbreviated SF-12 as an anchor in the current analysis.[14,15]

The purpose of this study was to determine whether the NDI is a valid patient-reported outcome measure for tracking changes in physical function after surgical intervention for cervical spondylotic myelopathy.

PATIENTS AND METHODS

Patient Selection

This study was approved by the Institutional Review Board at Thomas Jefferson University Hospital. Each author certifies that his or her institution approved the human protocol for this investigation and that all investigations were conducted in conformity with ethical principles of research.

A retrospective cohort review from a single, large orthopaedic practice with 7 fellowship-trained spine surgeons was conducted on patients treated surgically for CSM between March 2014 and February 2015, using the following CPT codes: Anterior Cervical Decompression and Fusion (22548,

22551, 22552, 22554, 22585), Posterior Cervical Decompression and Fusion (22590, 22595, 22600, 22614), and Cervical Laminoplasty (63050, 63051). Cervical myelopathy was diagnosed based on patient history of dexterity loss, ataxia, and/or bowel and bladder symptoms, along with correlative physical examination and imaging study findings.[16] Patients younger than 18 years, and those who underwent surgery for trauma, tumor, infection, or revision were excluded from the final cohort. Patients with fewer than 11 months of follow-up were also excluded from the analysis.

PRO Measures

The NDI composite score is calculated by equally weighing each of the 10 domains. Each domain is scored from 0 to 5 points, then summed and multiplied by 2 to give a composite score out of 100 points. Larger scores indicate increasing disability, whereas a score of zero indicates no disability.

The SF-12 is a measure of physical and mental health–related quality of life. It is an abbreviated, 12-question version of the SF-36 and is self-administered. The 12 questions belong to 1 of 8 health dimensions: physical function, mental health, role emotional, role physical, social function, bodily pain, vitality, and general health. These dimensions are summarized by 2 components: a physical component score (PCS) and a mental component score (MCS).[13] Within the general US population, both the PCS and MCS were rescaled to each have a mean of 50 and standard deviation of 10 points—higher scores indicate better health-related quality of life.[17,18]

In the present study, quality of life and disability due to neck pain were evaluated using the SF-12 PCS and NDI scores, All PROs were collected in the clinic during postoperative visits or via a web-based application (OBERD, Universal Research Solutions, Columbia, MO).[13,17,19,20] Baseline outcomes were collected preoperatively and again, at least 11 months or later, postoperatively.

Statistical Analysis

To test for differences between baseline (preoperative) and follow-up (postoperative) scores for each domain of the NDI, the composite NDI, and PCS separately, Wilcoxon signed-rank tests were performed. To compare the strength of the relationship of (Δ)PCS to (1) the composite (Δ)NDI

score and (2) the (Δ)NDI domains, 2 separate linear regressions were performed. Both regressions assumed (Δ)PCS was the dependent variable and controlled for age, sex, follow-up (months), surgery type (Anterior Cervical Discectomy and Fusion [ACDF], Posterior Cervical Discectomy and Fusion [PCDF], Anterior/Posterior Cervical Discectomy and Fusion [A/PCDF], laminoplasty). In the first regression, (Δ)NDI was included as an independent variable. In the second regression, all 10 domains of (Δ)NDI were included as independent variables. Backwards stepwise elimination removed nonsignificant (Δ)NDI domains—with a threshold for retention of $P < .15$ in the final model. Model results for (Δ)NDI and retained domains were summarized by β coefficients with 95% confidence intervals. To compare the fit of each regression, the adjusted r^2 was reported. Data were reported by the mean and SD. All statistical analyses were performed using the Statistical Package for the Social Sciences (SPSS) version 24 (IBM Corp, Armonk, NY).[21] Statistical significance was determined as $P < .05$.

RESULTS

Patient Demographics

A total of 118 patients underwent decompression surgery for CSM. A total of 63 patients (53%) were male, and the mean age was 58.1 ± 12.3 years (range, 26–81 years). The following surgical procedures were undertaken: 83 anterior cervical discectomy and fusion, 24 posterior cervical discectomy and fusion, 7 anterior-posterior cervical discectomy and fusion, and 4 laminoplasties. Preoperative and follow-up PRO data were complete for 88 patients. Average follow-up was 21 ± 5 months (range, 11.1–28.8 years). Demographic data for the entire CSM cohort are summarized in Table 1.

NDI Domains, NDI Composite, and PCS Improvement Scores

Overall, patients reported a significant improvement at follow-up in each of the NDI domains ($P < .009$), the total NDI composite ($P < .001$), and PCS ($P \leq .001$) scores. In the final model of the backwards stepwise regression for (Δ)PCS, the only retained (Δ)NDI components were change in work and recreation. The specific parameters for that final model were work (β $= -2.419$ [-3.831, -1.006]; $P = .001$) and recreation (β $= -1.354$ [-2.640,

Table 1. Demographics of cervical spondylotic myelopathy cohort.

	Value
Cohort size	118
Age, y, mean (SD)	58.1 (12.3)
Sex, n (%)	
M	63 (53.4)
F	55 (46.6)
Follow-up, mo, mean (SD)	21.0 (5.0)
Surgery type, n (%)	
ACDF	83 (70.3)
PCDF	24 (20.3)
A/PCDF	7 (5.9)
Laminoplasty	4 (3.5)

Abbreviations: ACDF, Anterior Cervical Discectomy and Fusion; PCDF, Posterior Cervical Discectomy and Fusion; A/PCDF, Anterior/Posterior Cervical Discectomy and Fusion.

-0.068]; $P = .039$; Table 2). The final regression had an adjusted $r^2 = 0.352$. Separately, the composite (Δ)NDI (β $= -0.223$ [-0.319, -0.127]; $P < .001$) was also a significant predictor for (Δ)PCS (Table 2). The adjusted r^2 of the regression model was 0.179. No other variable in either regression model was found to significantly predict (Δ)PCS over time.

DISCUSSION

There are conflicting data on the performance and applicability of the NDI in patients undergoing surgical or nonsurgical treatment for cervical spondylotic myelopathy.[10] CSM, the most common degenerative condition of the spine, tends to present with muscle weakness, gait imbalance, and urogenital dysfunction, among other symptoms.[16] It is a progressive disorder, and although surgery is designed to prevent the progression of the disease, it may not lead to complete resolution of symptoms. Based on this trend, surgery is often preformed when there is compression on the spinal cord, but the symptoms are mild. Furthermore, when patients present late and have severe, longstanding myelopathic symptoms, surgery is unlikely to result in a dramatic improvement of neurologic function. Although the SF-36 and the abbreviated SF-12 show excellent sensitivity to changes in physical and mental health status in patients undergoing surgery for CSM, no validation studies have been done for NDI in patients with CSM despite its widespread use.[13,22] As health care shifts from volume-based to value-based reimbursements, it is critical that appropriate metrics be used to establish the value of care.

The present analysis found all domains of the NDI, NDI composite, and PCS scores to signifi-

Table 2. Preoperative and postoperative Neck Disability Index (NDI) component scores.

	Score, Mean (SD)		Wilcoxon Signed-Rank Test P Value	Multivariable Linear Regression[a]		
	Preoperative	**Postoperative**		**β (95% Confidence Inverval)**	**P Value**	**Adjusted r^2**
NDI domains						0.352
Pain	1.99 (1.32)	1.16 (1.21)	<.001[b]	—	—	
Personal care	1.09 (1.13)	0.54 (1.03)	<.001[b]	—	—	
Lifting	2.65 (1.47)	1.87 (1.66)	<.001[b]	—	—	
Reading	1.49 (1.32)	0.86 (1.05)	<.001[b]	—	—	
Headaches	1.17 (1.46)	0.88 (1.23)	.008[b]	—	—	
Concentration	1.06 (1.09)	0.57 (0.95)	<.001[b]	—	—	
Work	2.42 (1.47)	1.33 (1.59)	<.001[b]	−2.419 (−3.831, −1.006)	.001[b]	
Driving	1.87 (1.53)	1.06 (1.44)	<.001[b]	—	—	
Sleep	2.41 (1.53)	1.28 (1.42)	<.001[b]	—	—	
Recreation	2.59 (1.53)	1.31 (1.50)	<.001[b]	−1.354 (−2.640, −0.068)	.039[b]	
NDI composite	37.56 (19.71)	21.74 (20.65)	<.001[b]	−0.223 (−0.319, −0.127)	<.001[b]	0.179
PCS	32.28 (8.89)	40.45 (10.49)	<.001[b]		c	

Abbreviation: PCS, physical component score.

[a]Regression models controlled for age, sex, follow-up (months), surgery type (Anterior Cervical Discectomy and Fusion [ACDF], Posterior Cervical Discectomy and Fusion [PCDF], Anterior/Posterior Cervical Discectomy and Fusion [A/PCDF], laminoplasty). Work and recreation were the only NDI subdomains retained in the backward stepwise regression.

[b]Statistical significance ($P < .05$).

[c]Delta PCS was the dependent variable in each regression.

cantly improve after surgery. Improvements in the work and recreation domains, as well as the composite NDI score, were found to be the only significant predictors of change in PCS after surgery for CSM. Even though the NDI pain component was not a significant predictor of PCS, on average this domain displayed significant improvement—a finding consistent with prior literature.[23] Most other subdomains followed the pattern of the pain domain by demonstrating improvement but not being significant predictors for (Δ)PCS in the linear regression model. The present study's outcome data are comparable to what has been reported in the literature with respect to NDI and PCS scores. Several studies have reported baseline NDI and PCS scores that were similar to the current study's preoperative values, reporting significant improvement in scores from baseline to 1 to 2 years postoperatively in both NDI and PCS components.[9,15,23]

Existing literature is relatively sparse regarding improvement in work and recreation domains for surgically treated CSM patients. However, multiple studies have found that upper limb function has better recovery after surgery, followed by lower limb function.[24–26] These improvements in limb function directly correlate with the SF-12 PCS score and may also explain why the work and recreation domains have stronger associations with PCS in this study. Here, the work and recreation domains made up 2 of 4 of the highest scores measured at baseline, and these questions—of all questions on the NDI—seem to be most similar to those asked in the SF-12 survey.

If NDI is to be used as an outcome measure in the surgically treated CSM population, then a better than moderate association with PCS in all or most all domains found on NDI is expected. Instead, only 2 domains, work and recreation, were significant predictors of improvement in physical function over time. This suggests that although other domains of the NDI may demonstrate statistical improvements after surgery, they do not necessarily translate to physical function changes. This is possibly related to the fact that patients with CSM exhibit signs of gait impairment—a symptom that is not addressed in any components of the NDI.[16]

This is one of the first studies to directly study the individual NDI domains instead of the aggregate NDI score. There have been several studies questioning the reliability, validity, performance, and applicability of the composite NDI score in patients with cervical radiculopathy and myelopathy.[10,27] Cleland et al[27] found the NDI to have adequate test-retest reliability, but poor construct validity in a population of patients with cervical radiculopathy. The authors found that the NDI was not an effective tool to identify changes in patients' perceived levels of disability when such a change had occurred.[27] In another study, Hung et al[10] found the NDI to have a high level of unexplained variance, which the authors interpreted as "poor unidimensionality"—considering that the NDI is supposed to measure one condition, a lower percentage of unexplained variance was expected.

Additionally, it has been suggested that the presence of psychologic distress may have confounding effects on NDI scores.[28]

If NDI proves to be ineffective in tracking physical function changes in patients receiving surgical intervention to address CSM, there needs to be a set of standardized outcome measures to appropriately assess this particular parameter. Kalsi-Ryan et al[22] claimed that because CSM has such a heterogenous presentation, a single outcome measure is not sufficient to quantify the broad range of neurologic deficits seen in this population. The authors identified additional outcome measures that characterize the deficits in the CSM population with greater validity, reliability, and responsiveness—including QuickDASH, Berg balance scale, GRASSP, grip strength, GAITRite analysis, and the 30-meter walk test—suggesting that with the use of these scales, clinicians would better be able to assess outcomes and monitor the natural course of the disease.[22] Recently, there has been the development of a new type of outcome measure—Patient Reported Outcomes Measurement Information System (PROMIS)—which is being used in the field of orthopaedics.[11] What distinguishes PROMIS outcomes from the traditional PRO measures is use of item response theory.[11] Item response theory uses computerized adaptive testing, which after starting at a predefined point on the scale, chooses each subsequent item based on answers to previous items.[11] Studies have shown that the psychometric characteristics of the PROMIS questionnaire compared favorably with a traditional outcome measure for the upper extremity.[29] Patient time burden was also significantly reduced when using the PROMIS questionnaire compared with the traditional questionnaire.[29] Investigation into how these instruments preform in the spine population is a potential area for future exploration—one that should serve to further elucidate the most appropriate outcome assessment tool in the investigation of physical function changes over time for patients undergoing surgery to prevent the progression of CSM.

Limitations to this study include the fact that data were retrospectively reviewed and that only the PCS portion of the SF-12 survey was used as an anchor. In order to have a more robust analysis, the use of additional anchors—including patient satisfaction with the surgery—would have been beneficial. Additionally, CSM patients often have a psychological component to their disability, so the use of the PCS as an anchor could possibly be confounded by not including a separate analysis of the MCS portion of the SF-12.

CONCLUSION

In summary, most domains of the NDI—including the pain subdomain—did not significantly predict PCS change after surgery for myelopathy. It is important for an outcome measure to consider all aspects of impairment in the CSM population and the full impact on functionality if it is to be used in this set of patients. Based upon these results, the individual item bank and composite scoring of the NDI are inappropriate for evaluating quality of life and physical function changes in patients who are surgically treated to prevent the progressive symptomatology of CSM.

REFERENCES

1. National Quality Forum. Patient Reported Outcomes (PROs) in Performance Measurement. 2013:1–35. http://www.qualityforum.org/Publications/2012/12/Patient-Reported_Outcomes_in_Performance_Measurement.aspx. Accessed December 15, 2018.

2. Patient Protection and Affordable Care Act, 42 U.S.C. § 18001 (2010). n.d.

3. Howard V, Mior S. The Neck Disability Index: a study of reliability and validity. *J Manipulative Physiol Ther*. 1991 Sep;14(7):409–415.

4. Vernon H. The Neck Disability Index: state-of-the-art, 1991-2008. *J Manipulative Physiol Ther*. 2008;31:491–502.

5. Ackelman BH, Lindgren U. Validity and reliability of a modified version of the Neck Disability Index. *J Rehabil Med*. 2002:284–287.

6. Whitmore RG, Ghogawala Z, Petrov D, Schwartz JS, Stein SC. Functional outcome instruments used for cervical spondylotic myelopathy: interscale correlation and prediction of preference-based quality of life. *Spine J*. 2013;13:902–907.

7. Fehlings MG, Santaguida C, Tetreault L, et al. Laminectomy and fusion versus laminoplasty for the treatment of degenerative cervical myelopathy: results from the AOSpine North America and International prospective multice. *Spine J*. 2017;17(1):102–108.

8. Davies BM, McHugh M, Elgheriani A, et al. Reported outcome measures in degenerative cervical myelopathy: a systematic review. *PLoS One*. 2016;11:1–12.

9. Al-Tamimi YZ, Guilfoyle M, Seeley H, Laing RJ. Measurement of long-term outcome in patients with cervical spondylotic myelopathy treated surgically. *Eur Spine J*. 2013;22:2552–2557.

10. Hung M, Cheng C, Hon SD, et al. Challenging the norm: further psychometric investigation of the Neck Disability Index. *Spine J*. 2015;15:2440–2445.

11. Brodke DJ, Hung M, Bozic KJ. Orthopaedic advances item response theory and computerized adaptive testing for

orthopaedic outcomes measures. *J Am Acad Orthop Surg*. 2016;24(11):750–754.

12. King JT, Roberts MS. Validity and reliability of the Short Form-36 in cervical spondylotic myelopathy. *J Neurosurg*. 2002;97:180–185.

13. Singh A, Gnanalingham K, Casey A, Crockard A. Quality of life assessment using the Short Form-12 (SF-12) questionnaire in patients with cervical spondylotic myelopathy: comparison with SF-36. *Spine (Phila Pa 1976)*, 2006;31:639–643.

14. Nilsdotter A, Bremander A. Measures of hip function and symptoms: Harris Hip Score (HHS), Hip Disability and Osteoarthritis Outcome Score (HOOS), Oxford Hip Score (OHS), Lequesne Index of Severity for Osteoarthritis of the Hip (LISOH), and American Academy of Orthopedic Surgeons (AAOS) Hip and Knee Questionnaire. *Arthritis Care Res (Hoboken)*. 2011;63(suppl 11):S200–S20S27.

15. Carreon LY, Glassman SD, Campbell MJ, Anderson PA. Neck Disability Index, short form-36 physical component summary, and pain scales for neck and arm pain: the minimum clinically important difference and substantial clinical benefit after cervical spine fusion. *Spine J*. 2010;10(6):469–474.

16. Lebl DR, Hughes A, Cammisa FP, O'Leary PF. Cervical spondylotic myelopathy: pathophysiology, clinical presentation, and treatment. *HSS J*. 2011;7:170–178.

17. Hanmer J, Lawrence WF, Anderson JP, Kaplan RM, Fryback DG. Report of nationally representative values for the noninstitutionalized US adult population for 7 health-related quality-of-life scores. *Med Decis Making*. 2006;26:391–400.

18. Ware J, Kosinski M, Keller S. A 12-Item short-form health survey: construction of scales and preliminary tests of reliability and validity. *Med Care*. 2016;34:220–233.

19. Smith MJ, Marberry KM. Reliability of a novel, web-based, shoulder-specific, patient-reported outcome instrument. *Curr Orthop Pract*. 2013;24:64–67.

20. Rodts MF, Glanzman R, Gray A, Johnson R, Viellieu D, Hachem F. Measuring outcomes in orthopaedics: implementation of an outcomes program in an outpatient orthopaedic practice. *Orthop Nurs*. 2014;33:331–339.

21. IBM SPSS software. https://www.ibm.com/analytics/spss-statistics-software. Accessed January 10, 2019.

22. Kalsi-Ryan S, Singh A, Massicotte EM, et al. Ancillary outcome measures for assessment of individuals with cervical spondylotic myelopathy. *Spine (Phila Pa 1976)*. 2013;38:S111–S122.

23. Auffinger BM, Lall RR, Dahdaleh NS, et al. Measuring surgical outcomes in cervical spondylotic myelopathy patients undergoing anterior cervical discectomy and fusion: assessment of minimum clinically important difference. *PLoS One*. 2013;8(6):e67408.

24. Houten JK, Cooper PR, Benzel EC, Sonntag VKH, Traynelis VC, Batzdorf U. Laminectomy and posterior cervical plating for multilevel cervical spondylotic myelopathy and ossification of the posterior longitudinal ligament: effects on cervical alignment, spinal cord compression, and neurological outcome. *Neurosurgery*. 2003;52(5):1081–1088.

25. Cheung WY, Arvinte D, Wong YW, Luk KDK, Cheung KMC. Neurological recovery after surgical decompression in patients with cervical spondylotic myelopathy–aprospective study. *Int Orthop*. 2008;32:273–278.

26. Chiles BW, Leonard M, Choudhri H, Cooper P. Cervical spondylotic myelopathy: patterns of neurological deficit and recovery after anterior cervical decompression. *Neurosurgery*. 1999;4:762–769.

27. Cleland JA, Childs JD, Whitman JM. Psychometric properties of the Neck Disability Index and numeric pain rating scale in patients with mechanical neck pain. *Arch Phys Med Rehabil*. 2008;89:69–74.

28. Young SB, Aprill C, Braswell J, Ogard WK, Richards JS, McCarthy JP. Psychological factors and domains of neck pain disability. *Pain Med*. 2009;10:310–18.

29. Tyser AR, Beckmann J, Franklin JD, et al. Evaluation of the PROMIS physical function computer adaptive test in the upper extremity. *J Hand Surg Am*. 2014;39:2047–2051.

Corresponding Author: Dhruv K.C. Goyal, BA, Rothman Orthopaedic Institute, 125 South 9th St, Ste 1000, Philadelphia, PA 19107. Phone: (937) 830-7110; Email: dhruvkcgoyal@gmail.com.