Asian Spine Journal
Clinical Study
Asian Spine J 2023;17(1):86-95 • https://doi.org/10.31616/asj.2021.0414

Validation of Neck Disability Index Severity among Patients Receiving One or Two-Level Anterior Cervical Surgery

Cara E. Geoghegan, Shruthi Mohan, Conor P. Lynch, Elliot D. K. Cha, Kevin C. Jacob, Madhav R. Patel, Michael C. Prabhu, Nisheka N. Vanjani, Hanna Pawlowski, Kern Singh

Department of Orthopaedic Surgery, Rush University Medical Center, Chicago, IL, USA

Study Design: Retrospective cohort.

Purpose: To evaluate the validity of established severity thresholds for Neck Disability Index (NDI) among patients undergoing anterior cervical discectomy and fusion (ACDF) or cervical disc arthroplasty (CDA).

Overview of Literature: Few studies have examined the validity of established NDI threshold values among patients undergoing ACDF or CDA.

Methods: A surgical database was reviewed to identify patients undergoing cervical spine procedures. Demographics, operative characteristics, comorbidities, NDI, Visual Analog Scale (VAS), and 12-item Short Form (SF-12) physical and mental composite scores (PCS and MCS) were recorded. NDI severity was categorized using previously established threshold values. Improvement from preoperative scores at each postoperative timepoint and convergent validity of NDI was evaluated. Discriminant validity of NDI was evaluated against VAS neck and arm and SF-12 PCS and MCS.

Results: All 290 patients included in the study demonstrated significant improvements from baseline values for all patient-reported outcome measures (PROMs) at all postoperative timepoints ($p<0.001$) except SF-12 MCS at 2 years ($p=0.393$). NDI showed a moderate-to-strong correlation ($r\geq0.419$) at most timepoints for VAS neck, VAS arm, SF-12 PCS, and SF-12 MCS ($p<0.001$, all). NDI severity categories demonstrated significant differences in mean VAS neck, VAS arm, SF-12 PCS, and SF-12 MCS at all timepoints ($p<0.001$, all). Differences between NDI severity groups were not uniform for all PROMs. VAS neck values demonstrated significant intergroup differences at most timepoints, whereas SF-12 MCS showed significantly different values between most severity groups.

Conclusions: Neck disability is strongly correlated with neck and arm pain, physical function, and mental health and demonstrates worse outcomes with increasing severity. Previously established severity categories may be more applicable to pain than physical function or mental health and may be more uniformly applied preoperatively for cervical spine patients.

Keywords: Cervical fusion; Neck Disability Index; Validity; Cervical arthroplasty

Received Oct 5, 2021; Revised Dec 27, 2021; Accepted Jan 5, 2022

Corresponding author: Kern Singh
Department of Orthopaedic Surgery, Rush University Medical Center, 1611 W. Harrison St., Suite #300, Chicago, IL 60612, USA
Tel: +1-312-432-2373, Fax: +1-708-409-5179, E-mail: kern.singh@rushortho.com

Asian Spine Journal • pISSN 1976-1902 eISSN 1976-7846 • www.asianspinejournal.org

Introduction

Neck-related pain and disability have a reported lifetime incidence of 67% in the general population [1]. When conservative management proves ineffective, anterior cervical discectomy and fusion (ACDF) or cervical disc arthroplasty (CDA) is used to treat cervical spine degeneration, which is often responsible for symptoms. Many clinicians opt to focus on patient-reported outcome measures (PROMs), which provide an understanding of symptoms from the patient's perspective, to accurately assess recovery from ACDF and CDA. The Neck Disability Index (NDI) is one such PROM that measures the level of disability due to pain among cervical spine patients using a set of 10 different questions [2,3].

Previous studies have verified the NDI questionnaire against other PROMs in ACDF and CDA populations [4,5]. While the NDI score is well validated as an accurate measure of patient symptoms, the literature is less clear about the context or relative severity of any given score. Several previous studies have categorized disability severity using the NDI. Vernon and Mior [6] assessed the reliability of NDI in whiplash patients and first stratified NDI severity as none (0–4), minimal (5–14), moderate (15–24), severe (25–34), and complete (≥35). Other studies on whiplash injury and ACDF patients have also divided baseline NDI scores according to severity; however, the threshold scores used in each study varied greatly [7-9] and the selection criteria for NDI score thresholds and their corresponding severity groups are not well documented.

To our knowledge, no studies have validated the accuracy and reliability of NDI severity categories in ACDF or CDA patients [10]. While previous studies have categorized patients, a single stratification method has yet to be statistically confirmed and established as more than arbitrary groupings of scores in those with degenerative disease of the cervical spine. Given that neck disability severity may play an important role in overall postoperative recovery, validation of specific categories may provide better context for NDI use during preoperative planning and for identification of symptoms that require greater focus in the postoperative period. Therefore, the present study aimed to evaluate the validity of NDI severity categories against various other PROMs among patients undergoing ACDF or CDA procedures.

Materials and Methods

1. Patient population

A retrospective review was performed to identify patients in a prospectively maintained surgical database who underwent a cervical spine procedure from December 2013 to February 2020. Inclusion criteria were primary or revision and one or two-level ACDF or CDA for degenerative spinal pathology. Exclusion criteria were procedures indicated for infectious, malignant, or traumatic etiologies, and patients who had not completed a preoperative NDI survey. All procedures were performed by the same attending spinal surgeon at a single academic institution. Approval from the institutional review board of Rush University Medical Center (ORA #14051301) and written informed patient consent were obtained prior to data collection. collection. All included individuals were prospectively consented before retrospective review was conducted.

2. Data collection

Data regarding demographics, operative characteristics, and comorbidities were collected for all patients included in the present study. Demographics included age, body mass index (BMI), sex, ethnicity, smoking status, American Society of Anesthesiologists (ASA) physical classification, and Charlson comorbidity index (CCI). Medical comorbidities included diabetes mellitus, history of myocardial infarction, hypertension, congestive heart failure, chronic obstructive pulmonary disease, arthritis, liver disease, renal failure, gastrointestinal bleeding, acute immune deficiency syndrome (AIDS), neurological disease, paraplegia, peripheral vascular disease, and metastatic cancer. Operative characteristics encompassed preoperative spinal pathology (herniated nucleus pulposus, central stenosis, foraminal stenosis), neuropathy (radiculopathy, myelopathy, myeloradiculopathy), number of spinal segments operated, operative duration (from skin incision to skin closure), estimated blood loss (EBL), postoperative length of stay, and postoperative day (POD) of discharge.

PROMs were administered preoperatively and at 6 weeks, 12 weeks, 6 months, 1 year, and 2 years postoperatively using a secure online Outcomes Based Electronic Research Database (OBERD, Columbia, MO, USA). PROMs included NDI, Visual Analog Scale (VAS) neck,

VAS arm, 12-item Short Form (SF-12) physical composite score (PCS), and SF-12 mental composite score (MCS).

3. Statistical analysis

Summary statistics were performed for patient demographics, comorbidities, and operative characteristics. Improvement in PROM scores from preoperative baseline values was evaluated at each postoperative timepoint using paired Student *t*-test. Convergent validity of NDI with each other PROM was evaluated at each timepoint using Pearson's correlation coefficient (*r*). Correlation strength was categorized as weak ($r<0.3$), moderate ($0.3\leq r<0.5$), or strong ($r\geq0.5$). Patients were categorized according to their level of disability at each timepoint using the following previously established threshold values for NDI: none (0–4); minimal (5–14); moderate (15–24); severe (25–34); and complete (≥35) [6]. NDI severity was categorized separately in this manner for each timepoint. Discriminant validity of NDI severity categories was evaluated using a one-way analysis of variance with post hoc Tukey testing to compare scores in all other PROMs (VAS neck, VAS arm, SF-12 PCS, and SF-12 MCS) among NDI categories. Differences in mean PROM scores between patients in each group were reported at each timepoint. An alpha value of 0.050 was set as the threshold for statistical significance in all tests. All statistical calculations were performed using Stata ver. 16.0 (Stata Corp., College Station, TX, USA).

Results

A total of 290 patients met the inclusion/exclusion criteria. The patients had a mean age of 48.7 years, mean BMI of 29.5 kg/m^2, and 60.0% were male. Most patients were of Caucasian ethnicity (77.5%), nonsmokers (85.9%), and had an ASA classification of ≥2 (79.1%). The mean CCI was 0.62 and the most common medical comorbidities were hypertension (23.1%), arthritis (9.7%), and diabetes mellitus (8.3%). None of the patients had a diagnosis of AIDS, neurological disease, paraplegia, peripheral vascular disease, or metastatic cancer (Table 1). Most patients had a preoperative spinal pathology of herniated nucleus pulposus (81.7%), experienced symptoms of myeloradiculopathy (85.2%), and had a single level surgery (64.1%). The mean operative duration was 57.9 minutes, mean EBL was 30.3 mL, and mean postoperative length of stay was 11.5 hours.

Table 1. Patient demographics (N=290)

Characteristic	Value
Age (yr)	48.7±9.5
Body mass index (kg/m^2)	29.5±5.7
Gender	
Female	40.0 (116)
Male	60.0 (174)
Ethnicity	
Caucasian	77.5 (224)
African American	9.3 (27)
Hispanic	9.0 (26)
Asian	2.4 (7)
Other	1.7 (5)
Diabetic status	
Non-diabetic	91.7 (266)
Diabetic	8.3 (24)
Smoking status	
Non-smoker	85.9 (249)
Smoker	14.1 (41)
American Society of Anesthesiologists classification	
<2	20.9 (53)
≥2	79.1 (237)
Charlson comorbidity index score	0.62±0.48
Medical comorbidity[a]	
Myocardial infarction	0.7 (2)
Hypertension	23.1 (67)
Congestive heart failure	0.7 (2)
Chronic obstructive pulmonary disease	1.7 (5)
Arthritis	9.7 (28)
Liver disease	0.3 (1)
Renal failure	0.3 (1)
Gastrointestinal bleed	0.3 (1)

Values are presented as mean±standard deviation or % (number).
[a]No patients had a past medical history of acquired immunodeficiency syndrome, neurological disease, paraplegia, peripheral vascular disease, or cancer metastasis

Most patients were discharged on POD 0 (Table 2).

NDI, VAS neck, VAS arm, and SF-12 PCS were significantly improved at all postoperative timepoints over 2 years ($p<0.001$, all). SF-12 MCS significantly improved over 1 year ($p\leq0.039$, all). VAS neck and SF-12 PCS demonstrated statistically significant, strong correlations with NDI at all timepoints ($p<0.001$, all). VAS arm demonstrated strong correlations with NDI preoperatively and at 12

Table 2. Perioperative characteristics (N=290)

Characteristic	Value
Spinal pathology	
Herniated nucleus pulposus	81.7 (237)
Central stenosis	58.6 (170)
Foraminal stenosis	10.6 (31)
Neuropathy	
Radiculopathy	11.4 (33)
Myelopathy	1.7 (5)
Myeloradiculopathy	85.2 (247)
No. of operative levels	
1 Level	64.1 (186)
2 Levels	35.9 (104)
Operative time (min)	57.9±16.0
Estimated blood loss (mL)	30.3±13.8
Length of stay (hr)	11.5±8.9
Day of discharge	
POD 0	75.5 (219)
POD 1	24.5 (71)

Values are presented as % (number) or mean±standard deviation.
POD, postoperative day.

weeks, 6 months, and 2 years postoperatively and showed moderate correlations at 6 weeks and 1 year (p<0.001, all). SF-12 MCS demonstrated strong correlations with NDI from 6 weeks to 2 years and moderate correlations at the preoperative timepoint (p<0.001, all) (Table 3). The mean scores for VAS neck, VAS arm, SF-12 PCS, and SF-12 MCS differed significantly among NDI severity groups at all timepoints (p≤0.009, all). No patients were categorized as having "complete" disability at 2 years (Table 4). Between-group comparisons of NDI categories revealed no significant differences in VAS neck scores between the "none" and "mild" groups preoperatively (1.5±0.70) and 2 years postoperatively (1.4±0.76), "moderative" and "severe" groups at 2 years (1.6±0.86), "moderate" and "complete" groups at 2 years (1.6±0.86), and "complete" and "severe" groups at all timepoints. VAS arm scores showed no significant differences between the "none" and "mild" groups preoperatively and at 6 weeks, 12 weeks, and 2 years, "none" and "moderate" groups at 6 weeks, "mild" and "moderate" groups at 6 weeks, 1 year, and 2 years, "mild" and "severe" groups at 1 year, "mild" and "complete" group at 1 year, "moderate" and "severe" groups at 12 weeks to 1 year, "moderate" and "complete" groups at 1 year, and "severe" and "complete" groups at all timepoints. VAS neck and VAS arm scores showed significant differences for all other between-group comparisons (Table 5). SF-12 PCS scores showed no significant differences between the "none" and "mild" groups at 2 years, "mild" and "moderate" groups at 2 years, "mild" and "severe" groups at 6 months, "mild" and "complete" groups at 6 months, "moderate" and "severe" groups from 6 weeks to 2 years, "moderate" and "complete" groups at 6 weeks to 2 years, and "severe" and "complete" groups at all timepoints. SF-12 MCS scores showed no significant differences between the "none" and "mild" groups at any timepoint, "none" and "moderate" groups preoperatively and at 2 years, "mild" and "moderate" groups preoperatively and at 2 years, "moderate" and "severe" groups at 6 months and 2 years, "moderate" and "complete" groups at 1 year, and "severe" and "complete" group at all timepoints. SF-12 PCS and SF-12 MCS scores significantly differed between NDI severity groups at all timepoints (Table 6).

Discussion

ACDF and CDA are established treatment options that show favorable results for the treatment of various cervical spine pathologies [11]. The increased use of PROMs to assess patients' perception of their health has led to NDI becoming the most widely used tool to measure an individual's associated neck disability [3]. Although NDI is a strongly validated measurement in the spine, the severity thresholds used to stratify patients based on their disability are not well established. Therefore, the present study evaluated the validity of previously established NDI severity thresholds against other well-established PROMs in patients undergoing ACDF or CDA.

NDI has an established correlation with various PROMs, making it a useful tool for understanding patients' disabilities and how they may affect other aspects of their health and quality of life. The present study showed that NDI demonstrated a moderate-to-strong correlation at most timepoints for VAS neck, VAS arm, SF-12 PCS, and SF-12 MCS. Steinhaus et al. reported a significant correlation between NDI and both Patient-Recorded Outcomes Measurement Information System physical function and SF-36 PCS in adults undergoing cervical spinal surgery [12]. Additional studies have shown NDI to be correlated with VAS, SF-36 MCS, and Hospital Anxiety and Depression Scale among a variety of other PROMs

Table 3. Patient-reported outcomes

PROM	Score	No. of patients	*p*-value[a]	Pearson coefficient (|*r*|)	Strength	*p*-value[b]
NDI						
Preop	41.3±18.9	290	-	-	-	-
6-Weeks	30.9±19.4	248	**<0.001**	-	-	-
12-Weeks	26.9±19.8	234	**<0.001**	-	-	-
6-Months	24.6±19.9	189	**<0.001**	-	-	-
1-Year	23.6±20.6	105	**<0.001**	-	-	-
2-Years	27.8±21.1	44	**<0.001**	-	-	-
VAS neck						
Preop	6.2±2.4	289	-	0.645	Strong	**<0.001**
6-Weeks	3.3±2.5	247	**<0.001**	0.766	Strong	**<0.001**
12-Weeks	2.9±2.4	236	**<0.001**	0.808	Strong	**<0.001**
6-Months	2.9±2.5	191	**<0.001**	0.786	Strong	**<0.001**
1-Year	3.2±2.7	105	**<0.001**	0.743	Strong	**<0.001**
2-Years	3.9±2.5	44	**<0.001**	0.755	Strong	**<0.001**
VAS arm						
Preop	5.9±2.6	289	-	0.533	Strong	**<0.001**
6-Weeks	2.8±3.0	249	**<0.001**	0.433	Moderate	**<0.001**
12-Weeks	2.7±2.9	234	**<0.001**	0.525	Strong	**<0.001**
6-Months	2.9±2.7	189	**<0.001**	0.633	Strong	**<0.001**
1-Year	3.3±3.0	104	**<0.001**	0.419	Moderate	**<0.001**
2-Years	5.7±10.2	44	**<0.001**	0.754	Strong	**<0.001**
SF-12 PCS						
Preop	34.1±8.3	251	-	0.557	Strong	**<0.001**
6-Weeks	36.5±9.3	173	**<0.001**	0.673	Strong	**<0.001**
12-Weeks	39.7±10.3	146	**<0.001**	0.719	Strong	**<0.001**
6-Months	41.5±10.6	127	**<0.001**	0.640	Strong	**<0.001**
1-Year	42.1±10.7	89	**<0.001**	0.742	Strong	**<0.001**
2-Years	40.7±11.5	60	**<0.001**	0.709	Strong	**<0.001**
SF-12 MCS						
Preop	45.9±12.7	251	-	0.439	Moderate	**<0.001**
6-Weeks	50.7±11.3	173	**<0.001**	0.663	Strong	**<0.001**
12-Weeks	49.4±11.8	146	**<0.001**	0.588	Strong	**<0.001**
6-Months	50.8±11.2	127	**<0.001**	0.597	Strong	**<0.001**
1-Year	50.0±11.8	89	**0.039**	0.751	Strong	**<0.001**
2-Years	47.9±11.2	60	**0.393**	0.515	Strong	**<0.001**

Values are presented as mean±standard deviation or number, unless otherwise stated. Boldface indicates statistical significance.
PROM, patient-reported outcome measure; NDI, Neck Disability Index; Preop, preoperative; VAS, Visual Analog Scale; SF-12, 12-item Short Form; PCS, physical composite score; MCS, mental composite score.
[a]Calculated using paired *t*-test to determine differences from preoperative values. [b]Calculated using Pearson correlation to determine relationship with NDI.

[3,13]. The present study provided further context for the NDI questionnaire by investigating the validity of the severity thresholds to understand both the limitations and capabilities of NDI as a meaningful clinical tool.

The use of NDI severity thresholds may help to quantify patients' perceived disability by providing additional context

Table 4. Mean PROM scores by NDI severity

PROM	None	Mild	Moderate	Severe	Complete	*p*-value[a]
VAS neck						
Preop	2.8±2.6 (9)	4.4±2.4 (62)	6.1±1.8 (116)	7.5±1.7 (84)	8.9±1.5 (18)	**<0.001**
6-Weeks	0.58±0.87 (31)	2.1±1.6 (100)	4.1±1.9 (65)	6.2±1.7 (43)	7.9±1.5 (7)	**<0.001**
12-Weeks	0.43±0.82 (51)	2.2±1.6 (82)	4.2±1.7 (68)	6.1±1.7 (25)	6.9±2.4 (8)	**<0.001**
6-Months	0.49±0.78 (49)	2.5±1.7 (72)	4.2±1.9 (48)	6.3±2.2 (12)	7.6±1.8 (8)	**<0.001**
1-Year	1.2±2.4 (33)	2.9±1.6 (39)	4.6±2.1 (18)	6.7±1.4 (13)	8.4±2.3 (2)	**<0.001**
2-Years	1.8±1.9 (12)	3.2±1.5 (13)	5.0±2.4 (11)	6.6±1.5 (8)	-	**<0.001**
VAS arm						
Preop	2.8±2.6 (9)	4.5±2.5 (62)	5.7±2.5 (116)	7.1±2.0 (84)	8.7±1.6 (18)	**<0.001**
6-Weeks	1.7±4.6 (32)	1.8±2.2 (1000)	2.9±2.4 (66)	5.1±2.3 (43)	6.3±2.6 (7)	**<0.001**
12-Weeks	1.3±2.7 (51)	2.0±2.4 (82)	3.4±2.4 (68)	4.9±2.8 (25)	6.7±3.5 (8)	**<0.001**
6-Months	0.85±2.1 (49)	2.6±2.2 (72)	3.9±2.5 (48)	5.6±2.8 (12)	7.4±1.4 (8)	**<0.001**
1-Year	1.7±3.3 (33)	3.7±2.5 (38)	4.2±2.4 (18)	5.0±2.9 (13)	7.4±2.9 (2)	**<0.001**
2-Years	0.83±1.5 (9)	2.7±2.5 (13)	3.9±2.4 (11)	7.1±1.5 (8)	-	**<0.001**
SF-12 PCS						
Preop	48.1±10.4 (7)	39.2±8.1 (53)	34.7±7.3 (98)	29.7±5.6 (77)	27.3±6.7 (16)	**<0.001**
6-Weeks	47.9±7.4 (23)	39.1±8.4 (76)	32.7±5.5 (47)	28.4±3.8 (29)	27.0±5.5 (5)	**<0.001**
12-Weeks	52.1±7.3 (33)	41.1±8.5 (55)	34.1±6.5 (47)	29.3±4.8 (13)	28.4±5.3 (5)	**<0.001**
6-Months	51.9±5.1 (37)	41.9±8.6 (50)	32.9±3.1 (28)	32.9±3.1 (5)	34.4±3.9 (4)	**<0.001**
1-Year	50.3±7.9 (27)	43.6±7.2 (32)	30.9±8.2 (15)	27.5±4.4 (9)	25.8±4.9 (2)	**<0.001**
2-Years	49.8±11.9 (7)	41.5±8.5 (12)	35.5±7.2 (9)	25.6±6.1 (6)	-	**<0.001**
SF-12 MCS						
Preop	55.7±4.5 (7)	51.5±10.1 (53)	48.6±11.6 (98)	39.9±12.5 (77)	35.7±11.6 (16)	**<0.001**
6-Weeks	58.5±5.3 (23)	55.3±8.0 (76)	49.8±9.4 (47)	38.5±11.2 (29)	30.6±10.4 (5)	**<0.001**
12-Weeks	55.1±7.2 (33)	53.3±9.3 (55)	46.9±10.8 (47)	35.1±6.4 (13)	26.8±6.1 (5)	**<0.001**
6-Months	56.2±5.9 (37)	54.2±10.2 (50)	45.7±9.6 (28)	36.2±7.7 (5)	28.4±8.9 (4)	**<0.001**
1-Year	57.6±5.2 (27)	52.7±9.6 (32)	42.1±11.7 (15)	30.9±8.4 (9)	28.4±10.7 (2)	**<0.001**
2-Years	53.3±8.4 (7)	51.8±8.5 (12)	43.6±12.1 (9)	36.5±10.2 (6)	-	**0.009**

Values are presented as mean±standard deviation (number of patients). Boldface indicates significance.
PROM, patient-reported outcome measure; NDI, Neck Disability Index; Preop, preoperative; VAS, Visual Analog Scale; SF-12, 12-item Short Form; PCS, physical composite score; MCS, mental composite score.
[a]Calculated using one-way analysis of variance.

for scores. There are several proposed scoring breakdowns in the published literature [10] and the present study used the thresholds established by Vernon and Mior [6], the original creators of the NDI tool, as follows: none (0–4); minimal (5–14); moderate (15–24); severe (25–34); and complete (≥35). However, the suggested thresholds have yet to be validated against other PROMs used in the cervical spine. McCarthy et al. [14] used a similar scoring system to validate NDI against SF-36 and reported percentage scores of 0%–20% (normal), 21%–40% (mild disability), 41%–60% (moderate), 61%–80% severe, and ≥80 (complete). In the present study, the NDI severity categories showed significant differences in mean VAS neck, VAS arm, SF-12 PCS, and SF-12 MCS at all timepoints. Furthermore, individuals categorized as complete disability reported higher VAS pain scores and lower mental health and physical function scores than those in the none and mild categories. This suggests that the threshold values originally proposed by Vernon and Mior [6] are not completely arbitrary and may help to improve preoperative education and understanding

Table 5. Discriminant validity of NDI vs. VAS

Variable	Preoperative	6-Weeks	12-Weeks	6-Months	1-Year	2-Years
VAS neck						
None vs. mild	1.5±0.70	**1.5±0.34**	**1.7±0.27**	**1.9±0.30**	**1.7±0.47**	1.4±0.76
None vs. moderate	**3.4±0.68**	**3.5±0.36**	**3.7±0.28**	**3.7±0.33**	**3.4±5.8**	**3.2±0.80**
None vs. severe	**4.7±0.69**	**5.6±0.39**	**5.7±0.37**	**5.8±0.52**	**5.5±0.65**	**4.8±0.87**
None vs. complete	**6.1±0.80**	**7.2±0.69**	**6.5±0.58**	**7.1±0.62**	**7.2±1.4**	-
Mild vs. moderate	**1.8±0.31**	**1.9±0.26**	**2.0±0.25**	**1.8±0.30**	**1.6±0.56**	**1.8±0.78**
Mild vs. severe	**3.1±0.33**	**4.0±0.30**	**3.9±0.35**	**3.8±5.0**	**3.7±0.63**	**3.4±0.86**
Mild vs. complete	**4.6±0.52**	**5.7±0.65**	**4.7±0.57**	**5.2±0.61**	**5.4±1.4**	-
Severe vs. moderate	**1.3±0.28**	**2.1±0.32**	**1.9±0.36**	**2.0±0.53**	**2.1±0.72**	1.6±0.86
Complete vs. moderate	**2.7±0.49**	**3.8±0.65**	**2.7±0.57**	**3.4±0.62**	3.8±1.5	-
Complete vs. severe	1.4±0.51	1.7±0.67	0.82±0.62	1.4±0.74	1.7±1.5	-
VAS arm						
None vs. mild	1.6±0.82	0.18±0.54	0.67±0.45	**1.7±0.41**	**2.0±0.66**	1.8±0.91
None vs. moderate	**2.9±0.80**	1.3±0.58	**2.1±0.47**	**3.1±0.45**	**2.5±0.82**	**3.0±0.95**
None vs. severe	**0.43±0.81**	**3.3±0.63**	**3.6±0.60**	**4.7±0.72**	**3.3±0.91**	**6.3±1.0**
None vs. complete	**5.9±0.94**	**4.7±1.1**	**5.4±0.96**	**6.6±0.85**	**5.7±2.0**	-
Mild vs. moderate	**1.2±0.36**	1.1±0.42	**1.4±0.42**	**1.3±0.42**	0.52±0.80	1.2±0.86
Mild vs. severe	**2.6±0.38**	**3.2±0.49**	**2.9±0.57**	**2.9±0.70**	1.3±0.89	**4.4±0.94**
Mild vs. complete	**4.3±0.62**	**4.5±1.1**	**4.7±0.94**	**4.8±0.84**	3.7±2.0	-
Severe vs. moderate	**1.4±0.33**	**2.1±0.53**	1.6±0.59	1.7±0.72	0.81±1.0	3.2±0.97
Complete vs. moderate	**3.1±0.58**	**3.4±1.1**	**3.4±0.94**	**3.5±0.85**	3.2±2.1	-
Complete vs. severe	1.6±0.60	1.3±1.1	1.8±1.0	1.8±1.0	2.4±2.1	-

Values are presented as Δmean±standard error. All values were calculated using post-hoc Tukey test. Boldface indicates statistical significance.
NDI, Neck Disability Index; VAS, Visual Analog Scale.

of postoperative recovery in patients.

The direct assessment of discriminant validity in the present study enabled a more direct comparison between different levels of severity and an understanding of whether these scoring thresholds could be used to correctly distinguish between patients with different levels of disability. The results of the present study demonstrate that scores are not as well differentiated at the extremes of the severity spectrum (i.e., none versus mild, severe versus complete). It is important to note that the lack of differentiation at the extremes may be partly due to the patient population included in the present study. Patients undergoing procedures for degenerative cervical pathologies tend to experience at least some disability and it is likely that this population did not include a significant number of patients with severe spinal cord damage.

Our results revealed that differences between NDI severity groups were not uniform for all PROMs. Assessment of discriminant validity yielded much more consistent results for VAS neck and arm than SF-12 MCS or SF-12 PCS, which could be due to the nature of the questionnaires. While SF-12 is a general quality of life measurement intended for use in a variety of patient populations, VAS may capture symptoms related to the individual's cervical pathology more effectively by only focusing on questions related to neck and arm pain. It is important to note the differences between SF-12 MCS and SF-12 PCS. While SF-12 PCS demonstrated significant results at all but one preoperative timepoint, SF-12 MCS showed no significant results between the none and mild, none and moderate, mild and moderate, and complete and severe groups preoperatively. This may be due to a stronger relationship between disability and pain or physical function compared with mental health. These findings are supported by those reported Stull et al. [15], who found that individuals with significant preoperative neck

Table 6. Discriminant validity of NDI vs. SF-12

Variable	Preop	6-Weeks	12-Weeks	6-Months	1-Year	2-Years
SF-12 PCS						
None vs. mild	**8.9±2.8**	**8.6±1.6**	**11.1±1.6**	**10.0±1.6**	**6.7±1.9**	8.4±4.1
None vs. moderate	**13.3±2.8**	**151±1.8**	**18.0±1.7**	**20.2±1.8**	**19.4±2.4**	**14.4±4.4**
None vs. severe	**18.3±2.8**	**19.5±1.9**	**22.8±2.4**	**18.9±3.5**	**22.8±2.9**	**24.2±4.8**
None vs. complete	**20.8±3.2**	**20.7±3.4**	**23.7±3.5**	**17.4±3.8**	**24.4±5.4**	-
Mild vs. moderate	**4.5±1.2**	**6.4±1.3**	**6.9±1.5**	**10.1±1.7**	**12.7±2.3**	6.0±3.8
Mild vs. severe	**9.5±1.2**	**10.8±1.5**	**11.8±2.3**	8.9±3.4	**16.2±2.8**	**15.9±4.3**
Mild vs. complete	**11.9±2.0**	**12.1±3.2**	**12.6±3.4**	7.4±3.8	**17.7±5.4**	-
Severe vs. moderate	**5.0±1.1**	4.4±1.6	4.8±2.3	1.2±3.5	3.5±3.1	9.8±4.6
Complete vs. moderate	**7.4±1.9**	5.7±3.3	5.7±3.4	2.7±3.9	5.1±5.6	-
Complete vs. severe	2.4±1.9	1.3±3.3	0.87±3.8	1.5±4.9	1.6±5.8	-
SF-12 MCS						
None vs. mild	4.3±4.6	3.2±2.1	1.7±2.0	1.9±1.9	4.9±2.2	1.5±4.7
None vs. moderate	7.1±4.5	**8.7±2.2**	**8.1±2.1**	**10.5±2.2**	**15.5±2.8**	9.7±4.9
None vs. severe	**15.8±4.5**	**20.0±2.4**	**19.9±3.0**	**19.9±4.2**	**26.8±3.4**	**16.9±5.5**
None vs. complete	**50.1±5.2**	**27.9±4.3**	**28.2±4.4**	**27.7±4.7**	**29.2±6.4**	-
Mild vs. moderate	2.8±1.9	**5.5±1.6**	**6.3±1.8**	**8.5±2.1**	**10.6±2.7**	8.2±4.4
Mild vs. severe	**11.5±2.0**	**16.8±1.9**	**18.1±2.8**	**17.9±4.1**	**21.9±3.3**	**15.4±4.9**
Mild vs. complete	**15.7±3.3**	**24.6±4.0**	**26.5±4.3**	**25.7±4.6**	**24.3±6.4**	-
Severe vs. moderate	**8.7±1.7**	**11.3±2.1**	**11.7±2.9**	9.5±4.3	**11.3±3.7**	7.1±5.2
Complete vs. moderate	**12.9±3.1**	**19.2±4.1**	**20.1±4.3**	**17.2±4.7**	13.7±6.6	-
Complete vs. severe	4.2±3.2	7.9±4.2	8.3±4.8	7.7±5.9	2.4±6.9	-

Values are presented as Δmean±standard error. All values were calculated using post-hoc Tukey test. Boldface indicates statistical significance.
NDI, Neck Disability Index; SF-12, 12-item Short Form; Preop, preoperative; PCS, physical composite score; MCS, mental composite score.

pain experienced significant improvement in neck pain, disability, and physical function following ACDF, whereas only a subset of their patient population experienced significant improvements in mental health, as measured by SF-12 MCS.

A comparison of the discriminant validity among the different PROMs showed that VAS neck yielded the most consistent differences among severity groups, even compared with VAS arm. This suggests that NDI may be a better representative of neck pain than arm pain, which is an important relationship to acknowledge when using the metric to assess a patient's symptomatology. Narain et al. [16] examined a cohort of patients undergoing ACDF and showed minimum clinically important difference achievement rates for VAS neck, VAS arm, and NDI of 55.4%, 36.9%, and 76.6%, respectively. This further highlights that while patients may recover well in terms of their neck pain and disability, it is more difficult for them to achieve clinically significant improvement in arm pain.

The present study revealed that PROM scores are not as well differentiated among severity groups at later time points. For example, there were no significant differences between the mild and complete groups at 6 months for SF-12 PCS or at 1 year for VAS arm. Additionally, while VAS neck demonstrated significant differences at all previous timepoints between the complete and moderate groups, there was not significant differentiation at 1 year. However, this may be because patients recover over time after their procedure. Andresen et al. found that patients undergoing ACDF demonstrated significant improvement in disability following surgery, with a mean preoperative NDI score of 40.0, which improved to 22.7 in the postoperative period [17]. Although disability likely improves postoperatively, discriminant analysis would still be expected to yield significant results if pain, physical function, and mental health also improved. However, it is

possible that although a patient's disability may drastically improve following surgery, there may be additional factors unrelated to disability that contribute to their continued pain and limited physical function or poor mental health. This suggests that NDI may be a useful tool for preoperative stratification, but as patients progress in their postoperative recovery the metric may become less useful for differentiating severity of disability. Additionally, the extent of recovery following surgery may limit the number of individuals experiencing severe or complete disability at later timepoints, making any comparison among differing severity groups more difficult.

The present study has some limitations that should be considered. First, the use of PROMs may introduce response or recall bias, as occurs with all self-reported questionnaires used to assess outcomes. The study included outcomes up to 2 years; therefore, it was limited by the number of patients that completed the questionnaires throughout the entire follow-up period. Furthermore, the use of a patient cohort that underwent surgery performed by a single surgeon at a single institution, as well as the limited number of individuals at the extremes of disability, may limit the generalizability of our results to other populations. Additional studies that include a multicentered, multiple surgeon cohort will strengthen the results of the present study. Finally, the included patient cohort specifically underwent anterior cervical operation at one or two levels; therefore, our findings may not be generalizable to individuals undergoing surgery at three or more levels, myelopathy patients receiving posterior cervical operation, particular revision procedures, and those receiving surgery for deformity, malignancy, trauma, or infectious etiology.

Conclusions

Neck disability, as measured by NDI, is strongly correlated with neck and arm pain (VAS), physical function (SF-12 PCS), and mental health (SF-12 MCS) and demonstrates generally worse outcomes with increasing severity. However, differences between groups were not uniform for all PROMs. Specifically, severity categories were not as applicable at extremes or longer postoperative timepoints and demonstrated more consistent results for VAS than SF-12. This suggests that NDI and its established severity thresholds may be more useful as a preoperative stratification tool than a way to assess disability during postoperative recovery. Further studies are required to develop a more clinically relevant severity threshold for patients undergoing cervical spinal surgery.

Conflict of Interest

No potential conflict of interest relevant to this article was reported.

Author Contributions

Conception and design: KS; data acquisition: KS; analysis of data: CEG, SM, CPL, EDKC, KCJ, MRP; drafting manuscript: CEG, SM, CPL, EDKC; administrative support: CEG, SM, MCP, MNV, HP; supervision: KS; and critical revision: all authors.

References

1. Todd AG. Cervical spine: degenerative conditions. Curr Rev Musculoskelet Med 2011;4:168-74.
2. Young IA, Cleland JA, Michener LA, Brown C. Reliability, construct validity, and responsiveness of the neck disability index, patient-specific functional scale, and numeric pain rating scale in patients with cervical radiculopathy. Am J Phys Med Rehabil 2010;89:831-9.
3. Vernon H. The Neck Disability Index: state-of-the-art, 1991-2008. J Manipulative Physiol Ther 2008;31:491-502.
4. Khechen B, Patel DV, Haws BE, et al. Evaluating the concurrent validity of PROMIS physical function in anterior cervical discectomy and fusion. Clin Spine Surg 2019;32:449-53.
5. Chang HK, Huang WC, Tu TH, et al. Radiological and clinical outcomes of 3-level cervical disc arthroplasty. J Neurosurg Spine 2019;32:174-81.
6. Vernon H, Mior S. The Neck Disability Index: a study of reliability and validity. J Manipulative Physiol Ther 1991;14:409-15.
7. Miettinen T, Leino E, Airaksinen O, Lindgren KA. The possibility to use simple validated questionnaires to predict long-term health problems after whiplash injury. Spine (Phila Pa 1976) 2004;29:E47-51.
8. Sterling M, Jull G, Kenardy J. Physical and psychological factors maintain long-term predictive capacity post-whiplash injury. Pain 2006;122:102-8.

9. Nandoe Tewarie RD, Bartels RH, Peul WC. Long-term outcome after anterior cervical discectomy without fusion. Eur Spine J 2007;16:1411-6.
10. MacDermid JC, Walton DM, Avery S, et al. Measurement properties of the neck disability index: a systematic review. J Orthop Sports Phys Ther 2009;39:400-17.
11. Findlay C, Ayis S, Demetriades AK. Total disc replacement versus anterior cervical discectomy and fusion: a systematic review with meta-analysis of data from a total of 3160 patients across 14 randomized controlled trials with both short- and medium- to long-term outcomes. Bone Joint J 2018;100-B:991-1001.
12. Steinhaus ME, Iyer S, Lovecchio F, et al. Which NDI domains best predict change in physical function in patients undergoing cervical spine surgery? Spine J 2019;19:1698-705.
13. Sundseth J, Kolstad F, Johnsen LG, et al. The Neck Disability Index (NDI) and its correlation with quality of life and mental health measures among patients with single-level cervical disc disease scheduled for surgery. Acta Neurochir (Wien) 2015;157:1807-12.
14. McCarthy MJ, Grevitt MP, Silcocks P, Hobbs G. The reliability of the Vernon and Mior neck disability index, and its validity compared with the short form-36 health survey questionnaire. Eur Spine J 2007;16:2111-7.
15. Stull JD, Goyal DK, Mangan JJ, et al. The outcomes of patients with neck pain following ACDF: a comparison of patients with radiculopathy, myelopathy, or mixed symptomatology. Spine (Phila Pa 1976) 2020;45:1485-90.
16. Narain AS, Hijji FY, Khechen B, et al. Risk factors associated with failure to reach minimal clinically important difference in patient-reported outcomes following anterior cervical discectomy and fusion. Int J Spine Surg 2019;13:262-9.
17. Andresen AK, Paulsen RT, Busch F, Isenberg-Jorgensen A, Carreon LY, Andersen MO. Patient-reported outcomes and patient-reported satisfaction after surgical treatment for cervical radiculopathy. Global Spine J 2018;8:703-8.