Outcomes and value in elective cervical spine surgery: an introductory and practical narrative review

Scott L. Zuckerman[1,2], Clinton J. Devin[1,2]

[1]Vanderbilt Department of Neurosurgery, Vanderbilt University Medical Center, Nashville, TN, USA; [2]Steamboat Orthopaedic and Spine Institute, Steamboat Springs, CO, USA

Contributions: (I) Conception and design: All authors; (II) Administrative support: CJ Devin; (III) Provision of study materials or patients: All authors; (IV) Collection and assembly of data: SL Zuckerman; (V) Data analysis and interpretation: All authors; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: Clinton J. Devin, MD. Steamboat Orthopaedic and Spine Institute, Steamboat Springs, CO, USA. Email: clintondevin@gmail.com.

Abstract: How we determine a successful clinical outcome and the value of a spine intervention are two major questions surrounding clinical spine research. Patient-reported outcomes (PROs), both LEGACY and Patient-Reported Outcomes Measurement Information System (PROMIS) measures, are becoming ubiquitous throughout the literature. Spine surgeons need a facile understanding of the financial landscape of their environment to influence change. In the current introductory, narrative review on outcomes and value in cervical spine surgery, we aim to: (I) define relevant outcome and cost terminology, (II) review recent cervical spine surgery literature, divided by specific pathology with a focus on LEGACY and PROMIS measures, and (III) discuss value and cost as they pertain to postoperative return to work and ambulatory surgery centers surgeries.

Keywords: Cervical spine surgery; cervical radiculopathy; cervical myelopathy; cervical deformity; patient-reported outcomes; Patient-Reported Outcomes Measurement Information System (PROMIS); value; return-to-work; cost-effectiveness

Introduction

Cervical spine operations have been helping patients with degenerative spine conditions for decades, yet what constitutes an acceptable surgical outcome has evolved. Patient-reported outcomes (PROs) have revolutionized the definition of surgical success, and the accuracy and validity of these instruments continue to improve with newer and more streamlined systems. Patient-Reported Outcomes Measurement Information System (PROMIS) measures with the use of computer adaptive tests (CATs) have allowed for less administrative burden without sacrificing precision.

With durable improvement in function, disability, and quality of life, cervical spine surgery has increased in recent years, and so have corresponding healthcare costs (1). As more surgery is done on an older, comorbid population,

costs have risen to approximately $2 billion per year (2,3). Especially in the context of suboptimal healthcare metrics compared to other countries, U.S. surgeons must now be fiscally responsible without sacrificing outcomes. To curtail spending, the Center for Medicare Services (CMS) is replacing traditional fee-for service models with value-based compensation models, such as bundled payments, where a single payment per procedure is provided for all hospital, physician, and postsurgical care (4). It is essential to understand the financial landscape of our hospital systems and gain a seat at the table in these ongoing deliberations.

The aim of the current narrative review is to provide a practical, concise summary of outcomes and value in elective cervical spine surgery. In studies specifically focusing on degenerative conditions, we aim to address the following objectives: (I) define relevant outcome and

cost terminology, (II) review recent cervical spine surgery literature, divided by specific pathology with a focus on LEGACY and PROMIS measures, and (III) discuss value and cost as they pertain to postoperative return to work and ambulatory surgery centers surgeries. After reading this practical review, it is our hope the reader has a cogent understanding of outcomes and value, and can use this information to improve measurement of his/her own practice and facilitate discussion with their hospital system.

Methods

A narrative review was performed. Studies were chosen through relevant PubMed searches prioritizing the largest and most recent studies. Preference was given to well known, large, multi-institution databases that represented care across many centers, in addition to larger single-center studies. Most literature was restricted to publication in the last 5 years, from 2015–2019.

Definitions

Databases
Several databases and registries are routinely used to report the state of elective cervical surgery and cervical spine disease (*Table 1*). The majority of these databases are housed in the U.S., along with select international registries. These range from elective, degenerative surgeries (QOD, AOSpine) to spinal deformity surgery (ISSG).

Patient-reported outcomes (PROs)
Incited by the realization that physical examination, radiologic interpretation, and physician assessment were poor proxies in estimating outcomes, PROs have become the current standard for clinical spine research (5,6). PROs are simply any measure of change that is quantified directly from the patient. PROs are called by many names, including but not limited to: patient-reported outcome metrics (PROMs), quality of life (QOL), health-related quality of life (HRQL), and others. Albeit confusing, these terms all mean the same thing—the patient's self-reported status of their health state. One important nuance is that PROs measure different constructs and are most often divided into 3 domains: (I) pain level, (II) the patient's function or amount of disability, and (III) general quality of life. Little correlation exists between these three domains, and they in fact measure different entities altogether (7). Furthermore, different parts of a single PRO can measure

each construct, for example, the EuroQol-5D (EQ-5D) questionnaire assesses general health state through five questions addressing mobility, self-care, daily activities, pain, and anxiety/depression. The Neck Disability Index (NDI) enlists 10 questions broken down into pain intensity and amount of disability related to several daily activities. Studies can sometimes evaluate specific items within a PRO to see which ones are most responsive (8), a term to be expanded on later.

LEGACY measures
LEGACY measures represent PRO questionnaires where all questions are administered to the patient in entirety, regardless of prior answers. For example, if a patient reports high disability on the NDI in sleep, work, and driving, the full question and answer options about self-care will still be asked, despite the obvious fact that the patient undoubtedly struggles with self-care given their previous answers. The most widely studied cervical LEGACY measures are the NDI, modified Japanese Orthopaedic Association scale (mJOA), Visual Analog Scale arm/neck (VAS-arm/neck), in addition to generic health questionnaires used in any spine and all medical research, such as the EQ-5D and SF-36 PCS/MCS (*Table 2*). The NDI has been shown to be the most reliable and well validated metric for disability in patients with cervical spine pathology (9,10).

PROMIS
Burdened by the cumbersome, expensive, and time-consuming nature of LEGACY PROs, shortening questionnaires has been recommended by several experts (7,11). Moreover, LEGACY measures are prone to floor/ceiling effects. Floor effects are when baseline functioning falls below what is capturable, and ceiling effects are when post-surgery improvements exceed what is capturable. Thus, the quantified change is unreliable due to the discrepancy in the patient's true status and what the PRO is able to measure (5).

Given these drawbacks, Patient-Reported Outcomes Measurement Information System (PROMIS) has come into favor to standardize clinical outcomes in a valid, reliable, and efficient manner. PROMIS represents a validated system of PRO measurements that use computer adaptive tests (CATs), which use the response from the previous question to select relevant follow-up questions, to refine a respondent's score. Questions are administered until the score is adequately precise (standard error less than 3.0 on the T-score metric), which usually results in

Table 1 Summary of common databases and registries of cervical spine surgery

Database	Location	Details
Quality Outcomes Database (QOD)	82 sites; United States	Prospective registry of over 37,000 total patients (4,000 cervical)
		Includes: all elective spine surgery
		Excludes: infection, tumor, fracture, trauma, deformity, pre-existing spine disease or injury
		Enroll continuously or on prespecified rotating cycle ensuring unbiased enrolment
International Spine Study Group (ISSG)	13 sites; United States	Prospective registry of patients age >18-years-old
		Cervical deformity defined by at least one of following baseline imaging factors: cervical kyphosis (C2–7 Cobb angle >10°), cervical scoliosis (C2–7 coronal Cobb angle <10°), C2–7 SVA (cSVA) of more than 40 mm or chin-brow vertical angle of 25° or more
		Collects several different Patient-reported outcomes (PROs)
DaneSpine	15 sites; Denmark	5 private and 10 public centers
		Danish national spine database
		Several different PROs collected
AO Spine Clinical Research Network	26 sites; Latin America, Europe, Asia Pacific, North America	Large international research network
		Enrolled from several registered, prospective observational trials, including: CSM-North America, CSM-International
		Several different PROs collected
Prospective Spine Treatment Outcomes Study (PROSTOS)	14 sites; United States	Prospective registry
		Part of Association for Collaborative Spinal Research (ACSR)
		Several different PROs collected

12 questions, drawn from pools of up to >100 questions (5). The value and elegance of CATs used in PROMIS is the precision gained through fewer questions, thereby reducing time to survey completion, patient fatigue, and administrative support (5). PROMIS scores range from 0 to 100 with a mean of 50 and standard deviation of 10. Higher scores mean more of the outcome that is measured, which can be desired or undesirable. For example, a high score on pain/depression means more pain/depression (negative outcome), yet a high score on physical function means more function (positive outcome). PROMIS measures are grouped to test a specific construct (physical function, emotional function), and though over 11 PROMIS domains have been used in spine surgery (12), this review will discuss the 6 most commonly used domains determined in a recent systematic review: physical function (PF), pain interference (PIf), pain intensity (PIt), pain behavior (PB), anxiety (Ax) and depression (Dp) (*Figure 1*).

PROMIS appears to be at the forefront of future spine PRO research. That said, the current challenge lies in understanding what constitutes a meaningful change. The review below summarizes correlations between LEGACY and PROMIS tools, but defining objective clinical change remains elusive. Furthermore, many insurance companies require a certain level of disability as measured by the Oswestry Disability Index (ODI) to approve surgery, a LEGACY measure. In leu of this ongoing shift, many centers are collecting both PROMIS and LEGACY scores to provide an understandable transition.

Patient satisfaction

Patient satisfaction is defined as how satisfied the patient is after surgical treatment.

Despite its apparent simplicity, post-treatment satisfaction has been shown to be a poor predictor of medical response and depends on many factors related to physician-patient interaction, patient expectations, preoperative health issues, and psychological distress (13-16). Furthermore, some evidence suggests that those who start out with significant disability and achieve dramatic change scores may be more likely to be dissatisfied with their outcome than those that start out with minimal disability

Table 2 Summary of LEGACY PRO instruments

Instrument	Audience	Questions	Construct	Details
NDI	Cervical	10	Disability	10 questions: recreation, sleep, driving, work, concentration, headaches, reading, lifting, personal care, and pain
				Questions scored 0–5 with 0 (no disability) and 5 (full disability)
				Score doubled to reflect percentage; ranges from 0% (best function) to 100% (worst function)
				Despite name, measures arm pain as well
mJOA score	Cervical	18	Disability	Measures upper extremity function [5], lower extremity function [7], sensory function [3], and bladder function [3]
				Questions address eating, shirt buttoning, walking, sensation loss, micturition difficulty
				Higher score [18] is better function, lower score [0] is worse function [opposite of NDI]
				Often divided into 3 categories approximately divided into: mild [>13], moderate [9–13], severe [<9]
NRS/VAS arm/neck	Cervical	2	Pain	NRS is a purely numerical 11-point scale; 0 (no pain) and 10 (worst pain imaginable)
				VAS is a visual 10 cm line; left "no pain" and right "worst pain imaginable"; distance from left is score in cm
				Sometimes classified into mild (0–3), moderate [4–7], severe [8–10], or iterations thereof
EQ-5D	General	5	Quality of health state	Designed for simple health and economical assessment; many languages; international
				5 dimensions of mobility, self-care, usual activities, pain/discomfort, anxiety/depression
				Each question divided into 5 levels of severity
				Scored into 245 distinct health states; translated to score 0 (death) to 1.0 (perfect health)
				Contains a VAS component but rarely used
SF-36 PCS/MCS	General	36	Quality of health state	Assesses global health functioning with breakdown into physical and mental health
				Validated in many disease states and allows comparison to other diseases

and have minimal improvement from their baseline (17,18). As such, experts suggest not using patient satisfaction alone to measure quality and outcomes (19,20).

The outcome of satisfaction has been measured in several ways (21), yet the most common method is the North American Spine Satisfaction (NASS) questionnaire. The NASS four-point satisfaction questionnaire has four levels: (I) surgery met my expectations, (II) I did not improve as much as hoped but would undergo same operation with same results, (III) surgery helped but I would not undergo the same operation with the same results, and (IV) I am same or worse as compared to before surgery. This ordinal scale is often dichotomized into the first two options as "satisfied" with surgery and the latter two as "not satisfied" with surgery.

MCID

Minimum clinically important difference (MCID) is the smallest meaningful change in a PRO that is perceived as beneficial to the patient (22). MCID was created as a clinical rather than statistical measure of improvement, thought to be more representative of clinical care than a statistical significance alone. An in-depth discussion of MCID (23) is beyond the scope of this review, but in brief, MCID analyses usually involve an anchor that dichotomizes the group into "responders" and "non-responders". A common anchor used is the health transition index (HTI) of the SF-36 that asks the patient how their current health compares to their health prior to surgery—choices are "worse", "same", "slightly better", or "markedly better". The HTI is often dichotomized into same/worse as non-responders,

PROMIS Measure	Items	Audience	Scoring	Measurement
Physical Function (PF)	165	General Appropriate for adult populations Universal rather than disease-specific	• Item category response ranges 1–5 • Score recorded in T-scores derived from sampling weighted to reflect the US population • Mean score is 50 • Standard deviation is 10 • A higher PROMIS T-score represents more of the concept being measured • For negatively worded concepts like depression or pain, a T-score of 60 is one SD worse than average (more depression/pain), while a T-score of 40 is one SD better than average (less depression/pain)	Measures self-reported capability rather than actual performance of physical activities Includes the functioning of one's upper extremities (dexterity), lower extremities (walking or mobility), and central regions (neck, back), as well as instrumental activities of daily living, such running errands
Pain Interference (PIf)	40			Assesses self-reported consequences of pain on relevant aspects of one's life Includes how pain hinders engagement with social, cognitive, emotional, physical, and recreational activities Incorporates items probing sleep and enjoyment in life
Pain Intensity (PIt)	3			Assess how much a person hurts A fairly homogeneous dimension and easy for adults to gauge Of note, CAT is not used, only a single raw score from 3 questions measured 1–5 Total scores range 3–15
Pain Behavior (PB)	20			Measures self-reported external manifestations of pain, behaviors indicate to others that an individual is experiencing pain These actions or reactions can be verbal or nonverbal, and involuntary or deliberate Include observable displays (sighing, crying), pain severity behaviors (resting, guarding, facial expressions, and asking for help), and verbal reports of pain
Anxiety (Ax)	29			Assess self-reported fear (panic), anxiety (worry, dread), hyperarousal (nervousness, restlessness), and somatic symptoms related to arousal (racing heart, dizziness) Anxiety is reflected by autonomic arousal and experience of threat
Depression (Dp)	28			Assesses self-reported negative mood (sadness, guilt), views of self (self-criticism, worthlessness), and social cognition (loneliness, interpersonal alienation), as well as decreased positive affect and engagement (loss of interest, meaning, and purpose) Somatic symptoms (changes in appetite, sleeping patterns) are not included

Figure 1 Summary of PROMIS measures.

and slightly better/markedly better as responders (24). Several definitions and techniques for MCID analysis exist, of which 5 will be discussed.

(I) Average change—the average change score from preoperative to postoperative status seen in responders.

(II) Minimum detectable change (MDC)—the smallest change considered to be above the measurement error, which is often equal to the upper value of the 95% confidence interval for non-responders.

(III) Change difference—the difference in the average preoperative to postoperative change scores in responders and non-responders.

(IV) ROC method—a receiver operating characteristic curve (ROC) is made to determine the magnitude of change in a PRO that produces the optimal combination of sensitivity/specificity. The area under the curve (AUC) is the ability of the MCID change value to appropriately differentiate responders vs. non-responders. Sensitivity is the proportion of patients who were responders and had a PRO score above the MCID (true positive). Specificity is the proportion of non-responders who had PRO score below the MCID threshold (true negatives). The AUC measures the probability that scores will correctly discriminate between

responders and non-responders.

(V) Expert panel method—surveys experienced surgeons and asks them to define MCID values based on their clinical experience.

The maturation of MCID analyses has led to some confusion and uncertain validity (23). Subjective and inconsistent anchors have complicated the understanding of published MCID values, and measurements derived from certain populations are often applied to populations with different pathologies. As an example, MCID values for lumbar disease and cervical disease are often interchanged. Instead of individual manuscripts reporting their own MCID values, previously published values are often used, and wide ranges have been reported, such as: NDI of 7.5–15 (24-26), mJOA of 1.1–2 (27,28), NRS-arm/neck of 2.5–4.1 (24), EQ-5D of 0.05–0.54 (29-31), SF-36 PCS of 4.1–8.1 (24,31), and SF-36 MCS of 4.7–5.7 (31-34). Rather than utilizing an absolute number for change score, we would propose utilizing a 30% change from baseline as meaningful improvement. This approach better takes into account the baseline PRO measure (35).

What it means to "Validate" a PRO

The quality of PRO instruments is measured by their reliability, validity, feasibility, floor/ceiling effects, responsiveness, and correlation. As researchers, we have become flexible with this language, and the exact meaning of each term can be easily lost. When used without rigor, these terms become nebulous. However, when we say a PRO is validated, several components must be iteratively evaluated.

Reliability is the instrument's ability to produce consistent and similar results under the same circumstance, reducing error and variation in measurements (36,37). Inter-rater reliability quantifies an evaluator's ability to provide consistent evaluation and is measured with interclass correlation (ICC), for continuous variables, or Fleiss' kappa coefficient, for categorical variables. However, since patients provide data directly with PROs, rather than evaluators drawing inferences, ICC or Kappa are often not needed. Validity refers to the instrument's ability to measure what it aims to measure. Example include the ability of the NRS to measure pain or the NDI to measure disability. Discussed earlier in the context of PROMIS, feasibility means simply the ease of administration, both in terms of time, cost, and complexity to patients and providers. Floor/ceiling effects, also explained earlier, become problematic when >15% of respondents achieve the lowest or highest possible score (38).

Lastly, but perhaps most important in the context of this review is responsiveness. Responsiveness is the sensitivity of an instrument to capture clinically relevant changes. Said in other words, responsiveness is the ability to detect change in a disease state with a meaningful change in scoring (37). Validity and responsiveness are similar, yet responsiveness describes specifically the ability of an instrument's scoring system to reflect clinically meaningful change. ROC graphs and AUC are commonly used to calculate responsiveness, with AUC of 1.0 indicating perfect discrimination, 0.70 adequate, and 0.50 random chance. Lastly, correlation between two PROs is commonly studied between LEGACY and PROMIS measures, and since most are continuous variables, the Pearson correlation coefficient can be used. An r value is provided with a P value, signifying strong or loose correlation between measures. Keep in mind, that negative correlations can still signify tightly correlated variables, albeit in different directions, which is inherent to the instrument.

Cost-effectiveness

Cost is commonly presented in direct costs, costs incurred from direct healthcare expenses (hospital costs, surgeon fees, anesthesia fees, physical therapy, medications), and indirect costs, societal costs incurred from caretaker time loss and time out of work. The more granular in accounting for the costs, the more accurate the true cost burden of a given procedure. For example, need for discharge to a facility can account for 25% of the overall direct costs (39,40). Interestingly, indirect costs can account for up to 40% of the overall cost burden (41), therefore a quicker return to work can have a dramatic impact on reducing costs to society. Briefly, cost-effectiveness is commonly measured by comparing total costs of two separate groups coupled with the average change score in EQ-5D between both groups, known as the quality-adjusted life years (QALY). A QALY of 1 is tantamount to a year in perfect health, whereas a QALY of 0 represents a year of poor health. With cost and EQ-5D data, the cost per QALY can be calculated, which leads to the primary endpoint of most cost studies, which is an incremental cost-effectiveness ratio score or ICER. The ICER is the difference in mean total costs between cohorts divided by the difference in mean QALYs-gained. ICER have been collapsed into, high value: ICER <$50,000/QALY, intermediate value: ICER $50,000–150,000/QALY, and low value: ICER >$150,000/QALY (42). Understanding the drivers of cost and determining those factors that are modifiable, such as smoking or preoperative opioid intake,

allows for strategies to reduce cost and improve care. Furthermore, as we move toward a bundled payment model, it will be important to understand those characteristics that consistently drive up costs.

Outcomes

The umbrella term "outcomes" can mean many different things, including motor, sensory, or sphincter function, imaging changes, presence of a complication, pain level, or patient satisfaction. While it has been well proven that surgery leads to improvement, a more relevant discussion of "outcomes" surrounds which PROs are most reliable, responsive, and valid, and how measurement of outcomes can be optimized in a way that benefits patients, providers, and payers. Herein, we divide these discussions of elective cervical spine surgery by specific pathology: radiculopathy, myelopathy, deformity, satisfaction, and PROMIS measures.

Cervical radiculopathy

Several studies have evaluated the responsiveness of LEGACY measures in cervical radiculopathy. Using the QOD to study 2,206 patients undergoing elective surgery for cervical radiculopathy, Khan *et al.* (18) found that the 12-month NDI had the highest association with 12-month patient satisfaction, and level of satisfaction increased with decreased 12-month NDI scores independent of baseline NDI. Interestingly, the magnitude of NDI improvement was less important than the final NDI itself, a theme seen later with myelopathy and mJOA scores. Other measures that were significantly associated with patient satisfaction in descending order of importance were: 12-month NRS-AP, 12-month EQ-5D, baseline NDI, 12-month NRS-NP, baseline NRS-AP, and ASA grade (18). NDI was thought to have better correlation with satisfaction because 7 of the 10 NDI questions are related to activities of daily life. Further scrutinizing the NDI, a prospective, single-institution cohort of 137 patients undergoing elective cervical spine surgery showed that work, recreation and pain intensity were the only individual NDI domains that independently predicted improvement in function. Though this was not in a dedicated radiculopathy cohort (44.6% had radiculopathy), the results support collection of specific NDI domains, and not just an aggregate score (8).

Symptom duration has also been shown to affect PRO responses. Tarazona and colleagues (43) retrospectively reviewed 216 patients who underwent ACDF for cervical radiculopathy with LEGACY measures. Multivariate analysis revealed that radiculopathy for more than 2 years predicted worse outcomes, specifically lower SF-12 PCS/MCS scores, higher NDI, and higher VAS-neck/arm pain, than the 6-month to 2-year duration.

MCID analyses have also been conducted. A 2018 retrospective study of the DaneSpine registry by Andresen *et al.* (16) evaluated 272 patients with radiculopathy who underwent mostly 1–2 level ACDFs, and 73.3% were much/somewhat better compared to before surgery, using the HTI of SF-36 as the anchor. Using logistic regression to predict patient satisfaction, achieving an MCID for VAS-neck pain and SF-36 PCS were significantly associated with satisfied patients. Interestingly, the reason NDI was not found to be significant was because of its tight correlation with VAS-neck pain. Notably, while VAS-neck pain significantly improved, only 59% achieved an MCID. In an operation that is targeted to ameliorate radicular symptoms, it is no surprise neck pain did not improve as much as arm pain, and is a reminder that neck pain can originate from many sources not addressed with surgery. Thus, appropriate expectations should be discussed preoperatively.

The topic of neck pain is worthy of further discussion. By conducting a retrospective study of the PROSTOS database to identify baseline characteristics that predicted postoperative status, Passias and colleagues (44) found that using the NRS-neck/arm scale, patients with an arm-to-neck ratio (ANR) <1 (more neck than arm pain) were less likely to reach improvement at 2-year NDI and SF-36 PCS scores, and conversely, ANR >1 (more arm than neck pain) was associated with increased odds of SF-36 PCS improvement. Moreover, higher baseline NDI (more disability) was associated with a reduced odds of postoperative neck pain improvement.

Summary of recent cervical radiculopathy outcome studies:

- ❖ Radiculopathy patients report high rates of improvement after surgery across several PROs;
- ❖ Improvement in NDI, SF-36, and VAS-neck/arm all correlate well with patient satisfaction;
- ❖ Final NDI matters more than change in NDI;
- ❖ Longer duration of symptoms is associated with decreased chance of meaningful improvement;
- ❖ More neck than arm pain is a poor prognosticator of improvement.

Cervical myelopathy

Compared to radiculopathy alone, degenerative cervical myelopathy (DCM) represents a more severe disease

state (45). The often-quoted goal of operating on cervical myelopathy is to halt progression rather than improve symptoms. Whereas cervical radiculopathy represents a more elective problem where symptom improvement is the norm, myelopathy is a more urgent condition that many believe to be on the spectrum of spinal cord injury.

Using the QOD, Khan and colleagues (46) evaluated 2,156 myelopathy patients to assess 3- and 12-month outcomes based on preoperative mJOA category of mild [≥14], moderate [9–13], or severe [<9]. A 2-point improvement in the mJOA was defined as clinically meaningful (47). On unadjusted, paired t-tests, improvement at 3-months was seen in all groups; however, after multivariable analysis, only the severe myelopathy group kept improving from 3- to 12-month. Said in other words, amount of preoperative myelopathy was a significant predictor of improvement from the 3- to 12-month period, where severe patients continued to improve over longer durations of time. In a similar but smaller study, Kopjar *et al.* (28) compared 60 very severe DCM patients (mJOA ≤8) to 188 severe DCM patients (mJOA 9–11) in an international study. Choice of surgery was up to the training clinician, and previously published MCID values were used: mJOA (1.1), SF-36 PCS (4.1), SF-36-MCS (5.7), NDI (7.5). While both cohorts improved in mJOA, NDI, and SF-36 PCS/MCS scores, at 2 years, 34% of the very severe group still had very severe disease, while 22% of the severe group still had severe disease, showing more improvement in the severe group. Similar to radiculopathy, longer duration of myelopathy was associated with poorer response to surgery. Interestingly, somewhat differing from the above study by Khan *et al.* (46), the majority of improvement was obtained by 6 months, but significant improvement after 6 months was not seen. Combining the results of these two studies, it appears the improvement in the 3–12 months period described by Khan *et al.* (46) may take place mostly during the 0–6 months period shown by Kopjar and colleagues. Furthermore, both studies confirm that opportunity for improvement in very severe myelopathy patients may be limited.

Though patients may improve on mJOA, how meaningful is this? With regards to satisfaction and myelopathy, Asher *et al.* (48) studied 1,963 DCM patients from the QOD to decipher which patients changed mJOA category of mild [≥14], moderate [9–13], or severe [<9]. At 1-year follow-up, 55% remained in the same category, 37% improved, and 7% worsened by one category. The variable most strongly associated with patient satisfaction was 12-month mJOA

score, regardless of preoperative mJOA, and the most satisfied patients were in the mild category at 12 months. Other predictors in descending order of importance were diabetes, older age, type of occupation, and longer duration of symptoms. These findings are important in relation to patient's expectations and happiness—as it stands, patients are not nearly as satisfied if they don't end up in the mild group. This is likely an unrealistic expectation for severe patients who are less likely to improve, and conversely, those with mild myelopathy need only halt the progression of symptoms to remain satisfied. These findings provided support for surgeons to operate on mild myelopathy rather than wait until the disease has become severe. Other smaller satisfaction studies have concluded that DCM patients after laminoplasty were most satisfied at 5 years when lower extremity function improved (49). And a similar studies showed high correlation of satisfaction with postoperative NDI scores (50).

Equally useful to knowing how patients fare after surgical intervention is predictors of outcomes. Gerling *et al.* (51) retrospectively reviewed 203 DCM patients and found that a history of cervical spine surgery was the strongest risk factors for developing a complication. Moreover, NDI and SF-36 PCS/MCS improved at 2 years by statistical significance (no MCID analysis was used), and having a complication did not negatively affect PROs at 2 years postoperatively.

Optimal surgical approaches and procedure selection in DCM patients have also been debated. To answer the question of anterior *vs.* posterior approach in patients with 3–5 levels of myelopathy, Asher and colleagues (52) queried the QOD database to compare 163 (67%) patients who underwent anterior approaches *vs.* 82 (33%) patients who underwent posterior approaches. Despite a shorter LOS in the anterior group of 1 *vs.* 3 days, no difference was seen in readmission, return to work (RTW), and the 12-month LEGACY PROs of NDI, EQ-5D, NRS-neck/arm, and mJOA. No MCID analysis was done, and the outcome of interest was raw score of each LEGACY measure. To answer a second debate between surgical approach, Fehlings and colleagues (32) followed CSM patients, 100 of which underwent laminoplasty and 166 underwent laminectomy and fusion, until 2 years postoperatively. Both groups showed significant improvement in mJOA, NDI, SF36-PCS/MCS, and after adjusting for group differences, no differences in any outcomes were seen before procedures. In both cohorts, MCID was achieved for NDI (7.5) and mJOA (1.11), except for SF-12 PCS/MCS, where only

laminoplasty achieved MCID (4.4/5.7 respectively).

Summary of recent cervical myelopathy outcome studies:

❖ Severely myelopathic patients may experience improvement over a longer period of time, in the 3–6 or 3–12 months range;

❖ Final mJOA scores in the "mild" group were more predictive of satisfied patients than improvement in mJOA, and mild patients who did not improve significantly were satisfied, supporting early surgery for myelopathy;

❖ Lower extremity function, and the ability to ambulate, are predictive of patient satisfaction after myelopathy surgery;

❖ Similar outcomes are reported between anterior and posterior approaches, yet anterior approaches are associated with shorter LOS;

❖ Similar outcomes have been reported between laminoplasty *vs.* laminectomy and fusion.

Cervical deformity

Less outcome information exists on cervical deformity patients. Ailon and colleagues (53) from the ISSG studied 77 patients who underwent cervical deformity surgery, and significant improvement was seen at 1-year in NDI, NRS neck pain, and EQ-5D, showing that despite the surgical morbidity, patients improve. Another study of the ISSG by Passias *et al.* (54) sought to answer if alignment improvement or myelopathy improvement drove PROs in 70 patients undergoing cervical deformity surgery. Myelopathy improvement was defined statistically and by MCID, and alignment improvement was defined by improvement in at least one Ames classification modifier grade from baseline to 1-year. A total of 44% improved in mJOA and 19% met 1-year mJOA MCID values. Group distributions were as follows: 16/70 (23%) alignment-only improvement, 13/70 (19%) myelopathy-only improvement, 18/70 (26%) alignment and myelopathy improvement, and 23 (33%) no improvement. Improvement in myelopathy symptoms and functional scores was associated with superior 1-year PRO data, yet no relationship was seen between 1-year PRO outcomes and cervical specific sagittal parameters. Of note, global, whole spine parameters of C2–S1 SVA, C7-S1 SVA showed significant correlation with 1-year mJOA, EQ-5D and NDI. These results emphasize the importance of myelopathy and whole spine alignment in driving PROs, rather than cervical specific alignment. In terms of alignment, studies of thoracolumbar deformity have found greater degree of deformity correction

correlated with several PROs (55,56). Moreover, one study showed preoperative C2–C7 SVA was heavily associated with both SF-36 PCS and with NDI at early follow-up in posterior cervical fusion cases (57).

Another ISSG study of 104 cervical deformity patients found that the C2-slope, defined as the angle between the lower C2 endplate and the horizontal plane, was a singular measurement that correlated well with 1-year LEGACY measures of NDI (r=0.63), mJOA (r=−0.65), and moderate correlation with NRS neck (r=0.49) and EQ-5D (r=0.50). The C2-slope had benefit in its simplicity (a single measure) and that its combined alignment at the OC junction and also the CT junction. An approximate cut-off for severe deformity of 17–20 ℃ degrees was proposed.

As seen in these chosen studies, PROs used in cervical deformity are the same PROs used in less invasive cervical operations. Passias and colleagues (58) asked if the same PROs used in more minor elective spine surgery were sensitive enough to be used in a cervical deformity population. In a total of 63 deformity patients, the authors reported that a one-point improvement in NDI increased the odds of achieving MCID on the mJOA and an improvement in EQ-5D by 0.1 increased the odds of reaching mJOA MCID at 1-year. However, when adjusting for radiographic, comorbidities, and surgical invasiveness, these correlations were lost. The authors concluded that since correlation with improvement in functional outcomes, defined by mJOA, were lost on multivariable analysis, a need for cervical deformity specific PROs measure may exist.

The morbidity of cervical deformity surgeries should be emphasized, as 64% of patients experience a complication (59). Passias *et al.* (59) reported a high accuracy for predicting a complication (AUC=0.79) with the following variables: higher baseline EQ-5D pain/depression, higher global SVA. Medical complications were predicted by male gender, baseline mJOA score, and cervical SVA, which surgical complication was predicted by higher EBL, lower anxiety scores, and larger global SVA. Obese patients were also found to have 5-times greater odds of developing infections after surgery (60).

Summary of recent cervical deformity studies:

❖ Patients improve after deformity surgery, though rates are lower than radiculopathy and myelopathy patientsl;

❖ Improvement after cervical deformity appears to be driven most by myelopathy and global spinal alignment, rather than cervical specific alignment,

though some smaller studies show correlation between cervical C2–7 SVA and SF-36 and NDI;

❖ Depressed patients have potential to improve after cervical deformity;

❖ Risk factors for complications after cervical deformity surgery include male gender, higher baseline mJOA, higher cervical and global SVA, blood loss, and obesity.

PROMIS

Spine studies of PROMIS are growing, and though most studies refer to a lumbar population, cervical studies are not far behind. The first iteration of PROMIS literature largely aimed to validate PROMIS by comparison with well tested LEGACY measures. Boody and colleagues (5) studied 59 patients who underwent cervical spine surgery and established validity between the NDI and SF-12 PCS/MCS with PROMIS PF, PIt, and PB. PROMIS measures demonstrated validity and responsiveness similar to LEGACY measures, with minimal floor/ceiling effects. Strongest correlations were seen when measures were grouped by function, such as PROMIS PF and SF-12 PCS (r=0.57), PROMIS PIt and NDI (r=0.61), and PROMIS PB and SF-12 MCS (r=–0.44). When split into radiculopathy and myelopathy subgroups, correlation remained high; however, myelopathy patients showed less clinical improvement, as expected. PROMIS required 3.2 minutes for all 3 measures, compared to 7.5 minutes for both LEGACY measures. The same study evaluated MCID in PROMIS, and with the proposed threshold of 50% of the standard deviation, PROMIS PB/PIt and NDI reached MCID threshold, yet PROMIS PF and SF-12 PCS/MCS did not. The authors proposed this was due to the variable improvement seen in early follow-up, especially for myelopathy patients.

In another dedicated cervical population, Moses *et al.* (61) aimed to test the association of PROMIS PF, PIt, and PIf with NDI and VAS neck/arm. All 3 PROMIS measures correlated well with LEGACY measures, strongest of which was NDI with all three PROMIS PF (–0.77), PIt (0.60), and PIf (0.79), interestingly which correlated better than NDI and VAS, two LEGACY measures. While NDI had floor effect of 7.1% at one end, no floor/ceiling effect was seen for three PROMIS measures.

A very recent study assessed the use of PROMIS exclusively in patients harboring cervical radiculopathy. Owen *et al.* (62) evaluated 75 patients undergoing surgery for cervical radiculopathy secondary to disc herniation

at a single center with completed baseline, 6-month and 1-year outcomes data. The authors correlated PROMIS PF with NDI and PROMIS PIt with VAS arm/neck pain at 1-year. PROMIS PF and NDI scores demonstrated strong negative correlation with r values of –0.81, –0.77, and –0.76 at each interval. However, PROMIS PIt had moderately positive correlation with VAS neck pain of 0.51, 0.61, 0.6, and VAS arm pain of 0.46, 0.47, 0.45. Specific to cervical radiculopathy, PROMIS PF correlated well with NDI at baseline and postoperative periods, while PROMIS PIt had moderate correlation with VAS neck/arm pain. In a devoted myelopathy population by the same group, Owen *et al.* (63) determined correlation between LEGACY NDI and mJOA with PROMIS PF in 60 patients who underwent surgery. PROMIS PF had good correlation with NDI at baseline (r=–0.69) and 6-months (r=0.76), and mJOA at baseline (r=0.61) and 6-months (r=0.72).

In 148 patients undergoing ACDF, Purivs *et al.* (64) determined validity and responsiveness of PROMIS PF, PIt, anxiety (Ax), depression (Dp), fatigue, social satisfaction, and sleep disturbance with LEGACY measures of NDI, SF-12, brief pain inventory (BPI), generalized anxiety disorder questionnaire (GAD-7), and patient health questionnaire (PHQ-8). All PROMIS domains showed moderate to strong correlation with NDI, SF-12, BPI, yet weak correlation was seen with VAS neck/arm. In terms of mental health LEGACY measures, all PROMIS measures correlated well with GAD-7 and PHQ-8 except for PROMIS PF, which measures function and not mood or pain. In terms of pain, VAS neck/arm pain had good correlation with PROMIS pain tests, but weak correlation with all other measures. With an anchor towards general health improvement after surgery, PROMIS MCID estimates were included: PF +4.5, PIt –5.2, Ax –5.7; Dp –4.6, fatigue –5.8, satisfaction with participation in social roles +4.4; and sleep disturbance –7.4.

In an even larger PROMIS study, Khalifeh and colleagues (65) reported PROMIS PF and PIf in 2,770 patients who underwent spine surgery, 1176 of which were cervical. PROMIS appeared to be most responsive at longer-term follow-up rather than short or intermediate. The radiculopathy patients experienced slightly greater improvements in PROMIS PF/PIf than myelopathy patients of +5.4 and +4.7, respectively. Though MCID was not done, these values were close to exceeding most MCID estimates of 3.5–5.5 points. Also in a mixed population, Sharma and colleagues (66) studied 42 cervical patients and correlated NDI, VAS neck/arm, and EQ-5D with PROMIS PF, PIt, and emotional distress (ED). PROMIS PIf was

positively correlated with VAS neck/arm pain at 3 months (r=0.407, r=0.347). PROMIS PF was negatively correlated with 3-month NDI (r=−0.703). There was no difference in responsiveness of LEGACY and PROMIS measures for physical function, and VAS neck/arm pain.

Summary of recent cervical studies using PROMIS measures:

- ❖ PROMIS correlates well (r>0.5) with most LEGACY measures;
- ❖ PROMIS PF appears to match well with NDI and mJOA, while PROMIS PIt and PIf match correlated with VAS arm/neck, though to a lesser extent;
- ❖ Less floor/ceiling affects are seen with PROMIS measures;
- ❖ In a radiculopathy patient, PROMIS PF and NDI correlated better than PROMIS PIt and VAS neck/arm pain.

Value

Return to work

A major source of value gained from spine operations is wage-earning from a successful return to work (RTW). RTW is also being utilized as a benchmark for providers, institutions, and payers to evaluate quality of surgery and patient recovery. Using the QOD, members of our group studied 4,689 patients who underwent cervical spine surgery in an effort to predict who returned to work at 3 months (3). A total of 82% returned by 3 months, and multivariable logistic regression revealed that those less likely to RTW were older (>56–65 years), were employed but not working or employed part-time, had a heavy/medium-intensity occupation, had workers' compensation suits, had higher baseline NDI, had myelopathy, and had more levels [3–5] fused. After consensus among all authors, 10 variables were chosen for the predictive nomogram: (I) age, (II) employment (full time *vs.* part time), (III) education level, (IV) employment status, (V) occupation intensity, (VI) workers' compensation, (VII) baseline NDI score, (VIII) symptom duration, (IX) predominant symptoms, and 10) number of levels fused. The nomogram formed out of these variables had an AUC of 0.812, signifying strong validity. The findings of this national, multi-site registry represent data from one of the largest studies to assess RTW after cervical spine surgery.

Several smaller, single-institutional studies have also discussed RTW after cervical spine surgery. Kim *et al.* (67) studied 269 patients from a single-institution, prospective registry and found that median RTW time was 35 days, and those less likely to return had labor-intensive jobs, higher ASA grade, history of CAD, and history of COPD. Additionally, those less likely to RTW had disc herniations compared to stenosis, underwent cervical corpectomy compared to posterior laminectomy and fusion, and endured longer operative times. In a similar study focusing on mental health of 104 patients who underwent elective, single-level ACDF for cervical myelopathy, those with worse mental health scores on the SF-12 MCS had no difference in return to work compared to those with better mental health scores.

In taking these results together, it appears work related factors might be most important in terms of predicting RTW at 3 months (3). Additional studies in the lumbar fusion population have shown that patients working at the time of surgery had 10-times the odds of returning to work after surgery, and time away from work preoperatively was strongest predictor of successful return to work (68). One can also not overlook the importance of workers' compensation, and how the impact of financial incentives are associated with lower odds of returning to work (69).

Ambulatory Surgery for ACDF

A dialogue of value in cervical spine surgery must include a discussion of ACDF in an ambulatory surgery center (ASC). Throughout the last decade, many studies have established the safety of ACDFs in an ambulatory setting by reporting similar morbidity profiles between operations performed in inpatient *vs.* ASC settings (70-75). Mullins and colleagues studied 1,123 patients undergoing ACDF and found no significant difference in complication rates between ambulatory and inpatient settings (4.1% *vs.* 3.0%, P=0.34) (72). A second study looking primarily at 2-level ACDFs found no differences in blood loss or complication rates between ASCs and inpatient settings. In some of the largest studies to date of 1,000 and 2,000 consecutive outpatient ACDF in all ages (4,10), low incidences of perioperative events such as neck hematoma, CSF leaks, or respiratory events in an ASC care setting were seen. Most importantly, when rare and potentially life-threatening adverse events due occur, they can be safely and effectively diagnosed, managed, and, if needed, transferred to an inpatient setting within the 4-hour ASC observation window.

The economic gains of ASCs have also been well-established. Purger and colleagues (76) compared outcomes and costs between 46,996 inpatient and 3,135 ambulatory

ACDFs in the states of California, Florida, and New York. Comparable morbidity and readmission rates were seen, while charges were significantly lower for the ambulatory setting ($33,362.51 *vs*. $74,667.04; P<0.0001) (76). Furthermore, Wohns *et al*. (77) found the cost of a single-level outpatient cervical disc arthroplasty was 84% less than an inpatient procedure. A recent narrative review by Sivaganesan *et al*. (73) summarized ample level 3 and 4 evidence to support ASC surgery for ACDF, though higher level evidence was lacking. The largest source of cost savings was not in procedure or instrumentation costs, as these are the same regardless of setting, but rather reduced facility fees and LOS. However, it must be noted that some have speculated the reason for these cost-savings may be due to a younger, healthier population undergoing surgery rather than the setting itself (78).

Despite growing evidence supporting the safety and economic value of ACDFs performed in an ASC environment (1,7), recent anecdotal reports have questioned whether outpatient ACDF surgery is safe (9). This debate was reinvigorated by two USA today articles, one in March 2018 titled, "How a push to cut costs and boost profits at surgery centers led to a trail of death" (8), and another in July 2018 titled, "Medicare reconsiders paying for seniors' spine operations at surgery centers" (9). These articles discussed deaths that occurred in patients who had undergone procedures at ASCs, including cervical spine surgery (8-10). Contemporaneously with these reports in July 2018, CMS reviewed 38 recently added procedures to the ASC Covered Procedures List (CPL) (5), 25 of which were spine procedures, including anterior interbody arthrodesis CPT code 22,551. In light of recent sensationalized anecdotes, it is possible that CMS and policy makers have succumbed to the availability cascade (79,80), where individual, widely publicized examples readily available to the public receive disproportionate attention and supersede the findings of established, balanced, empirical data.

Another reason for the CMS review may have been a disproportionate focus on flawed studies. As an example, Arshi *et al*. (3) produced a retrospective, case-control study comparing 1 and 2-level inpatient (10,964) versus outpatient (1,215) ACDFs using a large insurance database. Patients undergoing outpatient ACDF were more likely at one year to undergo either a posterior revision (5.51% *vs*. 4.11%) or anterior revision (3.59% *vs*. 3.11%) at the same or adjacent level, and more acute renal failure in the outpatient ACDF cohort was seen. Despite a large sample, the codes pertained

to outpatient surgery in an inpatient setting rather than a true ASC. More importantly, the outcome of "revision surgery" is simply not relevant to the uniqueness and benefit of an ASC. Frequency of adjacent segment spine procedures is most strongly influenced by patient and surgical factors rather than anything related to the facility. Simply put, the outcome did not match the study question. Lastly, insurance-based databases studies are simply not specifically designed to capture granular clinical data in unambiguous patient cohorts.

Regardless of one's vantage point on the ASC ACDF discussion, patient safety is the chief concern, with zero tolerance for complications and unforeseen events. The tragic cases outlined in the media re-emphasize the importance of maintaining the highest standards for ASC care while promoting protocols to avoid catastrophic complications. Ideas for standardization of improvement include attending level anesthesiologists present at all time, pathways for most severe complications (neck hematoma, aspiration, neurologic injury), and minimum distance to hospital setting.

Value of specific procedures

Additional studies have compared the value of certain cervical spine procedures to each other. The literature is replete with comparison of ACDF *vs*. arthroplasty, and though a full discussion of outside the scope of this review, studies on value are relevant. A cost comparison by Radcliff *et al*. (81) with a health claims database showed that in 6,635 ACDF *vs*. 327 arthroplasty, the reoperation rate was significantly increased in the ACDF group, while the index surgery and 90-day global window costs were lower in the arthroplasty group due to lower readmission rates and lower reoperation rates. For these reasons, costs paid by insurers were less for arthroplasty ($34,979) *vs*. ACDF ($39,820). Additional literature has echoed these findings, as one study endorsed cervical arthroplasty was more cost-effective at 7 years (82). However, some studies report similar and conflicting results depending on the cost-model utilized (83).

Bundled payments are becoming a preferred method of payment, where the hospital is given a single payment to handle all care, regardless of complication status. A study comparing bundled payments of 94,031 ACDFs *vs*. 5,010 posterior cervical foraminotomies was conducted by Alhourani *et al*. (4), where bundled payments were calculated by all payments from index hospitalization to 90 days postoperatively. Amidst gross underrepresentation by the foraminotomy group, median 90-day ACDF costs

were much higher ($31,567 *vs.* $18,412), as were index hospitalization and physician/hospital payments; however, no difference was seen in post-discharge payments. This is not surprising given the graft and hardware associated with ACDF, compared to non-fusion, foraminotomy group. The authors concluded that posterior foraminotomies may be more cost-effective in appropriate patients. Interestingly, the factors most closely tied to payments were age and comorbidity scores.

The importance of comorbidities in driving costs was also seen in two similar studies. One study by Chotai and colleagues (84) analyzed 445 patients undergoing elective ACDF from a single-institution registry and found that preoperative history of anticoagulation medication significantly contributed to 90-day costs, likely a surrogate for patient health status. Other drivers were length of surgery, number of levels, length of stay, imaging, and readmissions due to complications. A second analysis from the same institutional registry addressed intersurgeon cost variability in ACDF patients. Sielatycki and colleagues (85) studied 431 ACDFs done by 5 different surgeons and calculated direct/indirect costs, PROs, and also predicted costs based on observes *vs.* expected cost differential. No differences in PROs were seen among surgeons, yet significant variation existed in 90-day costs even after adjusting for patient comorbidities. Patient age was a significant driver of cost. One strength of this study was that the authors adjusted for patient comorbidity status with new comorbidity-adjusted pre-directed costs. Armed with this information, hospitals and practices can see which surgeons are above or below their expected costs, and aim to curb spending in surgeons who are above their expected charges.

Revision surgery has also been analyzed with respect to cost implications, and as expected, revisions significantly add to cost, making most procedures not cost-effective. A single-institution, prospective registry assessed the cost-effectiveness of 115 patients undergoing revision cervical surgery and while all PROs improved at 12-months, total costs ranged between $21,294 and $23,914 (86). There was no difference in costs for each of the revision subgroups of same-level recurrent disease, pseudarthroses, and adjacent segment disease. Importantly, the analysis did not take into account the index surgery and subsequent revision, only the single episode of revision surgery. In a cervical deformity population, Horn *et al.* (87) used the ISSG to assess the cost-effectiveness of revision surgery after cervical deformity correction. In 89 patients with a mean level fused of 7.7 and 34% combined approaches, the mean cost for

index surgery was $44,318 with cost per QALY of $27,267, signifying a cost-effective procedure. However, 11 revision surgeries occurred with longer mean levels fused (10.3) and average cost of $41,510, and a cost per QALY of $28,138, and when summing both surgeries, the total cost per QALY was $55,405, drastically reducing the value. Though the index cervical deformity case appeared to be cost-effective, if a reoperation occurs due to distal junctional failure, these are no longer cost-effective.

Conclusions

Cervical spine surgery successfully treats patients suffering from debilitating radiculopathy, myelopathy, and deformity. Outcomes show durable improvement across many PROs, and novel PROMIS measures are advancing the way we collect PRO data. In terms of value, predictors of return to work can be used to decrease indirect costs, and when done safely, ACDFs in the ambulatory setting can also lead to cost-savings. It is our hope that this introductory review can help surgeons foster a strong understanding of outcomes and value in the landscape of cervical spine surgery.

Acknowledgments

Footnote

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: https://creativecommons.org/licenses/by-nc-nd/4.0/.

References

1. Wang MC, Kreuter W, Wolfla CE, et al. Trends and variations in cervical spine surgery in the United States: Medicare beneficiaries, 1992 to 2005. Spine (Phila Pa 1976) 2009;34:955-61; discussion 962-3.

2. Patil PG, Turner DA, Pietrobon R. National trends in surgical procedures for degenerative cervical spine disease: 1990-2000. Neurosurgery 2005;57:753-8; discussion 758.

3. Devin CJ, Bydon M, Alvi MA, et al. A predictive model and nomogram for predicting return to work at 3 months after cervical spine surgery: an analysis from the Quality Outcomes Database. Neurosurg Focus 2018;45:E9.

4. Alhourani A, Sharma M, Ugiliweneza B, et al. Ninety-Day Bundled Payment Reimbursement for Patients Undergoing Anterior and Posterior Procedures for Degenerative Cervical Radiculopathy. Neurosurgery 2019;85:E851-9.

5. Boody BS, Bhatt S, Mazmudar AS, et al. Validation of Patient-Reported Outcomes Measurement Information System (PROMIS) computerized adaptive tests in cervical spine surgery. J Neurosurg Spine 2018;28:268-79.

6. Revicki DA, Cella DF. Health status assessment for the twenty-first century: item response theory, item banking and computer adaptive testing. Qual Life Res 1997;6:595-600.

7. DeVine J, Norvell DC, Ecker E, et al. Evaluating the correlation and responsiveness of patient-reported pain with function and quality-of-life outcomes after spine surgery. Spine (Phila Pa 1976) 2011;36:S69-74.

8. 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.

9. Westaway MD, Stratford PW, Binkley JM. The patient-specific functional scale: validation of its use in persons with neck dysfunction. J Orthop Sports Phys Ther 1998;27:331-8.

10. 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.

11. Hung M, Clegg DO, Greene T, et al. Evaluation of the PROMIS physical function item bank in orthopaedic patients. J Orthop Res 2011;29:947-53.

12. Haws BE, Khechen B, Bawa MS, et al. The Patient-Reported Outcomes Measurement Information System in spine surgery: a systematic review. J Neurosurg Spine 2019;30:405-13.

13. Wallace AS, Freburger JK, Darter JD, et al. Comfortably numb? Exploring satisfaction with chronic back pain visits. Spine J 2009;9:721-8.

14. Chotai S, Sivaganesan A, Parker SL, et al. Patient-Specific Factors Associated With Dissatisfaction After Elective Surgery for Degenerative Spine Diseases. Neurosurgery 2015;77:157-63; discussion 163.

15. Abtahi AM, Brodke DS, Lawrence BD, et al. Association between patient-reported measures of psychological distress and patient satisfaction scores after spine surgery. J Bone Joint Surg Am 2015;97:824-8.

16. Andresen AK, Paulsen RT, Busch F, et al. Patient-Reported Outcomes and Patient-Reported Satisfaction After Surgical Treatment for Cervical Radiculopathy. Global Spine J 2018;8:703-8.

17. Mummaneni PV, Bydon M, Alvi MA, et al. Predictive model for long-term patient satisfaction after surgery for grade I degenerative lumbar spondylolisthesis: insights from the Quality Outcomes Database. Neurosurg Focus 2019;46:E12.

18. Khan I, Sivaganesan A, Archer KR, et al. Does Neck Disability Index Correlate With 12-Month Satisfaction After Elective Surgery for Cervical Radiculopathy? Results From a National Spine Registry. Neurosurgery 2019. [Epub ahead of print].

19. Bederman SS. Commentary: Patient satisfaction in spinal surgery: should we be addressing customer service or health improvement? Spine J 2013;13:507-9.

20. Bible JE, Kay HF, Shau DN, et al. What Patient Characteristics Could Potentially Affect Patient Satisfaction Scores During Spine Clinic? Spine (Phila Pa 1976) 2015;40:1039-44.

21. Swarup I, Henn CM, Gulotta LV, et al. Patient expectations and satisfaction in orthopaedic surgery: A review of the literature. J Clin Orthop Trauma 2019;10:755-60.

22. Brozek JL, Guyatt GH, Schunemann HJ. How a well-grounded minimal important difference can enhance transparency of labelling claims and improve interpretation of a patient reported outcome measure. Health Qual Life Outcomes 2006;4:69.

23. Chung AS, Copay AG, Olmscheid N, et al. Minimum Clinically Important Difference: Current Trends in the Spine Literature. Spine (Phila Pa 1976) 2017;42:1096-105.

24. Carreon LY, Glassman SD, Campbell MJ, et al. 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;10:469-74.

25. Badhiwala JH, Witiw CD, Nassiri F, et al. Patient phenotypes associated with outcome following surgery for mild degenerative cervical myelopathy: a principal component regression analysis. Spine J 2018;18:2220-31.

26. Young BA, Walker MJ, Strunce JB, et al. Responsiveness of the Neck Disability Index in patients with mechanical neck disorders. Spine J 2009;9:802-8.

27. Fehlings MG, Kopjar B, Ibrahim A, et al. Geographic variations in clinical presentation and outcomes of decompressive surgery in patients with symptomatic degenerative cervical myelopathy: analysis of a prospective, international multicenter cohort study of 757 patients. Spine J 2018;18:593-605.

28. Kopjar B, Bohm PE, Arnold JH, et al. Outcomes of Surgical Decompression in Patients With Very Severe Degenerative Cervical Myelopathy. Spine (Phila Pa 1976) 2018;43:1102-9.

29. Parker SL, Adogwa O, Paul AR, et al. Utility of minimum clinically important difference in assessing pain, disability, and health state after transforaminal lumbar interbody fusion for degenerative lumbar spondylolisthesis. J Neurosurg Spine 2011;14:598-604.

30. Le QA, Doctor JN, Zoellner LA, et al. Minimal clinically important differences for the EQ-5D and QWB-SA in Post-traumatic Stress Disorder (PTSD): results from a Doubly Randomized Preference Trial (DRPT). Health Qual Life Outcomes 2013;11:59.

31. Parker SL, Godil SS, Shau DN, et al. Assessment of the minimum clinically important difference in pain, disability, and quality of life after anterior cervical discectomy and fusion: clinical article. J Neurosurg Spine 2013;18:154-60.

32. 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 multicenter studies. Spine J 2017;17:102-8.

33. Zhou F, Zhang Y, Sun Y, et al. Assessment of the minimum clinically important difference in neurological function and quality of life after surgery in cervical spondylotic myelopathy patients: a prospective cohort study. Eur Spine J 2015;24:2918-23.

34. 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:e67408.

35. Asher AL, Kerezoudis P, Mummaneni PV, et al. Defining the minimum clinically important difference for grade I degenerative lumbar spondylolisthesis: insights from the Quality Outcomes Database. Neurosurg Focus 2018;44:E2.

36. Teles AR KK, ABOSb, Falavigna A. Why and how should we measure outcomes in spine surgery? Journal of Taibah University Medical Sciences 2016;11:91-7.

37. McCormick JD, Werner BC, Shimer AL. Patient-reported outcome measures in spine surgery. J Am Acad Orthop Surg 2013;21:99-107.

38. Terwee CB, Bot SD, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol 2007;60:34-42.

39. Goltz DE, Ryan SP, Howell CB, et al. Patients at Risk for Exceeding CJR Cost Targets After Total Ankle Arthroplasty. Foot Ankle Int 2019;40:1025-31.

40. Snyder DJ, Bienstock DM, Keswani A, et al. Preoperative Patient-Reported Outcomes and Clinical Characteristics as Predictors of 90-Day Cost/Utilization and Complications. J Arthroplasty 2019;34:839-45.

41. Jonsson E, Olafsson G, Fritzell P, et al. A Profile of Low Back Pain: Treatment and Costs Associated With Patients Referred to Orthopedic Specialists in Sweden. Spine (Phila Pa 1976) 2017;42:1302-10.

42. Dubois RW. Cost-effectiveness thresholds in the USA: are they coming? Are they already here? J Comp Eff Res 2016;5:9-11.

43. Tarazona D, Boody B, Hilibrand AS, et al. Longer Preoperative Duration of Symptoms Negatively Affects Health-related Quality of Life After Surgery for Cervical Radiculopathy. Spine (Phila Pa 1976) 2019;44:685-90.

44. Passias PG, Hasan S, Radcliff K, et al. Arm Pain Versus Neck Pain: A Novel Ratio as a Predictor of Post-Operative Clinical Outcomes in Cervical Radiculopathy Patients. Int J Spine Surg 2018;12:629-37.

45. Nikolaidis I, Fouyas IP, Sandercock PA, et al. Surgery for cervical radiculopathy or myelopathy. Cochrane Database Syst Rev 2010;CD001466.

46. Khan I, Archer KR, Wanner JP, et al. Trajectory of Improvement in Myelopathic Symptoms From 3 to 12 Months Following Surgery for Degenerative Cervical Myelopathy. Neurosurgery 2019. [Epub ahead of print].

47. Tetreault L, Nouri A, Kopjar B, et al. The Minimum Clinically Important Difference of the Modified Japanese Orthopaedic Association Scale in Patients with Degenerative Cervical Myelopathy. Spine (Phila Pa 1976) 2015;40:1653-9.

48. Asher AL, Devin CJ, Weisenthal BM, et al. Effect of Modified Japanese Orthopedic Association Severity Classifications on Satisfaction With Outcomes 12 Months After Elective Surgery for Cervical Spine Myelopathy. Spine (Phila Pa 1976) 2019;44:801-8.

49. Ohya J, Oshima Y, Takeshita K, et al. Patient satisfaction with double-door laminoplasty for cervical compression myelopathy. J Orthop Sci 2015;20:64-70.

50. Schroeder GD, Coric D, Kim HJ, et al. Are patient-reported outcomes predictive of patient satisfaction 5 years after anterior cervical spine surgery? Spine J 2017;17:943-52.

51. Gerling MC, Radcliff K, Isaacs R, et al. Two-Year Results of the Prospective Spine Treatment Outcomes Study: An Analysis of Complication Rates, Predictors of Their Development, and Effect on Patient Derived Outcomes at 2 Years for Surgical Management of Cervical Spondylotic Myelopathy. World Neurosurg 2017;106:247-53.

52. Asher AL, Devin CJ, Kerezoudis P, et al. Comparison of Outcomes Following Anterior vs Posterior Fusion Surgery for Patients With Degenerative Cervical Myelopathy: An Analysis From Quality Outcomes Database. Neurosurgery 2019;84:919-26.

53. Ailon T, Smith JS, Shaffrey CI, et al. Outcomes of Operative Treatment for Adult Cervical Deformity: A Prospective Multicenter Assessment With 1-Year Follow-up. Neurosurgery 2018;83:1031-9.

54. Passias PG, Horn SR, Bortz CA, et al. The Relationship Between Improvements in Myelopathy and Sagittal Realignment in Cervical Deformity Surgery Outcomes. Spine (Phila Pa 1976) 2018;43:1117-24.

55. Grosso MJ, Hwang R, Mroz T, et al. Relationship between degree of focal kyphosis correction and neurological outcomes for patients undergoing cervical deformity correction surgery. J Neurosurg Spine 2013;18:537-44.

56. Ames CP, Blondel B, Scheer JK, et al. Cervical

57. Tang JA, Scheer JK, Smith JS, et al. The impact of standing regional cervical sagittal alignment on outcomes in posterior cervical fusion surgery. Neurosurgery 2015;76 Suppl 1:S14-21; discussion S21.

58. Passias PG, Horn SR, Oh C, et al. Evaluating cervical deformity corrective surgery outcomes at 1-year using current patient-derived and functional measures: are they adequate? J Spine Surg 2018;4:295-303.

59. Passias PG, Oh C, Horn SR, et al. Predicting the occurrence of complications following corrective cervical deformity surgery: Analysis of a prospective multicenter database using predictive analytics. J Clin Neurosci 2019;59:155-61.

60. Passias PG, Poorman GW, Horn SR, et al. Effect of Obesity on Radiographic Alignment and Short-Term Complications After Surgical Treatment of Adult Cervical Deformity. World Neurosurg 2019;125:e1082-8.

61. Moses MJ, Tishelman JC, Stekas N, et al. Comparison of Patient Reported Outcome Measurement Information System With Neck Disability Index and Visual Analog Scale in Patients With Neck Pain. Spine (Phila Pa 1976) 2019;44:E162-7.

62. Owen RJ, Khan AZ, McAnany SJ, et al. PROMIS correlation with NDI and VAS measurements of physical function and pain in surgical patients with cervical disc herniations and radiculopathy. J Neurosurg Spine 2019;1-6.

63. Owen RJ, Zebala LP, Peters C, et al. PROMIS Physical Function Correlation With NDI and mJOA in the Surgical Cervical Myelopathy Patient Population. Spine (Phila Pa 1976) 2018;43:550-5.

64. Purvis TE, Neuman BJ, Riley LH, et al. Comparison of PROMIS Anxiety and Depression, PHQ-8, and GAD-7 to screen for anxiety and depression among patients presenting for spine surgery. J Neurosurg Spine 2019;1-8.

65. Khalifeh JM, Dibble CF, Hawasli AH, et al. Patient-Reported Outcomes Measurement Information System physical function and pain interference in spine surgery. J Neurosurg Spine 2019;1-10.

66. Sharma M, Ugiliweneza B, Beswick J, et al. Concurrent Validity and Comparative Responsiveness of PROMIS-SF Versus Legacy Measures in the Cervical and Lumbar Spine Population: Longitudinal Analysis from Baseline to Postsurgery. World Neurosurg 2018;115:e664-75.

67. Kim EJ, Chotai S, Wick JB, et al. Factors Associated With

Return-to-Work Following Cervical Spine Surgery in Non-Worker's Compensation Setting. Spine (Phila Pa 1976) 2019;44:903-7.

68. Anderson PA, Schwaegler PE, Cizek D, et al. Work status as a predictor of surgical outcome of discogenic low back pain. Spine (Phila Pa 1976) 2006;31:2510-5.

69. Atlas SJ, Tosteson TD, Blood EA, et al. The impact of workers' compensation on outcomes of surgical and nonoperative therapy for patients with a lumbar disc herniation: SPORT. Spine (Phila Pa 1976) 2010;35:89-97.

70. Adamson T, Godil SS, Mehrlich M, et al. Anterior cervical discectomy and fusion in the outpatient ambulatory surgery setting compared with the inpatient hospital setting: analysis of 1000 consecutive cases. J Neurosurg Spine 2016;24:878-84.

71. Gornet MF, Buttermann GR, Wohns R, et al. Safety and Efficiency of Cervical Disc Arthroplasty in Ambulatory Surgery Centers vs. Hospital Settings. Int J Spine Surg 2018;12:557-64.

72. Mullins J, Pojskic M, Boop FA, et al. Retrospective single-surgeon study of 1123 consecutive cases of anterior cervical discectomy and fusion: a comparison of clinical outcome parameters, complication rates, and costs between outpatient and inpatient surgery groups, with a literature review. J Neurosurg Spine 2018;28:630-41.

73. Sivaganesan A, Hirsch B, Phillips FM, et al. Spine Surgery in the Ambulatory Surgery Center Setting: Value-Based Advancement or Safety Liability? Neurosurgery 2018;83:159-65.

74. Tally WC, Tarabadkar S, Kovalenko BV. Safety and feasibility of outpatient ACDF in an ambulatory setting: A retrospective chart review. Int J Spine Surg 2013;7:e84-7.

75. Vaishnav A, Hill P, McAnany S, et al. Safety of 2-level Anterior Cervical Discectomy and Fusion (ACDF) Performed in an Ambulatory Surgery Setting With Same-day Discharge. Clin Spine Surg 2019;32:E153-9.

76. Purger DA, Pendharkar AV, Ho AL, et al. Outpatient vs Inpatient Anterior Cervical Discectomy and Fusion: A Population-Level Analysis of Outcomes and Cost. Neurosurgery 2018;82:454-64.

77. Wohns R. Safety and cost-effectiveness of outpatient

cervical disc arthroplasty. Surg Neurol Int 2010;1:77.

78. Mundell BF, Gates MJ, Kerezoudis P, et al. Does patient selection account for the perceived cost savings in outpatient spine surgery? A meta-analysis of current evidence and analysis from an administrative database. J Neurosurg Spine 2018;29:687-95.

79. Solomon GS, Sills A. Chronic traumatic encephalopathy and the availability cascade. Phys Sportsmed 2014;42:26-31.

80. Barr WB. An evidence based approach to sports concussion: confronting the availability cascade. Neuropsychol Rev 2013;23:271-2.

81. Radcliff K, Zigler J, Zigler J. Costs of cervical disc replacement versus anterior cervical discectomy and fusion for treatment of single-level cervical disc disease: an analysis of the Blue Health Intelligence database for acute and long-term costs and complications. Spine (Phila Pa 1976) 2015;40:521-9.

82. Kim JS, Dowdell J, Cheung ZB, et al. The Seven-Year Cost-Effectiveness of Anterior Cervical Discectomy and Fusion Versus Cervical Disc Arthroplasty: A Markov Analysis. Spine (Phila Pa 1976) 2018;43:1543-51.

83. Warren D, Andres T, Hoelscher C, et al. Cost-utility analysis modeling at 2-year follow-up for cervical disc arthroplasty versus anterior cervical discectomy and fusion: A single-center contribution to the randomized controlled trial. Int J Spine Surg 2013;7:e58-66.

84. Chotai S, Sivaganesan A, Parker SL, et al. Drivers of Variability in 90-Day Cost for Elective Anterior Cervical Discectomy and Fusion for Cervical Degenerative Disease. Neurosurgery 2018;83:898-904.

85. Sielatycki JA, Chotai S, Wick J, et al. Intersurgeon Cost Variability in Anterior Cervical Discectomy and Fusion. Spine (Phila Pa 1976) 2018;43:1125-32.

86. Kim EJ, Chotai S, Wick JB, et al. Patient-Reported Outcomes and Costs Associated With Revision Surgery for Degenerative Cervical Spine Diseases. Spine (Phila Pa 1976) 2018;43:E423-9.

87. Horn SR, Passias PG, Hockley A, et al. Cost-utility of revisions for cervical deformity correction warrants minimization of reoperations. J Spine Surg 2018;4:702-11.