Outcome Assessment in Neuromuscular Spinal Deformity

Richard E. Bowen, MD, Mark F. Abel, MD, Vincent Arlet, MD, Desmond Brown, MD, Douglas C. Burton, MD, Phyllis D'Ambra, RN, MPA, Laura Gill, MD, Dale V. Hoekstra, MD, Lawrence I. Karlin, MD, James Raso, MASc, James O. Sanders, MD, and Frank J. Schwab, MD

Abstract: Patient-based outcome measures are important tools quantifying the disease-specific and/or global quality of life (QOL) effects of spinal deformity treatment. In patients with neuromuscular disorders such as cerebral palsy, muscular dystrophy, and myelomeningocele, treatment effects must be differentiated from underlying disease functional impairments. In general, the goals of spinal surgery in these patients are to improve QOL by enhancing sitting balance and posture, improving lung and gastrointestinal function, and reducing pain and deformity. In selected patients, improving ambulation and hand function may also be realistic surgical goals. QOL measures specific to both the neuromuscular diagnosis and spinal deformity provide higher quality information on treatment outcomes for a particular patient than standard radiographic measures. This article reviews patient-based outcome measures in spinal deformity patients with neuromuscular disorders, including their development and use in comparative outcome studies in the recent literature.

Key Words: neuromuscular scoliosis, outcome assessment, myelodysplasia, cerebral palsy, muscular dystrophy

(*J Pediatr Orthop* 2012;32:792–798)

The benefit of using patient- (or parent-) based outcome measures to assess the results of medical intervention is to understand whether treatments actually provide their purported benefits in the respondent's eyes. Clinicians are increasingly asked to assess the cost effectiveness of treatment.[1–3] This is particularly relevant to surgical management of patients with spinal deformity and chronic neuromuscular disorders such as cerebral palsy (CP), muscular dystrophy, or myelomeningocele, all of which impact baseline physical function and frequently exhibit spinal deformity progression.[4–9] The overall medical and surgical care of these patients is expensive,[10] surgical implant costs vary widely,[3,11] with mixed reports describing the benefits of spinal deformity surgery.[12–14] Another goal of measuring treatment outcomes therefore is to justify the cost of treatment to society by demonstrating outcomes in terms of incremental cost effectiveness. Furthermore, outcome measures are an important part of a physician's critical assessment of his or her own treatment outcomes, which is an essential part of the practice improvement process.[1] The purpose of this document is to provide a focused review of the scientific literature related to clinical outcomes in the management of neuromuscular spinal deformities.

The term "neuromuscular" includes a large group of heterogenous disorders with varying degrees of chronic motor, sensory, and central nervous system deficits. Although treatment outcomes of spinal deformity in patients with "neuromuscular" disorders were historically lumped together, different diagnoses impact patient function in variable ways, and patient function within each diagnosis is often quite different from patient to patient. Some common features frequently include the inability to walk, impairment of pulmonary and gastrointestinal function, and dependence on a caregiver for activities of daily living. The common goals for these patients are to improve seating balance and posture, hygiene, pulmonary, and gastrointestinal function. In patients with higher baseline function, surgical goals may also be used to increase activity participation and ambulatory level.[2,5] An often present but rarely discussed issue is the presence of pain from seating or ribs pushing against the pelvis, although pain is more difficult to assess in patients who are more cognitively impaired. So, although clear that outcome measures need to consider the specific neuromuscular condition and patient functional abilities, many surgical goals are shared across a wide variety of patients.

Neuromuscular disorders impact health in chronic and progressive ways. Medical interventions in these patients often have implications far into the future. Treatment goals, therefore, need to have a long-term component, and measuring treatment outcomes should be done both in the short-term and long-term.[15,16] This is particularly challenging in this group of patients, who are diagnosed early in life, often undergo surgery in the early teens, and may experience benefits or complications from treatment decades later. Moreover, the effects of treatment need to be separated from the changing functional status over time due to the nature of the underlying diagnosis. Thus, both the short-term and long-term benefits of surgery and the changing baseline functional status of untreated patients

From the Department of Orthopaedic Surgery, Geffen School of Medicine at UCLA, Los Angeles, CA.

The authors declare no conflict of interest.

Reprints: Richard E. Bowen, MD, Los Angeles Orthopaedic Hospital, 2400 S. Flower St., Los Angeles, CA 90007. E-mail: rbowen@mednet.ucla.edu.

without spinal deformity need to be addressed in outcome measures to truly assess the impact of surgery.

GENERAL OVERVIEW OF DIFFERENT TYPES OF MEASURES

Many different "outcomes" can be measured. There are several outcome measure categories in the neuromuscular spinal deformity literature. Radiographic measurements have long been used by physicians as outcome measures in spinal deformity surgery. However, measuring proxies such as Cobb angle correction, spinal imbalance, pelvic obliquity, trunk rotation may only minimally correlate or potentially negatively correlate with patient outcomes from large surgeries. However, even these outcomes can be difficult to obtain in neuromuscular curves because of difficulty in obtaining reproducible upright radiographs. Improvement of these measures after surgery does not necessarily correlate with patient-perceived functional improvement.[12–14] A second group of outcome measures are performance measures. Performance measures require the patient to perform a specific task which is evaluated for speed or efficiency. Examples include the Jebsen Hand Test for children,[17] and pulmonary spirometry. Depending on the baseline cognitive function of the patient, these tests can be difficult to administer.

The primary focus of this paper is on patient—or parent/caregiver—outcome measures (patient-based measures). These are questionnaires administered by paper and pencil, computer, or structured interview. They can measure general health-related quality of life (HRQOL) or be specific to scoliosis in a particular patient population. There are over 100 of these measures for use in spinal deformity alone.[3] However, few measures focus specifically on patients with neuromuscular disorders. Depending on the baseline cognitive abilities of the patient, these measures are either completed by the patient or by their primary caregiver (often a parent).[16] This type of outcome measure is subjective rather than physiological or performance based, but is the only type of outcome measure that directly quantifies the quality of life (QOL) of the patient.[18] Ideally, use of a combination of physiological, performance, and patient-based outcome measures gives a global picture of the patient's health status.

OUTCOME MEASURES IN NEUROMUSCULAR SPINAL DEFORMITY: THE VALIDATION PROCESS

Because the outcomes literature originated in the social sciences, the terms used are often unfamiliar to orthopaedic surgeons rooted in biological or physical sciences. Psychometric properties refer to the theory and technique of measuring knowledge, ability, attitudes, and personality traits through outcomes questionnaires. A brief overview of commonly used psychometric terms is given below.

Validity

Validity for an outcomes instrument generally means that responses to the questionnaire reflect what they are intended to measure. In the case of neuromuscular scoliosis, this means an instrument proposing to measure the function or comfort of someone with a neuromuscular scoliosis actually does measure function or comfort. All the tests of validity aim at identifying how well an instrument reflects what it intends to show.

There are different types of validity measures. Content validity is the concept that a questionnaire targeted at a certain topic should cover the breadth (content) of the topic. For example, in a questionnaire for neuromuscular scoliosis, the questionnaire should cover a range of issues such as mobility, seating difficulties, back discomfort, pulmonary function. If an instrument is valid, its results should correspond to the current standard. Criterion validity means that another criterion, such as an instrument or outcome, corresponds with the instrument under development. Predictive validity indicates the results of the instrument are reasonably predictable of another outcome (pulmonary function, sitting balance, etc).

In psychometrics a concept is referred to as a construct. A construct could be any idea such as "scoliosis leads to a poor self image" or "patients with neuromuscular scoliosis cannot do activities because of their deformity." Testing construct validity requires examining whether this concept, as measured in the instrument, stands up to statistical scrutiny. For "scoliosis," this can be done by using physical measurements such as the scoliometer, rib prominence, and trunk shift, or radiographic measurements like the Cobb angle, or from instruments that measure the patient's concept of the curve and its consequences. Convergent validity, as a specific component of construct validity, examines whether 2 measurements aiming at a similar "construct" behave similarly. Worded differently, measures of constructs that theoretically should be related to each other are, in fact, observed to be related to each other. For example, both the Cobb angle and an instrument measuring disability from scoliosis reflect the "construct" that larger curves are worse than smaller curves. An instrument showing the disability from curves should correspond at least generally to Cobb angle magnitude. Discriminate validity means the measure can discriminate between different constructs. For example, can a questionnaire distinguish between a patient with disability from a neuromuscular scoliosis versus that from their hip dislocation or pulmonary status?

Reliability

A fundamental rule is that a measurement cannot be valid unless it is first reliable or consistent. This is just as true for outcomes instruments as it is for other measures such as scoliosis severity (eg, Cobb angle measurements). Reliability can vary from person to person or time to time. It may also differ upon whether the effects measured are small or large.

There are several statistical methods of assessing reliability. All methods aim at assessing the amount of

error or noise and separating the noise from the true measurements (the signal). Selection of the method depends largely upon the type of measurement being used. The 3 methods are the Pearson correlation coefficient, the interclass correlation coefficient (ICC), and the κ value. Although most orthopaedists are familiar with the Pearson correlation coefficient, the correlation coefficient evaluates aggregate scores rather than corresponding scores and often overestimates reliability. The ICC compares individual scores to each other. Similar to the Pearson coefficient, it ranges from −1 to +1 and is interpreted similarly (Table 1).[19] Although the ICC is good for continuous variables, it is not designed for ordinal data such as Likert-scaled outcomes questionnaires. A better statistic for ordinal data are the κ statistic. When test-retest results are compared for the number of exact matches, exact matches can occur by chance alone. The κ statistic accounts for this chance factor. Landis and Koch[20] have published widely accepted criteria of κ coefficients (Table 1), although other criteria are also used.[21]

Health scales typically measure many variables to measure various aspects of a disease. For example, a scale on neuromuscular scoliosis may ask about difficulties with sitting, balance, and breathing. If all of the scales are contributing to the picture of scoliosis, you would expect them all to worsen as scoliosis worsens. If the scores do not correlate to some degree, you would conclude there is either a problem with our understanding of the disorder or there is a problem with the questions. Conversely, very high correlation of the 2 implies that the items may be redundant making only one of them necessary. Internal consistency is commonly measured with Cronbach α.

Bias

Questionnaires can be biased in various ways, limiting their utility. Some of the common biases are "recall bias" when a patient is asked about past states, such as "how much have you improved from surgery" or "how was your position six months ago." These biases are very common when physicians take histories from patients, and these types of questions must be viewed with caution. In neuromuscular scoliosis, 2 particular biases are floor and ceiling effects as the variety of patients is so large. If a patient is very high functioning (near the top or ceiling of the questionnaire), then changes may not occur, and similarly patients with low functioning may not worsen on a questionnaire even with severe complications (floor effect). Other common biases are listed in Table 2.

TABLE 1. Interpretation of κ and Interclass Correlation Coefficient Values With Respect to Outcome Instrument Reliability

κ Value	Interclass Correlation Coefficient	Reliability
$\kappa > 0.80$	0.9-1.0	Almost perfect
$\kappa = 0.61$-0.80	0.7-0.89	Substantial
$\kappa = 0.41$-0.60	0.5-0.69	Moderate
$\kappa = 0.21$-0.40	0.5-0.69	Fair
$\kappa = 0.00$-0.20	0.25-0.49	Slight
$\kappa < 0.00$	0.0-0.24	Poor

Responsiveness

If a deformity improves, then the instrument scores measuring disability from the deformity should also improve. If they do not improve, then the instrument may be insufficiently sensitive, or it may measure the wrong items (lack of validity). Similarly, an instrument may improve when there was no improvement, in which case the stability and validity of the instrument must be questioned. In a responsive instrument, the degree of responsiveness to treatments such as bracing or surgery should correspond to the actual improvement the patient experiences. The responsiveness of an instrument is also called its sensitivity to change. Responsiveness can be measured using variables such as effect size, standardized response mean, or relative efficiency. Changes over multiple time points can be modeled as growth curves using more complex modeling.

REVIEW OF PATIENT-BASED HRQOL MEASURES

General HRQOL Outcome Questionnaires

In the 1990s the movement to quantify outcomes lead the American Academy of Orthopaedic Surgeons and the Pediatric Orthopaedic Society of North America to commission a work group to develop functional health outcome scales assessing upper extremity function, transfers and mobility, physical function and sports, comfort, happiness and satisfaction, and expectations of treatment. The result was the Pediatric Outcomes Data Collection Instrument (PODCI),[22] a 108-item questionnaire available in 3 forms; a parent format (kids aged 0 to 10 y), a parent report format for adolescents (kids aged 11 to 18 y), and an adolescent report format (kids aged 11 to 18 y). Extensive validation was performed to assess reliability, content, and discrimination. Groups assessed included patients with common pediatric disorders such as fractures and scoliosis, and others with chronic progressive disorders including spastic diplegic and quadriplegic CP, myelodysplasia, and osteogenesis

TABLE 2. Types of Biases Potentially Present in the Development of Outcome Questionnaires

Optimizing	Difficult to interpret questions causing the patient to "optimize" the answer
Satisfying	In longer burdensome questionnaires, doing what it takes to finish the questionnaire
Social desirability	Respondent's subconscious desire to want a result to be better than it is
End aversion	Not wanting to answer at the extremes—also called "central tendency bias"
Positive skew	Tendency to skew results toward the positive
Halo-effect	Overall positive experience tends to cause all answer to reflect this even if the individual items would not have matched this level
Framing	How a question is worded will change the response to the answers

imperfecta. Testing across 470 patients found reliability to be good (internal reliability = Chronbach $\alpha > 0.8$). Concurrent validity was assessed by comparing the PODCI to the Child Health Questionnaire (CHQ). They found that the PODCI seemed to have an advantage over the CHQ in assessing upper extremity function and comfort, but that the CHQ performed better at younger ages. This was thought to be due to the Pediatric Orthopaedic Society of North America questionnaire having greater specificity of tasks than the CHQ and that younger children were not developmentally able to perform certain tasks at such a young age. The PODCI was developed to assess orthopaedic conditions specifically. The responsiveness was tested only in children sustaining femur fractures. This group did show significant improvement after healing as would be expected.

Lerman et al[23] administered the PODCI to patients with adolescent idiopathic scoliosis, juvenile idiopathic scoliosis, congenital scoliosis, and congenital kyphosis and compared the scores obtained with those of children without orthopaedic disability. They found the scores could discriminate children with spinal deformity from those without, with the former having lower scores in the domains of upper extremity function, transport, sports, and global function. Since then many studies have used the PODCI over the years. The assessment of ambulatory children with CP is perhaps the most common group studied with this instrument.[24,25] However, the questionnaire shows floor effects and poor responsiveness in the severely disabled group (Gross Motor Function Classification System class 4 and 5), which is the very group most likely to get severe spinal deformities. Many items in the questionnaire, such as ability to ride a bike, climbing stairs, lifting heavy objects, are not applicable for this group of patients. Thus, the PODCI is only moderately helpful in evaluating the impact of scoliosis surgery in neuromuscular patients. The expectation and satisfaction sections are also very limited, and many investigators question the validity of these domains.

Pencharz et al[26] prospectively compared the distributions, construct validity, and discriminative ability of the activity scale for kids (ASK), Child Health Questionnaire Parent Form (CHQ-PF 28), and the PODCI. They administered these 3 self-reported outcome measures to 210 children and parents of children with musculoskeletal disorders between the ages of 5 and 17 years, including patients with more severe disabilities such as CP and Duchenne muscular dystrophy. Seventy-nine percent (166 patients) completed the questionnaires. They found that the PODCI and ASK had good correlation, whereas the CHQ-PF 28 had weaker correlations with the ASK and the PODCI. They also found that the CHQ-PF 28 had less discriminatory ability between different diseases than the PODCI or ASK. This was thought to be due to the generic nature of the items in the CHQ-PF 28. The CHQ-PF 28 was originally designed to measure a general pediatric population with a variety of illnesses and disabilities. Although there was a high correlation between the ASK and the PODCI, examination of the "limits of agreement" between the 2 measures showed variance in score for individual children suggesting that they may be measuring different aspects of physical disability.

Despite its potential limits in measuring outcome in neuromuscular spinal deformity, Jones et al[27] have used the PODCI to evaluate short-term results of spinal fusion to treat spinal deformity in a group of totally involved patient with CP. Improvements in function were not demonstrated, but improvements in pain, comfort, frequency of feeling sick and tired, and satisfaction were all realized. The authors concluded that their prospective study confirmed the findings of previous retrospective studies in patients with CP with identifying improvements after surgery for spinal deformity.

Muscular Dystrophy and Flaccid Neuromuscular Conditions

Bridwell et al[28] evaluated a series of 48 patients with flaccid neuromuscular disorders (specifically Duchenne muscular dystrophy and spinal muscular atrophy) from 1985 to 1995. The purpose of the study was to evaluate whether patients and their caregivers thought that surgical treatment of scoliosis was beneficial in light of the general downhill course of these diseases. The authors constructed a questionnaire consisting of 20 questions focusing on several factors including function, pain, satisfaction, QOL, and cosmesis. Retest-retest reliability was assessed by testing and retesting 28 of their 48 subjects over a 6-month interval.

The questionnaire covered a range of issues specific for progressive flaccid neuromuscular scoliosis and included questions related to mobility, ease of care, seating difficulties, use of bracing, physical, and pulmonary function. The range of questions was broad and seemed to cover most of the concerns specific to neuromuscular patients (content validity). To determine disease specificity, this questionnaire was also administered to patients with idiopathic scoliosis. There was a statistically significant difference between the groups related to questions involving walking ability, pulmonary function, and perineal care, which are more common areas of difficulty for patients with neuromuscular scoliosis. No concurrent or construct validation for this questionnaire has been performed. Discriminate validity, which would require comparing a group with the disease and scoliosis against a group with the disease and no scoliosis has not been assessed. Furthermore, the questionnaire addresses responsiveness by structuring particular questions in 2 parts; first asking about a specific variable and then how this variable had changed after the intervention. Ultimately, responsiveness could not be established because the questionnaire was administered retrospectively with probable recall bias. Nevertheless, this is a promising questionnaire and further work continues to refine it.

Wright et al[29] recently developed a Muscular Dystrophy Spine Questionnaire. They included parent, patient, and clinician input for item development and ranked the importance of questions, then narrowed down the questionnaire to the 28 most important items. They

demonstrated test/retest repeatability. Construct validity was established by correlating scores to pulmonary function test results (forced vital capacity measurements) and to PODCI and ASK scores. To date, the MDSQ has not been used to measure the responsiveness in treatment of spinal deformity. It, too, shows promise in quantifying outcomes in patients with Duchenne muscular dystrophy.

Myelomeningocele

Wai et al[30] sought out to develop a reliable instrument to evaluate physical disability from scoliosis as they relate to children with spina bifida and their families. The study population included English speaking children between the ages of 7 and 16 years with spina bifida cystica and scoliosis (defined as at least 10 degrees fixed lateral curvature). Children who had spine surgery within the last year before the study were excluded. The result of their efforts is the Spina Bifida Scoliosis Questionnaire (SBSQ).

Question content was established by reviewing the literature and interviewing content experts including orthopaedic surgeons, registered nurses, physical therapists, and occupational therapists. From these interviews they developed a pool of questions related to sitting, walking, dressing ability, and other functional tasks that might be affected and related to QOL in patients with spina bifida. In addition, families of patients with this disorder generated a list of items related to function and answered open-ended questions also related to function and QOL. Test-retest reliability was analyzed using the ICC that was found to be 0.88. (ICC ranges between 0 and 1, >0.75 indicate good to excellent reliability.) Concurrent and construct validity in this study was done by comparison with the ASK. Pearson correlation coefficient was used to test this relationship. Excellent correlation between the ASK and the SBSQ was obtained ($r = 0.86$.) The SBSQ correlated marginally better with parents and children's assessment of disability related to the back but as expected the ASK correlated better with overall global function of the child.

The investigator also tried to establish that the SBSQ could distinguish between children with different sitting and ambulation abilities. The 1-way analysis of variance found the mean SBSQ scores for sitters and ambulators were significantly different. In conclusion, the SBSQ shows promise as an outcome instrument for patients with spina bifida and scoliosis but at this time no clinical study using this questionnaire has shown its ability to demonstrate a change after scoliosis surgery. The SBSQ has been used by the authors to suggest that scoliosis surgery in this population does not improve QOL.[31]

CP

The Caregiver's Priorities and Child Health Index of Life with Disabilities (CPCHILD) Questionnaire was developed to evaluate interventions in the management of patient's with CP.[32] The goal was to measure outcome as it relates to the improvement of QOL and where possible, the function of these patients as opposed to radiologic and clinical outcomes. The developers proposed to establish that CPCHILD is a reliable, valid proxy measure of functional and health status, caregiver burden, and HRQOL in children with severe CP (Gross Motor Function Classification System class 4 and 5). In this study, primary caregivers were defined as the person most responsible for the care of the child and had lived with the child for at least 6 months prior the administration of the questionnaire.

The population studied were children with CP aged 5 to 18 years (N = 77) recruited from spasticity and orthopaedic clinics. Items were generated from interviews with health care providers and caregivers of the study population. Caregivers rated the importance of each item's contribution to their child's QOL on a 6-point scale. None of the items were rated below the threshold. Items related to comfort and communication were considered most important. The final questionnaire consisted of 36 items addressing 6 main areas: (1) personal care; (2) positioning, transfer, and mobility; (3) communication and social interaction; (4) comfort, emotions, and behavior; (5) health; and (6) quality of life. Reliability was assessed by test-retest of caregivers of patients (N = 52) with severe CP over a 2-week interval. Forty-one reported no change over the interval. The ICC was used to assess reliability and found to be 0.97 (95% confidence interval, 0.95-0.99) with a range from 0.88 to 0.96 for the 6 categories in the questionnaire. Thus, the questionnaire demonstrated excellent reliability.

TABLE 3. Commonly Used Disease-specific Patient-based Outcomes Questionnaires for Use in Neuromuscular Spinal Deformity Patients

Study	Diagnosis	Name of Questionnaire	Valid?	Reliable?	Responsive to Scoliosis?
Bridwell et al[28]	Flaccid neuromuscular conditions	NA	Yes*	Yes	Unknown
Wright et al[29]	Duchenne muscular dystrophy	Muscular Dystrophy Spine Questionnaire	Yes	Yes	Unknown
Wai et al[30]	Myelomeningocele	SBSQ	Yes	Yes	Unknown
Narayanan et al[16]	Cerebral palsy	CPCHILD	Yes	Yes	Unknown
Watanabe et al[33]	Cerebral palsy	NA	Yes*	Unknown†	Unknown

*Used some questions from previously validated questionnaires.
†Previous version of questionnaire was reliable.
CPCHILD indicates Caregiver's Priorities and Child Health Index of Life with Disabilities; NA, not available; SBSQ, Spina Bifida Scoliosis Questionnaire.

Construct (convergent) validity was achieved by administering the CPCHILD to ambulatory children with CP. It was observed that the mean score for non-ambulatory children was significantly worse than those who were ambulatory. Further validation is being performed to compare CPCHILD with other validated measures such as the CHQ-PF50 and caregiver assistance scales of the PEDI. In conclusion, this questionnaire was developed to evaluate the general health and function of children with severe CP. Although content validity was good, questions specifically addressing the impact of spinal deformity on these children, for example, back discomfort/pain, pulmonary function, use of brace and improvement with surgical intervention would be required to increase the validity in evaluating outcome in neuromuscular scoliosis. Nevertheless, CPCHILD is currently undergoing further validation to establish generalizability and responsiveness.

Watanabe et al[33] performed a retrospective clinical outcome study in patients with CP with a neuromuscular patient-evaluation questionnaire. Satisfaction was high (92%). Sitting balance and cosmesis were >90%; QOL improved in 71%. Functional improvements were more limited (8% to 40%), though there was functional improvement after surgery. Patients less satisfied had more complications, less correction, or the presence of hyperlordosis in the lumbar spine postoperatively. The outcome instrument used was a modified version of a validated questionnaire developed for patients with Duchenne muscular dystrophy by the same authors. The same problems exist with this questionnaire as with the previous one (concurrent or construct validation, discriminate validity, and responsiveness). Nevertheless, the questions contained in the outcome instrument have face validity and represent an attempt to measure QOL from a caregiver's perspective in this patient population.

The disease-specific patient-evaluation questionnaires appropriate for patients with chronic neuromuscular conditions and spinal deformity are summarized in Table 3.

SUMMARY

There is no gold standard for outcome assessment for patients with neuromuscular scoliosis. Patient-based measures are but one of many assessments that should be used in the assessment of surgical results in these patients. Neuromuscular scoliosis represents a diverse group of patients that frequently suffer from progressive spinal deformity, which affects HRQOL in a diagnosis-specific manner. Although desirable, development of a single patient-based outcome measure that covers all aspects of neuromuscular spinal deformity is unlikely to occur because of disease heterogeneity. Two potential strategies for the future are to continue the development and validation of diagnosis-specific spinal deformity questionnaires. Another would be to use elements from existing instruments that are generalizable across different neuromuscular diagnoses and create an instrument that could be used in a wider variety of patients. Studies in these patients have yet to adequately document preoperative to postoperative changes in disease-specific health status. Nevertheless, given the costs and potential complications resulting from the intervention, the medical community needs to continue the pursuit of quantifying the impact of spinal deformity surgery on patients with neuromuscular conditions. An important step in this direction is the universal acceptance of established outcomes instruments. Further work should examine the incremental cost effectiveness of treatment modalities with a long-term perspective.

REFERENCES

1. Fisher CG, Wood KB. Introduction to and techniques of evidence-based medicine. *Spine*. 2007;32(suppl):S66–S72.
2. Suk M, Norvell DC, Hanson B, et al. Evidenced-based orthopaedic surgery: what is evidence without the outcomes? *J Am Acad Orthop Surg*. 2008;16:123–129.
3. Sarwark J, Sarwahi V. New strategies and decision making in the management of neuromuscular scoliosis. *Orthop Clin North Am*. 2007;38:485–496, v.
4. Kinali M, Main M, Eliahoo J, et al. Predictive factors for the development of scoliosis in Duchenne muscular dystrophy. *Eur J Paediatr Neurol*. 2007;11:160–166.
5. Majd ME, Muldowny DS, Holt RT. Natural history of scoliosis in the institutionalized adult cerebral palsy population. *Spine (Phila Pa 1976)*. 1997;22:1461–1466.
6. Müller EB, Nordwall A, Odén A. Progression of scoliosis in children with myelomeningocele. *Spine (Phila Pa 1976)*. 1994;19:147–150.
7. Saito N, Ebara S, Ohotsuka K, et al. Natural history of scoliosis in spastic cerebral palsy. *Lancet*. 1998;351:1687–1692.
8. Trivedi J, Thomson JD, Slakey JB, et al. Clinical and radiographic predictors of scoliosis in patients with myelomeningocele. *J Bone Joint Surg Am*. 2002;84:1389–1394.
9. Yamashita T, Kanaya K, Kawaguchi S, et al. Prediction of progression of spinal deformity in Duchenne muscular dystrophy: a preliminary report. *Spine (Phila Pa 1976)*. 2001;26:E223–E226.
10. Centers for Disease Control and Prevention. Improved National Prevalence Estimates for 18 Selected Major Birth Defects – United States, 1999–2001. *Morbidity and Mortality Weekly Report*. 1995;44:694–699.
11. Barsdorf AI, Sproule DM, Kaufmann P. Scoliosis surgery in children with neuromuscular disease: findings from the US National Inpatient Sample, 1997 to 2003. *Arch Neurol*. 2010;67:231–235.
12. Cassidy C, Craig CL, Perry A, et al. A reassessment of spinal stabilization in severe cerebral palsy. *J Pediatr Orthop*. 1994;14:731–739.
13. Kalen V, Conklin MM, Sherman FC. Untreated scoliosis in severe cerebral palsy. *J Pediatr Orthop*. 1992;12:337–340.
14. Mohamad F, Parent S, Pawelek J, et al. Perioperative complications after surgical correction in neuromuscular scoliosis. *J Pediatr Orthop*. 2007;27:392–397.
15. Mercado E, Alman B, Wright JG. Does spinal fusion influence quality of life in neuromuscular scoliosis? *Spine*. 2007;32(suppl):S120–S125.
16. Narayanan UG, Fehlings D, Weir S, et al. Initial development and validation of the Caregiver Priorities and Child Health Index of Life with Disabilities (CPCHILD). *Dev Med Child Neurol*. 2006;48:804–812.
17. Van Heest AE, Ramachandran V, Stout J, et al. Quantitative and qualitative functional evaluation of upper extremity tendon transfers in spastic hemiplegia caused by cerebral palsy. *J Pediatr Orthop*. 2008;28:679–683.
18. Streiner DL, Norman GR. *Health Measurement Scales: A Practical Guide to the Development and Use*. 2nd ed. Oxford, England: Oxford University Press; 2003.
19. Munro BH. *Statistical Methods For Health Care Research*. 3rd ed. Philadelphia: Lippincott-Raven; 1997:224–245.
20. Landis JR, Koch GG. An application of hierarchical kappa-type statistics in the assessment of majority agreement among multiple observers. *Biometrics*. 1977;33:363–374.

21. Garbuz DS, Masri BA, Esdaile J, et al. Classification systems in orthopaedics. *J Am Acad Orthop Surg*. 2002;10:290–297.
22. Daltroy LH, Liang MH, Fossel AH, et al. The POSNA pediatric musculoskeletal functional health questionnaire: report on reliability, validity and sensitivity to change. *J Pediatr Orthop*. 1998;18:561–571.
23. Lerman JA, Sullivan E, Haynes RJ. The Pediatric Outcomes Data Collection Instrument (PODCI) and functional assessment in patients with adolescent or juvenile idiopathic scoliosis and congenital scoliosis or kyphosis. *Spine*. 2002;27:2052–2057.
24. Damiano DL, Gilgannon MD, Abel MF. Responsiveness and uniqueness of the pediatric outcomes data collection instrument compared to the gross motor function measure for measuring orthopaedic and neurosurgical outcomes in cerebral palsy. *J Pediatr Orthop*. 2005;25:641–645.
25. McMulkin ML, Baird GO, Gordon AB, et al. The pediatric outcomes data collection instrument detects improvements for children with ambulatory cerebral palsy after orthopaedic intervention. *J Pediatr Orthop*. 2007;27:1–6.
26. Pencharz J, Young NL, Owen JL, et al. Comparison of three outcomes instruments in children. *J Pediatr Orthop*. 2001;21:425–432.
27. Jones KB, Sponseller PD, Shindle MK, et al. Longitudinal parental perceptions of spinal fusion for neuromuscular spine deformity in patients with totally involved cerebral palsy. *J Pediatr Orthop*. 2003;23:143–149.
28. Bridwell KH, Baldus C, Iffrig TM, et al. Process measures and patient/parent evaluation of surgical management of spinal deformities in patients with progressive flaccid neuromuscular scoliosis (Duchenne's muscular dystrophy and spinal muscular atrophy). *Spine*. 1999;24:1300–1309.
29. Wright JG, Smith PL, Owen JL, et al. Assessing functional outcomes of children with muscular dystrophy and scoliosis: the Muscular Dystrophy Spine Questionnaire. *J Pediatr Orthop*. 2008; 28:840–845.
30. Wai EK, Owen J, Fehlings D, et al. Assessing physical disability in children with spina bifida and scoliosis. *J Pediatr Orthop*. 2000;20: 765–770.
31. Wai EK, Young NL, Feldman BM, et al. The relationship between function, self-perception, and spinal deformity: implications for treatment of scoliosis in children with spina bifida. *J Pediatr Orthop*. 2005;25:64–69.
32. Narayanan UG, Fehlings DL, Weirs S, et al. Caregiver Priorities and Child Health Index of Life with Disabilities: initial development and validation of an outcome measure of health status and well being in children with severe cerebral palsy. *Dev Med Child Neurol*. 2006;48:804–812.
33. Watanabe K, Lenke LG, Daubs MD, et al. Is spine deformity surgery in patients with spastic cerebral palsy truly beneficial?: a patient/parent evaluation. *Spine (Phila Pa 1976)*. 2009;34:2222–2232.

