J Neurosurg (Pediatrics 3) 103:207–212, 2005

Potentially useful outcome measures for clinical research in pediatric neurosurgery

Paul Klimo Jr., M.D., M.P.H., and John R. W. Kestle, M.D., M.Sc.

Department of Neurosurgery, Division of Pediatric Neurosurgery University of Utah, Salt Lake City, Utah

✓The choice of outcome (or outcomes) and their measurement are critical for a sound clinical trial. Surgeons have traditionally measured simple outcomes such as death, duration of survival, or tumor recurrence but have recently developed more sophisticated measures of the effect of an intervention. Many outcome measures require a lengthy maturation process, which includes a determination of the instrument's validity, reliability, and sensitivity; thus, using established instruments rather than creating new ones is recommended. The authors illustrate several guidelines for the determination of appropriate outcome measures by using examples from their experience and describe several outcome measures that can be used in pediatric neurosurgery. These include general outcome measures such as the Pediatric Evaluation of Disability Inventory and the Functional Independence Measure for Children, which measure physical function and independence in chronically ill and disabled children as well as disease-specific measures for hydrocephalus (Hydrocephalus Outcome Questionnaire), cerebral palsy (gross motor function and performance measures), head injury (Pediatric Cerebral Performance Category and Children's Coma Scale), and oncology (Pediatric Cancer Quality-of-Life Inventory).

Key Words • outcome research • outcome measure • assessment • pediatric neurosurgery

Clinical research is an integral part of current neurosurgical practice. One of the most important components of a clinical research project is the choice of the outcome and the way in which it is measured. In some areas of neurosurgical interest, the outcome is obvious (for example, survival time), but in others (for example, spasticity), the development and testing of an outcome measure have required years of work and investigation.[40] In this article, we describe the parameters by which outcome measures are assessed and the process of developing an outcome measure when an appropriate one is not available. To assist clinical investigators in pediatric neurosurgery, we also describe a number of outcome measures that may be useful to them.

Outcome Measures

An outcome measure refers to any tool used to evaluate the effect of a disease and/or treatment on the health status of a patient. Traditionally, surgeons have been interested in outcomes that represent pathological, physiological, or clinical variables, such as death, pain, tumor growth, or bone fusion. As the science of outcome measurement has evolved, two advances have emerged. First, scales have been developed to improve the precision of the "traditional" outcome measures. For example, pain can be scored in a reproducible way, rather than by being reported as "improved" or "unchanged." Ambulation can be graded (American Spinal Injury Association motor score) rather than being recorded as present or absent. Second, "nontraditional" outcome measures, especially those that attempt to capture information on other domains within the patient's life, have been developed.

Depending on the study, the outcome(s) of interest may be related to clinical signs and symptoms, imaging findings, functional abilities (physical, cognitive, social), survival, quality of life, patient satisfaction, and/or cost of care. Often there are several outcomes of interest, but every attempt should be made to identify a single primary outcome. This focus will allow an investigator to define a specific study hypothesis, calculate a sample size (which will determine budget and feasibility), and draw clear, specific conclusions.

Many factors should be considered when selecting a primary outcome, depending on the study question, but for clinical questions the following two principles should be kept in mind: 1) What is important to the patient? To answer this, patients or patient advocacy groups should ideally be involved in the decision, because their opinions may not always be obvious. 2) What would make me change my practice? In other words, is the outcome of sufficient importance to justify the risk, convenience, and cost of the treatment?

Abbreviation used in this paper: GCS = Glasgow Coma Scale.

Evaluating and Designing Outcome Measures

Designing, testing, and implementing outcome measures other than simple dichotomous variables, such as death, is a lengthy and complex process.[20,44,50] After deciding which outcomes should be measured, the clinical investigators should review the relevant literature to find outcome measures that have already been designed and evaluated by others. This approach saves considerable time and effort and can make the results of the study more comparable with other studies that use similar outcome measures.

When selecting an outcome measure from the literature (or developing a new one), three important characteristics should be considered: reliability, validity, and responsiveness. Reliability is the extent to which the same result is obtained when a measure is applied more than once. Interrater reliability refers to the degree of agreement between different observers at the same point in time. Intrarater reliability measures the degree of agreement by the same observer at different points in time. Reliability is usually assessed first in the development of an outcome measure, because other parameters are irrelevant if the measure is not reliable. The validity of an outcome measure is the extent to which it measures what it is supposed to measure. Face validity is a subjective judgment about whether a measurement makes sense intuitively (that is, does it appear to measure what it is intended to measure?). Content validity is a judgment about whether the instrument samples all of the relevant or important content or domains.

Demonstrating validity requires more than peer judgments, however. If an accepted outcome measure already exists (a gold standard), then one can administer it along with the new one and measure the correlation between the two. This pairing is known by various terms such as convergent, criterion, or concurrent validity. When no gold standard exists, one must link the measure under investigation to some other attribute by a hypothesis or construct. This paradigm is called construct validity (for example, if a new scale is being developed to measure the severity of subarachnoid hemorrhage, one would expect it to be correlated to the patients' Hunt and Hess grades).

Responsiveness is the ability of an instrument to measure a meaningful or clinically important change in a clinical state. Responsiveness is particularly important if a measure is used to follow individual patients over time. We refer readers to the excellent article by Cohen and Marino[15] and other articles that apply these concepts to functional and other outcome measures.[2,3,19,38]

Health status measures can be either generic or specific.[44] Generic or global outcome measures are used to assess health status across disease states but may not specific enough to detect important changes in a particular condition. They can be viewed as health profiles and are designed to be applicable to a broad range of different interventions, diseases, and populations. Generic measures often have well-established reliability and validity; however, they may not be as responsive as disease-specific measures and may not focus on the more important outcomes of a particular disease. Conversely, disease-specific outcome measures are designed to concentrate attention on domains specific to a particular disease, population, or area of functioning.[25] Their advantages are improved responsiveness and greater willingness of physicians to use them in clinical studies. Their primary disadvantage is in their inherently limited applicability.

In addition to these parameters, the following questions should be considered when choosing or developing an outcome measure: 1) Is the measure acceptable to patients (that is, what is the burden to the patients)? For example, if it is a questionnaire, will they complete it or is it too long, complex, or boring? 2) Does it matter to patients? Is it something that will make a difference in their survival, health, or quality of life? 3) Is it feasible (that is, what is the administrative burden)? Can I assess this outcome in my practice? If the outcome is supposed to be measured after 5 years of follow up, will I be able to track the patients for that long? 4) Does it matter to healthcare professionals? Will the results, based on this outcome measure, result in a change in clinical practice?

Avoiding Observer Bias: An Example From Hydrocephalus Research

In the Shunt Design Trial and the Endoscopic Shunt Insertion Trial, the outcome of interest was shunt failure.[31,32] Initially, shunt failure was to be defined simply as the need for a shunt revision, but the following two potential problems were identified with this plan: 1) Different surgeons have different indications for reoperation. 2) The surgeons were not blinded to the study. This setup resulted in a potential observer bias, which occurs when the person assessing the outcome knows which treatment was given. This knowledge can influence their assessment of the outcome. Because of this potential observer bias, "need for reoperation" would not be an appropriate definition of the primary outcome in the Shunt Design Trial and the Endoscopic Shunt Insertion Trial.

The literature at the time did not include a definition of shunt failure. It appeared that shunt failure and shunt revision surgery were synonymous. We therefore developed definitions by consensus among the investigators on the basis of history, physical findings, and ancillary tests. Separate definitions were developed for shunt obstruction, shunt overdrainage, loculated ventricles, and shunt infection.[17] At the time of follow up, if the patient met any of the definitions in a blinded review of a follow-up examination, he or she was said to have reached the selected primary study end point: shunt failure.

To remove the potential observer bias, the presence or absence of shunt failure was blindly reviewed in all study patients. When a patient was seen at follow up, the clinical notes, data forms, and imaging studies were copied and sent to the data coordinating center, where they were blinded (patient names, center names, and valve information were removed). The blinded material was then reviewed by two of the investigators and, when necessary to resolve differences, by an independent committee. The blinded review process determined whether the patient met the definition of the study end point (shunt failure) and was used for the primary analysis and conclusions of the study. This process also allowed an assessment of whether the surgeons' assessments of outcome were biased. The shunt survival curves based on the blinded review were compared with the curves based on the surgeons' unblinded decision to operate or not. For each of the three valves, the shunt survival rate based on the surgeons' decision appeared slight-

ly better, but the differences were not large enough to alter the primary study conclusions.[30] Although the adjudication process was laborious, it ensured that observer bias did not influence the results. This is important because observer bias is always a potential threat, and its magnitude and direction are not predictable before a study.

Outcome Measures Used in Pediatric Neurosurgery

Researchers face several methodological challenges when developing and applying outcome measures in pediatrics.[14,24,29,39,42] Any study of outcomes in pediatric patients must account for the physical growth from infancy through adolescence and the changes in behavioral, social, and psychological development. An outcome measure appropriate for a teenager is unlikely to be appropriate for a toddler or baby. Thus, outcome measures should be chosen to be specific to each stage of child development (infancy, early childhood, late childhood, and adolescence). Most of the validated health status measurement scales in the literature have been developed for adults. Some have been modified from their adult form to be used in children, and a few have been developed specifically for children. In a review of the literature, Forrest and colleagues[23] found few studies that examined the health effects of preventive, diagnostic, long-term management, and curative services delivered to children and adolescents. Although much progress has been made in the field of pediatric outcomes research, the unique challenges with pediatrics have prevented widespread use of these instruments.[24]

Examples of Outcome Measures

Considering the work required to develop a validated outcome measure from scratch, an investigator's first step should be to look for an established one that might be useful. Several outcome measures that are potentially useful in pediatric neurosurgery are outlined.

Generic Outcome Measures

Survival. Survival is the primary outcome in most oncology and trauma studies. For conditions in which death is unlikely, this measure is not sensitive enough to detect important changes, thus other clinical outcomes are more important.

Pediatric Evaluation of Disability Inventory. This scale measures physical function and independence in chronically ill and disabled children by assessing self-care mobility and social function. Its reliability and validity in children whose ages fall between 6 months and 7 years have been thoroughly evaluated.[21,28,41,52]

Functional Independence Measure for Children. This scale requires a trained observer to assess the degree of assistance required by disabled patients. It was originally developed for adults, but the reliability of a modification[26] has been demonstrated for use in nondisabled children 6 months to 8 years of age and in developmentally disabled children aged 6 months to 12 years. It has been applied to children with spina bifida.

Motor Control Assessment. This measurement requires a clinician to assess motor control skills. It has been validated in children between 2 and 5 years of age with mild-to-severe physical disability and has good demonstration of validity and reliability.[47]

Short Form–36. This popular self-administered questionnaire measures physical and social function, role limitations caused by physical or emotional problems, and general health perception. Its advantage is its wide use in many diseases, but it is limited in pediatric use to adolescents and older teenagers.[9]

Sickness Impact Profile. This 136-item scale was designed for use across many disorders to measure perceived health status.[10] It was not designed for use in children but may be suitable for use with teenagers.

Rand Health Status Measure for Children. This was the first published attempt to assess children's health-related outcomes on a large scale.[18] It is designed to evaluate four core domains, as well as general health perception and behavior problems, through questions to the parents. There are versions for children 0 to 4 years of age and those 5 to 13 years of age.

Functional Status II-R. This measure evaluates parents' perceptions of the impact of their child's (age 0–16 years) illness on physical, social, and psychological functions.[48] An item is scored as having an impact on functional status if the parent states that the problem results from the child's current illness.

Child Health and Illness Profile-Adolescent Edition. This is the first instrument that deals specifically with health assessment in adolescents (11–17 years of age).[52] It assesses comfort/discomfort, satisfaction with health, risk, disorder, achievement of social expectations in major role activities, and resilience. It has been shown to be reliable and to have construct validity.[52]

Other global measures of health outcomes addressing children include the School Function Assessment, Child Health Questionnaire, Questionnaire for Identifying Children with Chronic Conditions, Abilities Index, Youth Quality of Life Instrument-Research Version, and Quality of Well-Being Scale.

Disease-Specific Outcome Measures

Spasticity Assessment

Gross Motor Function Measure. This measure was designed by a trained observer specifically for children with cerebral palsy and has been used in surgical studies of spasticity.[49] It assesses gross motor skills and specifically measures the quantity of each particular skill, not the quality.[46]

Gross Motor Performance Measure. A trained clinician can assess the quality of movement of patients with cerebral palsy in terms of alignment, coordination, dissociated movements, stability, and weight transfer.[11] It has also undergone validity and reliability testing.

Ashworth Scale. This commonly used simple scale grades the degree of spasticity in an extremity from 0 to 4.[4] It has a high level of interrater reliability.[7]

Epilepsy Assessment

Quality of Life in Newly Diagnosed Epilepsy Instrument. This quality-of-life measure was developed for use with recently diagnosed epilepsy patients.[1] It has good validity

and reliability in adults and may be appropriate in older children or teens with epilepsy.

Seizure Severity and Side Effects Scales for Childhood Epilepsy. These scales are completed by the parents of children 4 to 16 years of age with chronic epilepsy as well as children with neurological comorbidity.[12] In a group of 80 children, the scales showed good internal consistency and good test–retest reliability.

Adolescent Psychosocial Seizure Inventory. This outcome instrument assesses the psychosocial problems of adolescent patients with epilepsy analogous to the Washington Psychosocial Seizure Inventory used in adults.[6]

Head Injury

Several outcome measures have been used in clinical trials of head injury. In 1991, recommendations for outcome measures in traumatic brain injuries were issued. These recommendations were not made for patients younger than 16 years of age but may be applicable to older children.[11] For studies of severe brain injury (GCS Scores 3–8), the Glasgow Outcome Scale or the Disability Rating Scale measured 6 months after the injury were recommended as the primary outcome measures. In moderate brain injury (GCS Scores 9–12), the Disability Rating Scale at 3 months after injury was recommended as the primary outcome. In addition, the following neuropsychological tests were recommended on the basis of their previous use and validity in brain injury research, their relevance to quality of life and capability of returning to work, and their brevity: Digital Symbol Substitution, Paced Auditory Serial Addition, Rey Complex Figure, Selective Reminding, Controlled Oral Word Association, Trail Making B, Wisconsin Card Sorting, Grooved Peg Board, and Neurobehavioral Rating Scale.

Scales more specific to children include the following:

Pediatric Cerebral Performance Category. This easy-to-complete scale provides useful information regarding probable outcomes of patients in the pediatric intensive care unit.[22] The scores in 200 children had good correlation with Stanford Binet Intelligence scores, the Bayley Scales of Infant Development, and the Vineland Adaptive Behavior Scale.

Children's Coma Scale. In this scale, the best verbal response portion of the GCS is modified for children.[27] Pediatric scores can be compared directly with those from the adult population. Reliability and validity data are not available. Other variations of the GCS include the Children's Coma Score by Raimondi, the Pediatric Coma Scale, and the Jacobi Comatose Management Score.

King's Outcome Scale for Childhood Head Injury. Another pediatric (patient age range 2–16 years) adaptation of the original Glasgow Outcome Scale, this measure has increased sensitivity at the milder end of the disability range. Validity and reliability testing have been reported.[16]

Outcome Score According to Kriel. This scale has two parts: cognitive (0–4 points) and motor (0–5 points) scores. This score is useful in neurotrauma studies with extended follow-up periods.[14]

Risk Score for Posttraumatic Epilepsy in Childhood. This score assesses the risk of development of posttraumatic epilepsy in children based on mechanism of injury, type of injury, and age.[33]

Outcome of Myelomeningocele

Spina Bifida Health-Related Quality of Life Instrument. Reliability of this instrument for patients with spina bifida was good, but validity was better in older children than in younger ones. Further testing of responsiveness is underway.[43]

Outcome of Hydrocephalus

Hydrocephalus Outcome Questionnaire. This 10- to 15-minute parent questionnaire measures quality of life in children 5 years of age and older with hydrocephalus. Excellent reliability and validity have been demonstrated.[36,37]

Imaging Measurements. Frontal and occipital horn ratio is a measurement of ventricular size on axial brain images (computerized tomography or magnetic resonance imaging) that more accurately reflects ventricular volume than the Evans ratio does. Its interrater reliability is very good. It is obtained by averaging the widest distance across the frontal and occipital horns and then dividing it by the largest biparietal diameter. Measurements can be taken from different slices to determine the maximum for each item in the equation.[35]

Pain Rating

Children's Global Rating Scale. This scale specifically for children measures their self-report of pain and fear. Its convergent and predictive validities have been demonstrated in an assessment of 145 children undergoing phlebotomy.[13]

Oncology Scales

The Pediatric Cancer Quality-of-Life Inventory. This instrument is not specific to neurooncology but is intended for use in the general population of pediatric cancer patients. Initial testing is promising, but further assessment of validity and reliability is required.[51]

Quality of Life Assessment Form—Children. This scale assesses the quality of life of children with cancer as reported by their parents.[8]

Neurological Severity Score.[5] This score assesses the neuropsychological and intellectual outcomes based on preoperative, perioperative, and postoperative medical data obtained from children with intracranial tumors, especially astrocytomas. It has shown good correlation with other neuropsychological scores and is a good tool for investigating the psychological impact of brain tumors and their treatment.

Conclusions

Any study, from a simple case series to a randomized multicenter study, should have well-defined outcomes that are clearly stated before the data are collected. Until recently, the outcomes of interest to most surgeons were those that were directly affected by the disease or the intervention, such as survival, recurrence, complications, and extent of resection. In an effort to understand the im-

pact of the disease and treatment on children better, there has been an effort to incorporate more elaborate outcome measures, especially ones derived from the patients or their families.

The process of designing and evaluating outcome scales is complex and has multiple stages. Each outcome measure undergoes a process of reevaluation until it becomes accepted as a useful measure. Despite the numerous methodological obstacles presented when designing and implementing outcome measures in the pediatric population, numerous generic and specific outcome measures can be applied in the field of pediatric neurosurgery.

References

1. Abetz L, Jacoby A, Baker GA, McNulty P: Patient-based assessments of quality of life in newly diagnosed epilepsy patients: validation of the NEWQOL. **Epilepsia 41:**1119–1128, 2000
2. Andresen EM: Criteria for assessing the tools of disability outcomes research. **Arch Phys Med Rehabil 81 (12 Suppl 2):** S15–S20, 2000
3. Andresen EM, Meyers AR: Health-related quality of life outcomes measures. **Arch Phys Med Rehabil 81 (12 Suppl 2):** S30–S45, 2000
4. Ashworth B: Preliminary trial of carisoprodol in multiple sclerosis. **Practitioner 192:**540–542, 1964
5. Ater JL, Moore BD III, Francis DJ, Castillo R, Slopis J, Copeland DR: Correlation of medical and neurosurgical events with neuropsychological status in children at diagnosis of astrocytoma: utilization of a neurological severity score. **J Child Neurol 11:**462–469, 1996
6. Batzel LW, Dodrill CB, Dubinsky BL, Ziegler RG, Conolly JE, Freeman RD, et al: An objective method for the assessment of psychosocial problems in adolescents with epilepsy. **Epilepsia 32:**202–211, 1991
7. Bohannon RW, Smith MB: Interrater reliability of a modified Ashworth scale of muscle spasticity. **Phys Ther 67:**206–207, 1987
8. Bowling A: Cancers, in **Measuring Disease: A Review of Disease-Specific Quality of Life Measurement Scales, ed 2.** Philadelphia: Open University Press, 2001, pp 23–68
9. Bowling A: A popular core (generic) measure for use in disease-specific studies of adults: The Short Form-36 (SF-36), in **Measuring Disease: A Review of Disease-Specific Quality of Life Measurement Scales, ed 2.** Philadelphia: Open University Press, 2001, pp 298–302
10. Bowling A: Rheumatological conditions, in **Measuring Disease: A Review of Disease-Specific Quality of Life Measurement Scales, ed 2.** Philadelphia: Open University Press, 2001, pp 221–246
11. Boyce WF, Gowland C, Hardy S, Rosenbaum PL, Lane M, Plews N, et al: Development of a quality-of-movement measure for children with cerebral palsy. **Phys Ther 71:**820–832, 1991
12. Carpay HA, Arts WF, Vermeulen J, Stroink H, Brouwer OF, Peters AC, et al: Parent-completed scales for measuring seizure severity and severity of side-effects of antiepileptic drugs in childhood epilepsy: development and psychometric analysis. **Epilepsy Res 24:**173–181, 1996
13. Carpenter PJ: New method for measuring young children's self-report of fear and pain. **J Pain Symptom Manage 5:**233–240, 1990
14. Christakis DA, Johnston BD, Connell FA: Methodologic issues in pediatric outcomes research. **Ambul Pediatr 1:**59–62, 2001
15. Cohen ME, Marino RJ: The tools of disability outcomes research functional status measures. **Arch Phys Med Rehabil 81 (12 Suppl 2):**S21–S29, 2000
16. Crouchman M, Rossiter L, Colaco T, Forsyth R: A practical outcome scale for paediatric head injury. **Arch Dis Child 84:** 120–124, 2001
17. Drake JM, Kestle J: Rationale and methodology of the multicenter pediatric cerebrospinal fluid shunt design trial. Pediatric Hydrocephalus Treatment Evaluation Group. **Childs Nerv Syst 12:**434–447, 1996
18. Eisen M, Ware JE Jr, Donald CA, Brook RH: Measuring components of children's health status. **Med Care 17:**902–921, 1979
19. Epstein RS: Responsiveness in quality-of-life assessment: nomenclature, determinants and clinical application. **Med Care 38 (9 Suppl):**II91–II94, 2000
20. Erickson P: Assessment of the evaluative properties of health status instruments. **Med Care 38 (9 Suppl):**II95–II99, 2000
21. Feldman AB, Haley SM, Coryell J: Concurrent and construct validity of the Pediatric Evaluation of Disability Inventory. **Phys Ther 70:**602–610, 1990
22. Fiser DH, Long N, Roberson PK, Hefley G, Zolten K, Brodie-Fowler M: Relationship of pediatric overall performance category and pediatric cerebral performance category scores at pediatric intensive care unit discharge with outcome measures collected at hospital discharge and 1- and 6-month follow up assessments. **Crit Care Med 28:**2616–2620, 2000
23. Forrest CB, Shipman SA, Dougherty D, Miller MR: Outcomes research in pediatric settings: recent trends and future directions. **Pediatrics 111:**171–178, 2003
24. Forrest CB, Simpson L, Clancy C: Child health services research. Challenges and opportunities. **JAMA 277:**1787–1793, 1997
25. Gerszten PC: Outcomes research: a review. **Neurosurgery 43:** 1146–1156, 1998
26. Granger CV, Hamilton BB, Kayton R (eds): **Guide for Use of the Functional Independence Measure for Children (WeeFIM) of the Uniform Data Set for Medical Rehabilitation.** Buffalo, New York: Research Foundation, State University of New York, 1989
27. Hahn YS, Chyung C, Barthel MJ, Bailes J, Flannery AM, McLone DG: Head injuries in children under 36 months of age. Demography and outcome. **Childs Nerv Syst 4:**34–40, 1988
28. Haley SM, Coster WJ, Faas RM: A content validity study of the Pediatric Evaluation of Disability Inventory. **Pediatr Phys Ther 3:**177–184, 1991
29. Kaplan SH, Greenfield S, Connolly GA, Barlow SE, Grand R: Methodologic issues in the conduct and interpretation of pediatric effectiveness research. **Ambul Pediatr 1:**63–70, 2001
30. Kestle J, Milner R, Drake D: An assessment of observer bias in the shunt design trial. **Pediatr Neurosurg 30:**57–61, 1999
31. Kestle J, Milner R, Drake J: The shunt design trial: variation in surgical experience did not influence shunt survival. **Pediatr Neurosurg 30:**283–287, 1999
32. Kestle JR, Drake JM, Cochrane DD, Milner R, Walker ML, Abbott R III, et al: Lack of benefit of endoscopic ventriculoperitoneal shunt insertion: a multicenter randomized trial. **J Neurosurg 98:**284–290, 2003
33. Kieslich M, Jacobi G: Incidence and risk factors of post-traumatic epilepsy in childhood. **Lancet 345:**187, 1995
34. Kriel RL, Krach LE, Jones-Saete C: Outcome of children with prolonged unconsciousness and vegetative states. **Pediatr Neurol 9:**362–368, 1993
35. Kulkarni AV, Drake JM, Armstrong DC, Dirks PB: Measurement of ventricular size: reliability of the frontal and occipital horn ratio compared to subjective assessment. **Pediatr Neurosurg 31:**65–70, 1999
36. Kulkarni AV, Drake JM, Rabin D, Dirks PB, Humphreys RP, Rutka JT: Measuring the health status of children with hydrocephalus by using a new outcome measure. **J Neurosurg (Pediatrics 2) 101:**141–146, 2004
37. Kulkarni AV, Rabin D, Drake JM: An instrument to measure

the health status in children with hydrocephalus: the Hydrocephalus Outcome Questionnaire. **J Neurosurg (Pediatrics 2) 101:**134–140, 2004
38. Liang MH: Longitudinal construct validity: establishment of clinical meaning in patient evaluative instruments. **Med Care 38 (9 Suppl):**II84–II90, 2000
39. Lollar DJ, Simeonsson RJ, Nanda U: Measures of outcomes for children and youth. **Arch Phys Med Rehabil 81 (Suppl 2):** S46–S52, 2000
40. Masur H, Papke K, Althoff S, Oberwittler C: **Scales and Scores in Neurology. Quantification of Neurological Deficits in Research and Practice.** Stuttgart: Thieme, 2004
41. Nichols DS, Case-Smith J: Reliability and validity of the Pediatric Evaluation of Disability Inventory. **Pediatr Phys Ther 8:** 15–24, 1996
42. Palmer RH, Miller MR: Methodologic challenges in developing and implementing measures of quality for child health care. **Ambul Pediatr 1:**39–52, 2001
43. Parkin PC, Kirpalani HM, Rosenbaum PL, Fehlings DL, Van Nie A, Willan AR, et al: Development of a health-related quality of life instrument for use in children with spina bifida. **Qual Life Res 6:**123–132, 1997
44. Patrick DL, Chiang YP: Measurement of health outcomes in treatment effectiveness evaluations: conceptual and methodological challenges. **Med Care 38 (9 Suppl):**II14–II25, 2000
45. Patrick DL, Deyo RA: Generic and disease-specific measures in assessing health status and quality of life. **Med Care 27 (Suppl 3):**S217–S232, 1989
46. Russell DJ, Rosenbaum PL, Cadman DT, Gowland C, Hardy S, Jarvis S: The gross motor function measure: a means to evaluate the effects of physical therapy. **Dev Med Child Neurol 31:** 341–352, 1989
47. Steel KO, Glover JE, Spasoff RA: The motor control assessment: an instrument to measure motor control in physically disabled children. **Arch Phys Med Rehabil 72:**549–553, 1991
48. Stein RE, Jessop DJ: Functional status II(R). A measure of child health status. **Med Care 28:**1041–1055, 1990
49. Steinbok P, Reiner AM, Beauchamp R, Armstrong RW, Cochrane DD, Kestle J: A randomized clinical trial to compare selective posterior rhizotomy plus physiotherapy with physiotherapy alone in children with spastic diplegic cerebral palsy. **Dev Med Child Neurol 39:**178–184, 1997
50. Testa MA: Interpretation of quality-of-life outcomes: issues that affect magnitude and meaning. **Med Care 38 (9 Suppl):** II166–II174, 2000
51. Varni JW, Katz ER, Seid M, Quiggins DJ, Friedman-Bender A, Castro CM: The Pediatric Cancer Quality of Life Inventory (PCQL). I. Instrument development, descriptive statistics, and cross-informant variance. **J Behav Med 21:**179–204, 1998
52. Wright FV, Boschen KA: The Pediatric Evaluation of Disability Inventory (PEDI): validation of a new functional assessment outcome instrument. **Can J Rehabil 7:**41–42, 1993

Manuscript received October 12, 2004.
Accepted in final form April 28, 2005.
Address reprint requests to: John Kestle, M.D., Division of Pediatric Neurosurgery, Primary Children's Medical Center, 100 North Medical Drive, Salt Lake City, Utah 84113. email: john.kestle@hsc.utah.edu.