CONGRESS OF NEUROLOGICAL SURGEONS
2008 ANNUAL MEETING
ORLANDO, FLORIDA * SEPTEMBER 20–25, 2008

Join us in Orlando Florida for the 2008 CNS Annual Meeting, as we explore our theme The Community of Neurosurgery: Wisdom from Shared Experience.

Announcing the Joint Meeting with the Brazilian Society of Neurosurgery!

Make the 2008 CNS Annual Meeting YOUR Neurosurgical Meeting!
Share your expertise with the community of neurosurgery by participating in this open exchange of information, issues and policies of greatest importance to neurosurgery today.

Engage in interactive learning with eight new IML sessions on specialty-based clinical science topics. Share your practice preferences and knowledge through pre-meeting surveys giving you a chance to drive the session or course content and guide presenters in their delivery.

Interact directly with presenters and faculty. Use your own handheld devices or laptops to submit questions directly to the presenter or faculty during all General Scientific Sessions. The Select Abstracts session on Monday afternoon also lets you speak directly with the abstract authors to learn more about their methodologies and findings.

Contribute to a shared understanding and influence future neurosurgical initiatives by participating in one of five Consensus Conference sessions, designed to formulate key policy and strategies for neurosurgery.

Submit your cases for review. Our “Cases and Coffee with the Masters” sessions let you submit your difficult cases prior to the meeting. Those selected will be discussed in small, specialty-based groups led by master neurosurgeons within the specialty.

Register Online by August 20, 2008!
www.cns.org

For More Information
info@1cns.org
www.cns.org
Toll Free: 877 517 1CNS
Phone: 847 240 2500
Dear Colleagues,

This issue of the CNSQ focuses on neurosurgical outcomes. The subject of outcomes is multifaceted and is of increasing significance in medicine as society attempts to maximize health care resources. Outcomes analysis is also playing an important role in changing the landscape of neurosurgical patient management, regulation, reimbursement, and education. In our practices today we must consider different types of measurement tools to assess patient outcomes. These measures are becoming more significant in everything that we do. Analyzing performance, evaluating results, and determining the efficacy of treatment methods is a hot topic of debate that is becoming a standard—and sometimes controversial—part of our profession.

This issue of the CNSQ was spearheaded by our Associate Editor James Harrop. Articles by Beverly Walters and Robert Harbaugh provide an excellent overview and background information with regards to the science of outcomes, the variety of outcomes being studied, the types of trials and research methods, and importance of quality measures and how they impact us. Subsequent pieces review outcomes related elements associated with specific areas of neurosurgery, including traumatic brain injury, cerebrovascular, spine, oncology, pediatric, and functional.

Other highlights of this issue include an analysis of the Spine Patient Outcomes Research Trial (SPORT) results and how they may affect patient care. In addition, we have asked Kevin Gibbons, the lead neurosurgeon on the case, to provide his account of the highly publicized spinal cord injury sustained by the Buffalo Bills’ Kevin Everett.

As always, we welcome your contributions and appreciate your feedback on this issue of the CNSQ.

Sincerely,

Ali R. Rezai, MD

Editor Contact Information:
info@1cns.org or 847.240.2500
CONTENTS

04 Editor’s Note
Ali R. Rezai

06 Neurosurgeons and the Knowledge Economy
Anthony L. Asher

NEUROSURGICAL OUTCOMES

10 Evaluating Outcomes in Neurological Surgery: a Science of Measurement
Beverly C. Walters

14 Outcomes Analysis and Quality Improvement in Neurosurgery
Robert E. Harbaugh

18 Outcomes Studies in Cerebrovascular Neurosurgery
Carlos E. Sanchez, Bob S. Carter and Christopher S. Ogilvy

21 Outcomes Research in Spinal Surgery
Daniel K. Resnick

23 Outcomes in Neurosurgical Oncology
Ian F. Parney and Andrew E. Sloan

27 Outcomes after Traumatic Brain Injury
Emily Rogers and David O. Okonkwo

29 Outcomes in Pediatric Neurosurgery
Sarah J. Gaskill

31 Outcomes Studies in Functional Neurosurgery
Ron L. Alterman

FEATURES

36 Degenerative Spondylolisthesis with Lumbar Spinal Stenosis: an Evaluation of the SPORT Results
Peter D. Angevine, Michael G. Kaiser and Paul C. McCormick

42 How is the Hospital Profiting on Your Cases?
James I. Ausman

44 In Memoriam: Julius Goodman, MD

45 The Medical Care of Kevin Everett: Setting the Record Straight
Kevin J. Gibbons

MEMBERSHIP NEWS

48 CNS Bylaws: Proposed Amendments, Applications in Progress
Recently there has been an extraordinary outpouring of scientific work on the processes of thinking and learning. The Learning Sciences are an interdisciplinary movement that has, over the last 30 years, resulted in new methodologies, ideas, and ways of thinking about the process of learning. On the basis of this work, the very nature of learning and disciplinary knowledge is being redefined. In particular, learning is now viewed as an active, continuous process, the ultimate objective of which is the production and application of new knowledge. Although elements of modern learning theory are being increasingly applied to instructional situations in various disciplines and at all educational levels, and despite the fact that many of these techniques appear to more effectively enhance understanding and new knowledge generation, basic learning science principles have largely failed to inform the design of medical education programs. This may be to the detriment of both ourselves and our patients.

Most of our graduate and postgraduate medical education experiences continue to employ traditional didactic techniques which are modeled on the predominant educational paradigm of the 20th century, known as Instructionism. This pedagogical approach was grounded in philosophies such as logical empiricism and behaviorism, which view disciplinary knowledge as a collection of observations about the physical world (facts) along with procedures (step-by-step guides) to help solve problems. Instructionist systems are characterized by a unidirectional flow of information from teacher to passively receptive students and are principally designed to foster the memorization and recall of compartmentalized facts and procedures. While instructionist techniques can allow information to be conveyed, they often fail to promote conceptual understanding, which is necessary for the development of independently usable knowledge, such as knowledge that is applicable in novel situations.

Our previous training in traditional educational systems effectively prepared us for the types of activities that characterized medical work throughout the last century, most notably, the recognition of specific clinical entities, the application of relevant facts to clinical decision making, and the efficient use of procedures to effect therapeutic outcomes. Most physicians are not, however, skilled at the types of activities that will be required to ensure our collective effectiveness and relevance in an increasingly technologically complex and economically competitive society. Those activities include: 1) developing
Our continued ability to effect meaningful change in our environments and retain autonomy as clinical decision makers will depend on our success in navigating and mastering the tools of a knowledge-based economy.

a refined awareness of one’s current knowledge, 2) being willing to actively take steps to supplement knowledge when it is deficient, 3) reflecting systematically on patterns of practice and identifying areas for improvement, 4) evaluating evidence critically, coupled with a deep understanding of scientific information relevant to practice, 5) becoming fully integrated into medical systems in order to improve the safety and effectiveness of care, and 6) actively participating in the development and dissemination of new knowledge.

Physicians in general and neurosurgeons in particular must recognize that we now live in an information economy, an economy characterized by “knowledge work.” Specifically, we need to recognize that most of the knowledge relevant to physicians and the patients we serve is local, situation-specific, related to actual practice, and often collaboratively generated. In that regard, we have to be willing to promote change within the complex systems in which we work. In addition, and most importantly, we need to more effectively gather information from our day-to-day work environments and convert it into useful, context-specific, and shareable knowledge. In other words, we must commit to continually furthering our individual and collective knowledge.

Our ability to remain effective instruments for change within medical systems will require embracing new paradigms of clinical practice and continuing medical learning. Specifically, we must treat learning and practice as a continuum in which individual and collective understanding and knowledge generation become part of our daily routines. The Congress of Neurological Surgeons, through programs such as Integrated Medical Learning (IML) (CNSQ December, 2006), has already begun the process of developing more scientific and, we hope, relevant learning programs in an effort to create a culture of continuous learning and self assessment among practicing neurosurgeons.

In this issue, we focus on the importance of outcomes data collection in advancing our understanding of the value and effectiveness of various therapeutic interventions. Outcomes projects present a tremendous opportunity for our specialty to collectively advance our understanding of the effectiveness of current practice and develop hypotheses for new avenues of investigation. Ultimately, the CNS leadership envisions clinical data collection efforts merging with individual and community learning programs into a coordinated spectrum of information management and clinical systems development.

In summary, medicine and society are undergoing fundamental changes. As physicians, we are being challenged by a variety of forces to optimize the quality and safety of care while accounting for the value of the medical enterprise. Daily we are faced with a complex set of fundamental choices, many of which relate to our ability and willingness to work with these massive amounts of complex data. Our continued ability to effect meaningful change in our environments and retain autonomy as clinical decision makers will depend on our success in navigating and mastering the tools of a knowledge-based society.

As individuals, we no longer can afford to work in relative intellectual isolation. We need to expand our professional lives in ways that allow for the creation of new knowledge and the continual furthering of our own understanding. As a specialty, we must recognize the value of shared experience and work to harness our collective intellectual potential. Through collaborative efforts such as outcomes programs, we can foster the kind of interchange necessary to stimulate innovation, build consensus, and harness the knowledge and experience of our fellow neurosurgeons for the good of our patients and society.
CNS MICHAEL L.J. APUZZO LECTURER ON CREATIVITY AND INNOVATION

MUHAMMAD YUNUS, PhD
Nobel Peace Prize Laureate and Economist

Professor Muhammad Yunus was awarded the 2006 Nobel Peace Prize for his efforts to use microfinance as an instrument in the struggle against poverty. Through his groundbreaking Grameen Bank and its small micro-credit loans, Professor Yunus has managed to translate his vision of lifting individuals out of poverty into practical action, benefiting the lives of millions of people worldwide. His inspiring life story is a testimony to the power of ideas.

Professor Yunus studied economics in the Vanderbilt University, USA and received his Ph.D. in Economics in 1970. He taught economics in the Middle Tennessee University from 1969 to 1972. Returning to Bangladesh in 1972, he joined the University of Chittagong as Head of the Economics Department. He started the Grameen Bank Project in 1976. It was transformed into a formal bank in 1983, and today it operates 1,781 branches providing credit to 5.6 million poor people residing in 60,815 villages in Bangladesh.

Grameen, Professor Yunus claims, is a message of hope, a programme for putting homelessness and destitution in a museum so that one day our children will visit it and ask how we could have allowed such a terrible thing to go on for so long. The World Bank recently acknowledged that Yunus' business approach to the alleviation of poverty has allowed millions of individuals to work their way out of poverty with dignity.

In September of 2007, Professor Yunus and the Grameen Bank partnered with Intel to launch the Intel World Ahead Program in Bangladesh. This program aims to provide greater accessibility to the world's underserved by creating access to fully capable, affordable PCs tailored to their needs, as well as developing local infrastructures to sustain this access and providing the educational support needed to enable people in this area to effectively utilize this technology. This project is the next logical step toward Yunus' ultimate vision—the total eradication of poverty from the world.

Register Online by August 20, 2008!
www.cns.org
For over twenty years, organized medicine has been focusing upon the effectiveness of care in a formalized way. This involves carefully designed and implemented clinical studies. The evaluation of success or failure of the interventions involves establishing clinically meaningful methods of measurement. Yet the notion of “outcomes studies” as something apart from the recognized paradigms of clinical research is at odds with scientific reality. Whether the studies fall under the category of randomized controlled trials, utilization research, data analysis, decision analysis, meta-analysis, or longitudinal research using prospective databases, the ultimate determination of outcomes is dependent upon some measure of patient condition that is considered clinically significant.

When evaluating treatments, it is important to understand the difference between “outcomes,” “outcomes measurement,” and “endpoints.” Outcomes are those features of a patient’s condition that are the focus of the therapeutic maneuver. Examples include achievement of fusion in spinal surgery, gross total resection of a brain tumor, obliteration of a cerebral aneurysm, resolution of hydrocephalus in shunt surgery, elimination or reduction of tremors in Parkinsonism, and functional recovery in neurotrauma.

Outcomes measurement is the mechanism whereby those goals of surgery are evaluated. Examples include flexion/extension radiographs, cerebral angiography, Unified Parkinson’s Disease Rating Scale (UPDRS), and Glasgow Outcome Score. And endpoints are those points in the patient’s history of care at which the outcomes of interest are (or are not) achieved, as measured by the appropriate criteria.

It is impossible to discuss outcomes and the ways in which they are measured without noting that there are different points of view about what constitutes a clinically meaningful outcome. As clinicians, we naturally favor measures that assess some facet of our direct intervention. For example, if we are treating a brain tumor, we will use outcomes based upon imaging studies that indicate the extent of remaining tumor following surgery. But to the patient, not only is length of survival important, but quality of life must also be taken into account. To the payer, including government agencies, the comparative cost of different therapeutic maneuvers is important. Neurosurgical research does not typically consider the latter two points of view.

“Outcomes studies” ask relevant questions about prognosis of disease in patients who have or have not been treated, or compare treatment strategies. Prognosis, or natural history, studies establish the baseline for expected patient outcomes when a given disease is present or possible and no intervention is undertaken. The possibility of disease is seen in circumstances in which some intervention is undertaken, and the intervention puts the patient at risk for a different disease or illness. Examples include surgical infection or deep vein thrombosis with prolonged surgery. Then the outcomes studies of interest would be directed toward preventive strategies during the surgery, such as use of prophylactic antibiotics, methods of skin preparation, mechanisms of hair removal or use of low-dose anticoagulants or sequential compression stockings. The outcomes of interest would include presence or absence of infection, in which the outcomes measure would be some microbiological assessment, or presence or absence of DVT, where the outcomes might be measured by the gold standard of vascular imaging or some surrogate such as impedance plethysmography.
Whether the studies fall under the category of randomized controlled trials, utilization research, data analysis, decision analysis, meta-analysis, or longitudinal research using prospective databases, the ultimate determination of outcomes is dependent upon some measure of patient condition that is considered clinically significant.

Quality in clinical research is defined by a combination of excellent (experimental) design and valid and reliable measurement of outcomes. While attention to study design and acknowledgement of a hierarchy of study types (i.e., Class I-III evidence) have evolved in clinical neurosurgery, very little attention has been focused on developing a cadre of scientifically valid, reproducible, and complete outcomes measures.

The potential liability for our profession is exemplified by the difficulties imposed by the pedicle screw litigation of the past. In this scenario, the crux of the case against organized neurosurgery, individual neurosurgeons, and device manufacturers revolved around poor studies with limited and (with respect to the patient perspective) irrelevant outcomes. The primary assertion from the plaintiff’s side regarding outcomes was that the studies, and those who performed them, neglected to focus on levels of pain and suffering when defining “treatment success.” This neglect posed a very real, and potentially devastating, threat to a procedure that has helped thousands of patients and pushed the frontiers of spinal technology and surgery forward. Therefore, it is essential to focus on developing and evaluating various outcomes measures to substantiate claims of effectiveness.

It can be useful to organize thinking about outcomes measurement by categorizing outcomes. These outcomes are arranged from simplest to most complicated, from the point of view of the institutions, clinicians, patients, and payers, respectively.

Auditable Outcomes
These measures are easily obtained from a medical record review and reflect factual material that is not open to questions of interpretation or validity. These outcomes measure occurrences in clinical practice that are usually unintended, with a possible negative connotation including death, unplanned return to the operating room, length of stay in the hospital, etc. They are of great interest to hospital administrators and payers but are of secondary or tertiary importance in clinical research, unless the study is one aimed at establishing a benchmark, decreasing utilization or providing general qualitative measures of care delivery.

Clinical Outcomes
Measurement by clinicians for clinicians usually is found in this category. These outcomes represent the goals of our surgical interventions, and help to define “meaningful” from the clinician’s point of view. Examples include:
- Death, morbidity (deficit, hemorrhage, infection, CSF leak, DVT, etc.)
- Neurological status (ASIA scores, Glasgow Coma Score, NIH Stroke Scale, etc.)
- Imaging (correction of deformity, stability, tumor recurrence, aneurysm obliteration, etc.)

These measures bring into discussion aspects of validity and reliability. Validity is a measure of how well the measurement reflects the true state of affairs, and reliability represents the ability of various people (or the same person) to assess the patient in the same way over repeated examinations.

Functional Outcomes
How well patients feel is determined by how much they can do and how closely their lives approximate their perception of normal. The primary areas of interest include:
- **Physical ability** can be measured in a variety of ways, from simple measures like muscle testing (MRC scale) to tests aimed at specific abilities, like the Functional Capacity Evaluation.
- **Work ability** is difficult to assess and is far more than the physical ability to carry out tasks, and has to take into account not only the stamina involved in consistent activity, but also the psychosocial aspects of employment. Because of this, most work-related outcomes are simple and binary, such as return to gainful employment or not.
- **Pain perception and Disability** are among the most important outcomes in spinal neurosurgery and peripheral nerve surgery. There are many measures of pain, including the ubiquitous Visual Analog Scale, and more sophisticated and well-validated measures like the McGill Pain Questionnaire. Pain and disability are linked in the Oswestry Disability Index, along with self-care.
- **Self-care ability** is the most important index of patient function and the independence of the patient is a key ingredient. One of the best established and researched outcomes measures of this sort is the Functional Independence Measure (FIM). As mentioned above, the Oswestry Disability Index also includes a section on self-care, along with pain and disability, recognizing the interaction and interdependence of these aspects of patient function. In brain tumor studies, the Karnofsky score is often used as an outcomes measure in addition to survival. The Barthel Index is another measure of function and self-care and has been used in cerebrovascular studies.

Quality of Life
Measurement of quality of life is the most complicated of outcomes of interest. Whatever particular measure is chosen, it must reflect all the aspects of a patient’s life. This includes physical, psychological, and social health domains. Within these categories, “physical” encompasses disability, functioning, and symptoms; “psychological” involves behavior and positive and negative affect; and “social” includes personal relations, roles, and work. Because these are so interrelated, they must all be measured.

Many health-related quality of life measures have been developed. These include:
- **SF-36, SF-12, SF-8** – These measures of well-being cover eight domains of function and can be used in any disease entity. Because they are general measures, they do not cover areas of specific dysfunction related to particular disease entities. For this reason, they
are often used in conjunction with other measures to round out the total picture of patient outcomes.

- **Quality of Well-Being Scale** – This measure inquires about the patient’s recent symptoms and functioning within three domains: mobility, physical activity, and social activity. It is usually used in combination with other disease-specific measures.
- **Health Utilities Index** – This measure is a general measure of quality of life that has been used extensively to establish population norms as a benchmark for patients with various illnesses. It includes measures in vision, hearing, speech, ambulation, dexterity, emotion, cognition, and pain. This measure is the least useful in disease-specific research situations, except perhaps in brain cancer research, where these functions are easily affected.

**Economic Measures**

The outcomes that are least frequently measured, but increasingly more important, are the complicated aspects of healthcare costs. The types of costs and benefits that go into the equation include direct medical and non-medical costs, indirect morbidity and mortality, and intangible costs.

In summary, there are a plethora of measures and measurement techniques that can aid in defining appropriate and desirable health outcomes and measuring treatment success. Each study must be designed with the outcomes clearly in mind, because sample size calculation is based upon frequency of outcomes in treated and non-treated patients. In order to provide satisfactory information about the efficacy and effectiveness of surgical procedures, outcomes must be chosen that are meaningful to patients, clinicians, and payers (including governments and society at large.) Only then can these measures help us answer the question, “Are we doing the best thing?” as it relates to clinical situations in our own practices.
Outcomes Analysis is used to investigate the effectiveness of surgical procedures, to track and improve healthcare quality, and to give patients meaningful data for making healthcare decisions. In addition, payers and government are increasingly likely to require outcomes monitoring for contracting and policy decisions. In this article I will summarize the methodology of outcomes studies and discuss why expert clinicians need to be involved to assure scientific rigor.

Types of Outcomes Studies
Under the umbrella term “outcomes research,” numerous methods are being used to assess the effectiveness of medical care.

Utilization Research
This area of outcomes research originated with John Wennberg and his colleagues at Dartmouth who studied the variability of utilization of healthcare resources across geographic regions. Their team reported large variations in the rate of surgical procedures and hospital utilization in small geographic areas, and observed that this variability did not correlate with mortality or other discernible measures of the health of the population. Such observations raised many questions regarding the rationale behind medical decisions and utilization of resources.

Data Synthesis Studies
Frequently, outcomes research involves synthesizing data from multiple studies. Even the most elegant and rigorous randomized controlled trial will leave many questions unanswered. A single study may not have sufficient numbers of patients to evaluate certain patient subgroups, and such studies often do not take into account important factors such as quality of life and costs. Various data synthesis methodologies can be used to address some of these shortcomings.

Meta-Analysis
Problems with an individual clinical trial may be addressed by combining the results of multiple trials. Meta-analysis is a statistical method for systematically evaluating and combining these results in order to obtain a sample size sufficient for performing subgroup analysis and determining clinically significant treatment effects. This methodology can also be used to explore why individual clinical trials may yield contradictory results. However, meta-analysis still relies on the available data and if patient-oriented outcomes have not been collected, this biostatistical tool will not be able to adequately assess the effectiveness of care.

Decision Analysis
Decision analysis methodology can also be used for evaluating the effectiveness of treatment alternatives using data generated in clinical trials. First, competing therapeutic strategies are specified and clinically relevant outcomes for each strategy are analyzed via a decision tree model. Then the probabilities of reaching the various clinical outcomes are determined using data generated from relevant clinical studies.

Outcomes are assigned a utility, usually expressed in Quality Adjusted Life Years (QALYs), based on patient or potential patient perceptions. The anticipated value of the various treatment strategies can then be calculated by multiplying the probability of reaching a clinical outcome by its utility. The stability of the results can be further analyzed using sensitivity analysis, varying the probabilities of reported outcomes according to the range established from clinical studies.

Prospective Outcomes Studies
Data synthesis methods differ considerably from traditional clinical research. This is not the case with prospective outcomes analysis. The goals of determining the effects of treatment and identifying prognostic indicators are the same as the goals of traditional research, and in addition, outcomes studies often employ observational study designs such as case series, case control studies or cohort studies, or experimental study designs such as randomized trials. There are, however, real differences in regard to which outcomes are assessed for individual patients and how outcomes rates are analyzed for surgeons and for society.

Patient Level Analysis
Outcomes studies focus on treatment outcomes that are patient-oriented. Many dif-
different kinds of patient data may need to be analyzed to give a complete picture of the effectiveness of medical care. Surgeons are most familiar with clinical outcomes such as morbidity and mortality, as these are the measures employed in traditional clinical research. In addition, it is often necessary to have a reliable measure of functional health status as an endpoint in outcomes analysis. Without knowing whether or not our interventions improve or preserve the ability of our patients to lead their lives as they wish, it is difficult to see how a “good outcome” can be determined. Validated generic and disease-specific instruments are available to assess functional health status following treatment.

Patient satisfaction outcomes measure the extent to which patients’ expectations of medical care are met. Because numerous environmental factors may affect patient satisfaction, this measure will often correlate poorly with clinical and functional outcomes so these data used in isolation are of limited value.

**Surgeon or Hospital Level Analysis and Risk Stratification**

Using patient-oriented outcomes and traditional clinical outcomes, one can explore outcomes rates among various surgeons or hospitals and analyze reasons for variation in performance. In order for this kind of analysis to be valid, it is necessary to use risk adjustment methods to account for possible differences in patient risk factors from one practice to another. The process of risk stratification is used to separate the effects and quality of treatment from other factors that may have a profound influence on patient outcomes. Numerous factors need to be assessed and expert clinician involvement is essential for accurate risk stratification.

**Population Level Analysis**

Many outcomes studies look at the effectiveness of care for a large population of patients. Such large database analyses have particular strengths. Because all patients undergoing a given procedure are included, there is little selection bias in this type of study. Large database studies also have real weaknesses. The accuracy of administrative databases for reporting outcomes can be questioned, and they often lack clinical detail needed for risk stratification. This may be why large database studies often reach different conclusions regarding the effectiveness of care than would be anticipated from the results of randomized trials.

**The Importance of Outcomes Studies**

The use of outcomes analysis to evaluate the effectiveness of medical care is of obvious importance to us, our patients, referring physicians, third party payers, and to society. Nevertheless, many clinicians have questioned the necessity of conducting outcomes research. It is vitally important that these clinicians, who know the most about their patients and their disease processes, take the lead in designing and carrying out reliable outcomes studies.

The use of outcomes analysis to evaluate the effectiveness of medical care is of obvious importance to us, our patients, referring physicians, third party payers, and to society. Nevertheless, many clinicians have questioned the necessity of conducting outcomes research. It is vitally important that these clinicians, who know the most about their patients and their disease processes, take the lead in designing and carrying out reliable outcomes studies.
are asymmetrical. After assignment to surgery there is a short period of time preoperatively during which the patient may elect for other medical treatment. Patients have a longer time to change from medical to surgical treatment. Statistical methods exist to deal with crossovers, but these methods ameliorate rather than eliminate the problem.

An RCT does not eliminate bias if endpoints are ambiguous and neither the patient nor the evaluator is blinded. Patients may experience a substantial placebo effect with surgery and investigators may harbor a surgical or non-surgical bias. Having someone other than the operating surgeon evaluate patients postoperatively does not eliminate bias, it only substitutes the evaluator’s bias for the surgeon’s bias. All observers bring their biases to the evaluation.

Surgical RCTs also suffer from problems with surgeon selection. In a study comparing aspirin to placebo it really doesn’t matter who writes the order to administer the agent. This is not the case with surgical trials, where the skill of the surgeon has profound effects on outcome. A study showing a benefit from surgery with a highly skilled group will not apply to patients whose surgeons fail to match these outcomes. Similarly, a study showing no surgical benefit may not be applicable if the study surgeons have outcomes significantly worse than one with superior skill.

A final issue with surgical RCTs is their cost in time, effort, and money. In order to have enough patients to properly power a study, large multi-center trials often are necessary. These are expensive, time consuming, and labor intensive, making it difficult or impossible to repeat a trial, even if there are serious concerns about the validity of the study. Furthermore, because RCTs often take many years to complete, their results may be meaningless if new technology has developed during the trial that could affect patient outcomes.

An Alternative for Neurosurgical Quality Improvement

Prospective tracking of surgeon- and procedure-specific process and outcomes data, with feedback to individual surgeons in a non-punitive environment, is the system most likely to improve the quality of neurosurgical care. The Congress of Neurological Surgeons, working in collaboration with the American Association of Neurological Surgeons and the American Board of Neurological Surgery, is establishing such a system. The system will be used not only for quality improvement efforts but also for maintenance of certification and pay for performance initiatives.

Summary

We are living in the era of evidence-based medicine. Numerous agencies are involved in measuring the “quality” of neurosurgical care and neurosurgeons are increasingly required to define and document the value of their interventions. Often, those evaluating the quality of neurosurgical care have little insight into the issues that arise in the care of our patients and, unfortunately, data from these studies will be used to determine government policy, patient referral, and reimbursement.

However, times of great threat are often times of great opportunity. If the neurosurgical community becomes enthusiastically involved in meticulous outcomes assessment, we will be able to document the value of our interventions, better inform our patients, and continue to improve the quality of neurosurgical care. CNSQ
W hile outcomes studies have driven decision making in cerebrovascular neurosurgery for several decades, newer trends have been directed towards developing more detailed patient-centered as well as disease-specific measurements. In this article we briefly address these outcomes studies by describing: 1) the purpose of outcomes studies 2) outcomes studies relevant to cerebrovascular neurosurgery and 3) the newer outcomes instruments for next generation studies in cerebrovascular disease.

As clinicians are exposed to an increasingly complex array of epidemiological instruments, choosing which outcomes measure to use becomes ever more important to understand the likely end result of a specific therapy or intervention. Outcomes measurements can be classified into several categories, including mortality, physiologic measures, clinical events, symptoms, functional measures, self-reported overall health, and patient experiences with care delivery. Ultimately the appropriate choice of instruments allows clinicians to better answer the single most important question any patient may ask regarding the proposed therapy: “How will I do?”

The correct tool permits patients, their families, and the public to better understand the effect of an intervention in terms that are familiar and relevant. This concept of “patient centered outcomes” moves away from physiologic measures and towards more practical measures like functional status and quality of life (QOL). And because functional status and QOL are so different, no single epidemiological instrument can simultaneously evaluate both.

Investigators also benefit from using a combination of general and specific measures. Broader measures allow comparisons of outcomes across different disease entities, (e.g., spinal surgery versus cranial surgery) and can be complemented further by disease-specific measures (e.g., NIH Stroke Scale for stroke patients). This combination is particularly powerful in describing outcomes.

Once the decision is made regarding what type of metric to use, there are several qualities to consider in choosing a particular measurement tool. The instrument’s feasibility is a major factor, and respondent burden and clinical appropriateness are clearly important.

As clinicians are exposed to an increasingly complex array of epidemiological instruments, choosing which outcomes measure to use becomes ever more important to understand the likely end result of a specific therapy or intervention.

As well. In addition, one must always ask: “Is this outcomes tool appropriate for the disease under study?”

Also, the quality of the outcomes measure is dictated by the reliability, validity, and responsiveness to change of the instrument in relation to the specific population and intervention studied. In other words, because any given cerebrovascular disease will only affect some dimensions of a patient’s health, only certain outcomes measures are relevant for any given intervention. For example, studying cognitive impairment after a spinal cavernous malformation resection is not useful for understanding the outcomes after surgery. Further information comes from using these measures in studies that have scientifically rigorous design along with the study of appropriate co-factors that allow highly specific patient outcomes prediction. A few types of these measures are discussed below.

Early and late mortality is relevant to almost every cerebrovascular disease treated, but even this measure is subject to loss to follow-up, insufficient follow-up duration, and errors in death assignment. Morbidity is much more complex to measure. The International Classification of Functioning and Disability defines morbidity through three categories: 1) impairments (signs or symptoms of illness), 2) restrictions in activities (observed behaviors, functions), and 3) restrictions in participation (social roles, participation). Advanced outcomes studies include all of these aspects of morbidity assessment for an overall view of treatment outcomes.

A brief survey of cerebrovascular outcomes instruments includes scales for stroke impairment such as the NIH stroke scale, Hemispheric Stroke Scale, and Canadian Neurological Scale, which are widely accepted for their simplicity and content validity with stroke patients. The Beck Questionnaire and the Zung Depression Scale have been used to study the prevalence of mood disorders in stroke patients. Other commonly used metrics include the MMSE, Wechsler Adult Intelligence Scale, Wechsler...
Memory Scale, and the Glasgow Coma Scale (GCS) for measuring cognitive impairment and level of consciousness.

When moving from impairment to activities scales, one should note that the scientific properties of the latter vary widely. The Glasgow Outcome Scale (GOS) and the Rankin Scale are commonly used for rapid and gross categorization in large studies, but have been criticized for low sensitivity for between-group differences as they offer very broad classifications of patients. When used alone, these scales offer little more than a general understanding of outcomes.

The Barthel Index, the most widely used activities of daily living (ADL) scale, describes basic measures of self-care and physical activity. However, it is best suited for patients with physical impairment. For example, for mild-to-moderate stroke or good grade SAH patients, the Barthel Index is prone to ceiling effects because it fails to tease out differences in cognitive function in patients who perform well physically. Instrumental ADL scales go one step further by measuring more complex cognitive and fine motor tasks that better predict how well a patient will live independently in the community. A variety of scales, including the Functional Activities Questionnaire and the Functional Independence Measure, have been used in the post-stroke population.

Participation and QOL measurements have only recently been developed for stroke patients. The Stroke Impact Scale and the Stroke-Specific Quality of Life Scales, designed for populations with mild to moderate stroke, remain untested in populations suffering severe stroke and/or in subarachnoid or intracerebral hemorrhage. More general outcomes instruments such as the SF-36 and Sickness Impact Profile have been applied broadly in stroke populations.

Disease-specific applications of these tools in cerebrovascular neurosurgery, both in established and future studies, demonstrate the benefit of particular instrument choices. In carotid atherosclerotic disease, the North American Symptomatic Carotid Endarterectomy Trial (NASCET), European Carotid Surgery Trial (ECST), and Veterans Administration Cooperative Study (VACS) used the primary endpoints of stroke and death to demonstrate the benefit of carotid endarterectomy (CEA) in selected symptomatic patients. Importantly, post-hoc analysis in these studies provided two important additional contributions. Co-factor analysis improved pre-operative risk stratification and activity measures demonstrated improved functional status in surgically treated patients with symptomatic stenosis. Co-factors proved to be equally important for asymptomatic carotid artery stenosis where the more marginal benefit of CEA is balanced against co-factors associated with higher surgical risk.

Moving to endovascular therapy for carotid disease, mortality and stroke morbidity outcomes were used to compare CEA to carotid angioplasty and stent placement in the Stent-Supported Percutaneous Angioplasty of the Carotid Artery versus Endarterectomy (SPACE) and Endarterectomy versus Angio-
plasty in patients with Severe Symptomatic Carotid Stenosis (EVA-3S) trials. One anticipates that detailed co-factor analysis will follow to define subgroups of patients favoring each therapeutic arm.

The impact of outcomes studies can occasionally surprise investigators. Surgical patients failed to show a reduced stroke rate and demonstrated a lower functional status in the EC-IC bypass study, although criticisms regarding the patient selection and exclusions have been widely published. Renewed interest in intracranial carotid disease is now reflected in the Carotid Occlusion Surgery Study, where both surrogates of outcomes and functional studies will be important.

Intracerebral hemorrhage (ICH) outcomes studies have evaluated mortality against several co-factors and correlated impairment at presentation with 30-day mortality rate for both medial and lateral hemorrhages. The high mortality rate for ICH is reflected in the use of gross measures for activity, like the GOS, but with advances in medical management, secondary factors like cognitive performance and QOL outcomes may be included in future studies.

Lesions prone to hemorrhage within the brain parenchyma utilize different tools. Caverno malformation outcomes studies draw upon a broader set of tools, including the Rankin scale, author-defined scales of neurological impairment and seizure outcomes, and KPS to evaluate neurological impairment and seizure control in addition to evaluating surrogate measures like rehemorrhage rates. Current artereovenous malformation (AVM) outcomes also include a broad set of instruments, where morbidity arises from hemorrhage, non-hemorrhagic progressive deficit, seizure activity, and headaches. Surgery has not been directly compared to natural history to assess overall outcomes, although GOS, Rankin disability scores and co-factor analysis help predict outcomes with surgical treatment.

Outcomes data for cerebral aneurysms therapy must take into account the differences between subarachnoid hemorrhage and the unruptured aneurysm population. Incidence and case-fatality rates developed from large-scale population-based studies confirm the association between higher case-fatality rates and older age. Associations between both initial level of consciousness and mortality rates and early intervention and lower mortality have also been reported. In addition, lower mortality rates have been demonstrated with the nonaneurysmal perimesencephalic variant SAH. The International Subarachnoid Aneurism Trial (ISAT) used a large randomized design with a modified Rankin Scale to compare mortality and dependency in surgical and endovascular treated groups.

A recent focus towards non-physical impairment in best-grade SAH patients reflects sensitivity towards more subtle but significant declines in functional status and return-to-work rates. Both the Beck Depression Inventory and the Zung Self-Rating Depression Scale are used with post-SAH patients. Some authors suggest a correlation between Reintegration to Normal Living scores, QOL, and depressive symptoms. Unlike the stroke literature, there is little information relating cognitive outcomes after SAH to general QOL measurements, but several studies have used instruments to measure cognitive outcomes. Neuropsychological outcomes between endovascular and surgical treatment of aneurysmal SAH showed no differences at one year using the Wechsler Adult Intelligence Scale, the Wechsler Memory Scale, the MMSE, and the Cambridge Neuropsychological Test Automated Battery. Activity measurement using the Barthel Index is subject to ceiling effects in Hunt-Hess I-III patients. QOL measurement changes are significant with the SF-36, particularly with reductions in role limitations due to physical and emotional problems.

The unruptured aneurysm population will benefit from measures that are sensitive to changes in what is a minimally impaired pre-treatment population. The Barthel Index can measure changes from independent to dependent functional status. The cost/benefit of treatment regarding protection of aneurysm rupture at the cost of partial cognitive degradation will push careful co-factor analysis to determine high-risk patients. In future analyses, detailed neuropsychological and QOL assessments should be included for both treated and untreated groups from ISUIA. The continuing application of disease-specific outcomes metrics will give neurosurgeons the best data for choosing the optimal therapy for their patients.

As new interventions emerge in cerebrovascular neurosurgery, outcomes studies will help define patient benefits. Improving the choice of outcomes measures in clinical studies design will also help us answer the question: “How am I going to do?” Selecting the best disease-specific and multidimensional outcomes measures for the studied intervention, applying the instruments within a methodologically rigorous trial design and studying co-factors within these changes will allow us to provide truly valuable evidence-based recommendations for our patients.
Outcomes research in spinal surgery is currently a hot topic. Lumbar fusion procedures have been particularly scrutinized, primarily due to data indicating increased numbers of fusion surgeries and also a wide variation in fusion rates across the country (N Engl J Med 350:722-726; Spine 31:2707-2714). While the importance and/or meaning of such data is unclear, critics of lumbar fusion have stated that the motivation for many fusion procedures is monetary and that industry is driving the process through incentives and kickbacks to surgeons (New York Times, December 31, 2003; N Engl J Med 350:722-726). The public outcry has been such that in November 2006, the Center for Medicare Services (CMS) held a medical care advisory committee (MCAC) meeting specifically devoted to the use of lumbar fusion in the Medicare population, the first time such a process has been applied to existing technology. Such hearings are usually reserved for the evaluation of new technologies, and recommendations are used to determine whether or not CMS will pay for a new device. The fact that this process was being applied to the issue of lumbar fusion was, to say the least, troubling.

Several of us, representing organized neurosurgery (AANS and CNS) through the spine section, had the opportunity to participate in the MCAC process. We partnered with representatives of the North American Spine Society, the Scoliosis Research Society, and the American Association of Orthopaedic Surgeons to form the “Combined Societies Task Force on Lumbar Fusion.” Charles L. Branch, Jr., David Polly, Steven Glassman, David Wong, and I were among the participants who presented the surgical side of the argument. During the proceedings, it became abundantly clear that there was no high quality evidence to support or refute the use of lumbar fusion in the Medicare population. Furthermore, what evidence existed regarding the non-Medicare population was of poor quality and had questionable relevance to the North American patient population. The surgical societies were charged with the task of developing better evidence to support lumbar fusion, with the implied consequence of inaction potentially including non-payment.

In response to this charge, the societies collaborated to form the “Lumbar Fusion Task Force.” This task force was created to address the concerns that were raised by the MCAC panel and to develop a mechanism whereby appropriate research could be fostered to either support or refute the efficacy of fusion as a treatment strategy for Medicare patients. Each of the participating societies agreed to contribute up to $10,000 to support the activities of the task force. Steven Glassman and I were put in charge and we quickly enlisted the help of Michael Kaiser, Zoher Ghogawala, and Paul McCormick from the ranks of the neurosurgical spine surgeons. Our initial plan was to organize spinal surgeons such that outcomes research could be catalogued, refined, and developed into properly performed ran-
domized trials that would produce information of high enough quality to satisfy the MCAC panel.

At the same time that the task force was being organized, the first of the publications describing the SPORT studies was released (JAMA 296:2451-2459). SPORT is a multi-million-dollar National Institute of Health (NIH)-sponsored randomized controlled multi-center study designed to determine the efficacy (more accurately described as being designed to refute the efficacy) of surgery for lumbar disc herniation. The results of the study are most interesting. Although patients who had surgery did much better than those who did not in every outcomes measures assessed and at every time point assessed, the conclusion drawn by the authors is that the difference was not significant. It is clear from reading the paper that the reason for this discrepancy related to a very high crossover rate. This crossover essentially eliminated the power of the randomization. The study is actually more accurately described as a large cohort study – a situation realized by the authors by the time that the second SPORT study was published in March, 2007 (N Engl J Med 356:2257-2270). As a cohort study, both SPORT papers are strongly supportive of surgery for properly selected patients.

The failure of the SPORT study to provide the type of evidence required by MCAC and other review panels was sobering to the members of our task force. How are we to develop randomized trials that have a meaningful chance of success, given the failure of a much higher profile and more expensive effort? Furthermore, we realized the reasons for the failure of randomization in SPORT (ethics and IRB requirements) are unavoidable in North America. The task force has entered into a conversation with the CMS, the NIH, and the Agency for Healthcare Research and Quality to attempt to determine how to answer basic questions regarding the utility of spinal surgery (given the restraints of a free society, patient preference, and limited resources). It is anticipated that outcomes research will play a large role in this process, and it is hoped that this research may be used for other purposes as well, such as maintenance of certification, credentialing, and quality of care reporting.

We have many challenges facing us when contemplating outcomes research in spinal surgery. Fundamental issues include defining patient types and indications, refining and validating outcomes measures, and negotiating the parameters of participation with the policy-making bodies and surgeons. For example, an ongoing project through the AANS/CNS Washington Committee is a pilot on lumbar stenosis. Robert Harbaugh was the motivating force behind the project and I and other spine section members (Robert Heary, David McKalip, and others) participated in the design and implementation. As the project has proceeded, numerous shortcomings have been noted. For example, only patients who undergo surgery are included in the database, eliminating any way to establish comparative efficacy between surgery and non-surgical treatments. The Oswestry Disability Index (Spine 25:2940-2952) is one of the outcomes measures used. This is a tool that is assumed to be valid for spinal surgery in general. However, its questions regarding weight lifting, sex life or dancing may not be all that helpful when assessing the results of lumbar stenosis surgery in an octogenarian widow with bad hips, COPD, and pelvic floor weakness. This limits the sensitivity and validity of the instrument in the elderly population.

Finally, as a participant in the development and implementation of the study and as an academic spinal surgeon, I am as highly motivated to enter information into the database as anyone could be. Yet I went to my nurse’s office 30 minutes ago and she showed me a pile of completed surveys that is well over an inch high and goes back several months. She simply has not had time to enter the data (nor have I) and no one is being paid extra to do so. It is imperative that some sort of “carrot” be associated with data entry into outcomes data sets, whether it be financial (data accepted as a substitute for process measures in P4P) or professional (required for MOC or membership in societies).

In summary, outcomes research is now part of the spinal surgery landscape. The prevalence of such research will likely greatly expand in the next decade. Participation in some sort of outcomes data-gathering project may eventually be required of every physician in the country. It is important that we spend the time, money and effort to organize our mechanisms and measures properly such that our members and society at large can benefit from the information gathered. Otherwise we will suffer the consequences of enforced participation in meaningless performance measure reporting, with simultaneously decreased reimbursement. CNSQ
Outcomes: definitions in neuro-oncology

The two main areas in neurosurgical oncology that we will focus on are perioperative outcomes and long-term outcomes in gliomas and metastatic brain tumors. Perioperative outcomes can be defined as the risks and benefits of surgery evident within thirty days of an operation. These include perioperative complications and changes in neurological function and performance status. Long-term outcomes can be defined as benefits or risks that become evident more than thirty days from surgery. This would include long-term neurological function and performance status, progression-free survival, and overall survival. As in all areas of outcome research, data supporting various conclusions can be categorized according to levels of evidence.

Malignant Gliomas

Perioperative Outcomes for Malignant Glioma Patients

Patients with malignant gliomas generally either have some form of stereotactic biopsy or undergo a more extensive craniotomy and resection of tumor. As one would expect, perioperative outcomes can vary depending on what type of surgery is performed.

Stereotactic biopsy is generally effective at obtaining a diagnosis, and complications are relatively uncommon. Diagnostic material is obtained in the vast majority of cases (over 90%) (J Neurosurg 81:165-168, 1994) but Level II evidence exists that the pathological diagnosis (mostly tumor grade) may change in up to 38% (Neuro-oncol 3:193-200, 2001). Hemorrhage is seen on imaging in over 50% after stereotactic biopsy but symptomatic bleeding is far less frequent (~5%). Permanent and/or severe new neurological deficits are uncommon (2.5 to 5%) and mortality after craniotomy and resection is similar, suggesting that it reflects natural history not specific procedural risks. The literature is largely prospective and can be considered Level I evidence of perioperative “prognosis.” Thus, it can be concluded that stereotactic biopsy is a relatively safe and reasonably accurate method to obtain a tissue diagnosis in malignant glioma patients. However, it must be remembered that there is essentially no chance for neurological improvement, an important aspect of perioperative outcome in this population.

Perioperative outcomes following craniotomy and resection of malignant gliomas have been similarly well described. The Glioma Outcomes (GO) Project has provided benchmark Level I evidence describing perioperative outcomes (J Neurosurg 98:1175-1181, 2003). Wound complications occurred in 10% of patients with newly diagnosed malignant gliomas undergoing craniotomy and resection. Similar rates (2.5 to 16%) have been reported in other, mostly Level II, studies. Systemic complications occur at about the same rate, happening in 9% of patients in the GO Project and 3.5 to 12.5% in earlier studies. Eight percent of first craniotomy patients in the GO Project suffered permanent neurological deficits perioperatively (7% – 21% in earlier studies). Thus, there is strong evidence supporting the conclusion that complications occur more frequently in patients undergoing craniotomy and resection for malignant glioma compared to stereotactic biopsy. However, there is also Level I evidence from the GO Project that patients undergoing craniotomy have a 50% chance of having improved neurological function postoperatively. This is an important distinction from stereotactic biopsy where the chance of neurological improvement is essentially zero.

Long-term Postoperative Outcomes in Malignant Glioma Patients

It is accepted that some form of surgery is necessary for malignant glioma patients to obtain a tissue diagnosis. The main outcome issue for discussion regarding malignant glioma surgery is this: does more extensive surgical resection provide a survival benefit? Many retrospective and prospective studies in the past two decades have suggested an association between extent of resec-
tion and survival, but this has not been universally accepted and has been based on Level II evidence at best (Principles of Brain Tumor Surgery, In press, Handbook of Clinical Neurology). More recently, Level I evidence has been published demonstrating that increased extent of resection is definitely correlated with prolonged progression-free survival for patients with glioblastomas. In 2006, Stummer et al. (Lancet Oncol 7:392-401, 2006) reported results of a randomized, controlled, multicenter trial comparing survival in 322 newly diagnosed glioblastoma patients undergoing standard microsurgical resection versus fluorescence-guided microsurgical resection using 5-aminolevulinic acid (5-ALA). 5-ALA was given intravenously preoperatively, resulting in intra-cellular accumulation of fluorescent porphyrins within glioma. Extent of resection in the 5-ALA arm was guided by the fluorescence seen intraoperatively in appropriate light. Patients with fluorescence-guided resection had less residual enhancing tumor on volumetric analysis of postoperative MRIs (65% gross total resection with fluorescence; 36% with standard resection). Furthermore, patients with fluorescence-guided resection had significantly longer times to progression (41% progression-free with 6 months with fluorescence; 21% with standard surgery). These findings are summarized in Table 1. While this study is perhaps the most compelling evidence that a greater degree of resection improves long-term postoperative outcome in malignant glioma patients, it has some shortcomings. In particular, the study was not sufficiently powered to answer the fundamental question of whether greater extent of resection is associated with prolonged overall survival.

**Diffuse low grade gliomas**

Diffuse low grade gliomas that we will discuss include WHO Grade 2 astrocytomas, oligodendrogliomas, and mixed oligoastrocytomas. Few studies have specifically addressed perioperative outcomes (complications, changes in neurological status) in these patients. What data we have can largely be extrapolated from the studies above for malignant gliomas, some of which included patients with lower grade tumors as well. Needless to say, this dilutes the levels of evidence upon which conclusions can be made. Despite this, it is likely that findings in these studies are a reasonable approximation of perioperative outcomes in diffuse low grade glioma patients.

**Long-term Postoperative Outcomes in Diffuse Low Grade Glioma Patients**

The question: “does increased extent of resection lead to improved survival?” and data quality issues for patients with diffuse low grade gliomas are similar to those with malignant gliomas. The central problem in answering this question (the lack of Level I evidence) is also identical. Two attempts in the past decade to develop clinical practice guidelines for patients with diffuse low grade gliomas have concluded that the only recommendation supported by the literature is that tissue diagnosis should be obtained prior to proceeding with treatment (eg. radiation or chemotherapy) (Br J Cancer 89:573-583, 2003; Neurosurg Focus 4:article 10, 1998). Otherwise, no strong evidence was available favoring any of the various standard management options (observation with serial MRI; biopsy +/- radiation and/or chemotherapy; surgical resection +/- radiation and/or chemotherapy).

Parney and Berger recently reviewed the literature from 1987 - 2007 in a similar manner to that described earlier for malignant gliomas (Principles of Brain Tumor Surgery, In press, Handbook of Clinical Neurology). Greater extent of resection was associated with prolonged survival in 7 of 9 studies involving 899 patients. This association was not demonstrated in two studies involving 1075 patients. As with malignant gliomas (and as found previously by those seeking to develop clinical practice guidelines), the studies were very heterogeneous. None met criteria for Level I evidence. A meta-analysis was not possible but, like malignant gliomas, a trend was seen toward increased overall survival with greater extent of resection. Of note, the magnitude of this increased survival was much greater for patients with diffuse low grade gliomas (years) than patients with malignant gliomas (months).

**Metastatic Brain Tumors**

When treating patients with metastatic brain tumors, the clinician must consider the etiology, extent, and prognosis of the primary cancer, the number, size, and location of the metastasis, as well as the patient’s performance status before making treatment decisions. In general, therapeutic decisions are usually guided by the following typical sce-

<table>
<thead>
<tr>
<th>Light Source</th>
<th>Gross Total Resection</th>
<th>6 Month PFS</th>
</tr>
</thead>
<tbody>
<tr>
<td>White Light</td>
<td>36%</td>
<td>21%</td>
</tr>
<tr>
<td>Fluorescence</td>
<td>65%</td>
<td>41%</td>
</tr>
</tbody>
</table>

Table 1. Lancet Oncol 7:392-401, 2006
simulations: patients with single brain metastasis, small numbers of brain metastasis (typically 2-4), and patients with multiple brain metastasis (>4). The etiology of the primary cancer must also be considered. Treatment considerations include surgery, stereotactic radiosurgery (SRS) both of which can be given with or without whole brain radiotherapy (WBRT). We will discuss outcomes for adult patients with metastatic brain tumors who present with these typical scenarios separately.

**Patients with Single Brain Metastasis**

The role of surgery for brain metastasis was somewhat controversial until the mid- to late 1990s. The seminal outcome studies establishing the efficacy of WBRT, which was considered the “standard of care” for most cases during this time, were done by the Radiation Therapy Oncology Group (RTOG). A recursive partitioning analysis (RPA) of three separate randomized controlled RTOG trials including over 1200 patients conducted between 1979 and 1993 classified patients with brain metastasis into three groups with statistically different prognoses (Int J Rad Onc Biol Phys 37:745-751, 1997) (Table 2). Notably, Class I patients, who had the best median survival of 7.1 months, were those < 65 years of age, with KPS >= 70, and a single brain lesion with well controlled primary disease. This Class I evidence serve as a basis of comparison for subsequent surgical trials. There are three randomized controlled studies comparing WBRT to surgery and WBRT. Patchell et al. (JAMA 280:1485-1489, 1990) demonstrated significantly improved median survival (40 wks vs 15 wks; p=0.01), functional independence (38 wks vs 8 wks; p=0.005), and local control (48% vs 80%; p=0.02) in 48 patients (KPS >= 70) with biopsy proven brain metastasis treated with surgery + WBRT (36 Gy) compared to those treated with biopsy + WBRT alone (JAMA 280:1485-1489, 1990) (Table 3). Perioperative and 30 day mortality was 4% and perioperative and 30 day morbidity was 8 and 17% respectively. Similar results were reported by Vecht et al. (Ann Neurol 33:583-590, 1993) in 63 patients with KPS>= 50, who also noted that this improvement in median survival (12 vs. 7 mo) applied only to patients with stable extracranial disease. These two trials have served as Class I evidence supporting the role of surgery in patients with single brain metastasis, particularly for patients meeting criterion for RPA Class I. Interestingly, a third randomized controlled study in 84 patients published later failed to confirm the benefit of surgery (Cancer 78:1470-1476, 1996) but consensus opinion had already established surgical resection for patients in RPA I as the treatment of choice. Patchell et al. (N Eng J Med 322:494-500, 1990) then established the role of postoperative WBRT after surgical resection in a randomized controlled trial of 95 patients with single lesions. WBRT increased local control (82% vs 30%; P<0.001), and brain control (86% vs 63%; p=0.01); and decreased neurologic death (14% vs. 44%; p=0.003), but did not alter survival.

There is also Level I evidence to support the use of SRS for the treatment of patients in this subgroup. As expected, perioperative and 30 day morbidity and mortality have been < 1% in these studies. Andrews et al. demonstrated and colleagues demonstrated the efficacy of SRS boost post-WBRT in a randomized controlled trial of 333 patients with 1-3 brain metastasis < 3 cm in diameter (Lancet 363:1665-1672, 2004). In particular, this study provided Level I evidence that SRS boost was effective in prolonging survival in patients with single brain metastasis (6.5 vs 4.9 mo). There is no Level I evidence comparing Surgery vs. SRS for the treatment of such patients. Level II studies in patients with oligometastatic brain lesions exists and will be discussed below. Thus, clinicians must use clinical judgment in making treatment decisions weighing the pros and cons of these approaches.

**Patients with Oligometastatic Brain Tumors (2-4)**

Several Phase III RTOG studies in the early to mid-1990s that increasing dose or fractionation scheme in patients with one or more brain metastasis provided no survival advantage compared to a conventional regimen of 30 Gy in 10 fractions. The role of surgery for multiple brain metastasis has also been somewhat controversial and there is no Level I evidence to support this approach. However, a retrospective single institutional cohort study demonstrated that patients with KPS >=70 who underwent imaging complete resection of 2-4 brain metastasis had the same median survival as patients undergoing imaging complete resection of a single lesion (14 mo);
Nevertheless, there remain significant gaps in our outcomes knowledge. Many of these questions will require randomized controlled studies to answer definitively and some of these are unlikely to be completed.

though patients who had some but not all lesions resected had median survival of 6 mo and did not appear to benefit.

There is also Level I evidence to support the use of SRS for the treatment of patients in this subgroup with low morbidity and mortality. The results of RTOG 95-08 reported by Andrews et al. (Lancet 363:1665-1672, 2004) also demonstrated that adding SRS boost to WBRT for patients with 2-3 brain metastasis improved survival in patients in RPA Class I (11.6 vs 9.6). While 30% of patients in both groups died of progressive brain metastasis, SRS boost improved local control, KPS at 6 mo, and decreased steroid use (Lancet 363:1665-1672, 2004). Conversely, a randomized controlled study of SRS +/- WBRT in 132 patient with 1-4 brain metastasis < 3cm in diameter demonstrated that addition of WBRT to SRS improved local control (53.2 vs 23.6%), without significant differences in 1 year survival, cause of death, or toxicity (JAMA 295:2483-2491, 2006).

In summary, there is Level I evidence to support the efficacy of both surgery and SRS +/- WBRT in patients with KPS > 70 with oligo metastatic brain disease and outcomes appear to be similar for both treatments at selected institutions. Nonetheless, clinicians must use clinical judgment in making treatment decisions weighing the pros and cons of these approaches. The advantages of surgery compared to SRS are the ability to confirm the diagnosis, eliminate mass effect, and in some cases improve tolerance to adjuvant therapy. It is thus preferentially used for larger (>3cm), superficial tumors, infratentorial lesions, or in combination with SRS. In contrast, SRS has minimal perioperative morbidity, but works best on smaller masses (<=3.5 cm) with minimal mass effect or edema. It can also be employed in patients who are not surgical candidates due to medical co-morbidity or advanced systemic disease and is easily applied to multiple brain metastasis. In many instances surgery and SRS are combined with or without WBRT.

**Patients with Multiple Brain metastasis (>4)**

There is no Level I evidence which defines optimal treatment of patients with more than 4 brain metastasis. WBRT remains the standard of care in most patients with a life expectancy of greater than three months on the basis of systemic disease. There is also considerable Level II evidence that SRS may be effective in up to ten brain metastasis if they are smaller than 3 cm and not associated with mass effect or edema.

**Conclusions and Future Outcomes Studies in Neuro-Oncology**

As the preceding discussion illustrates, recent studies have made significant contributions to our understanding of outcomes in neurosurgical oncology. Nevertheless, there remain significant gaps in our outcomes knowledge. Many of these questions will require randomized controlled studies to answer definitively and some of these are unlikely to be completed. However, these obstacles are not insurmountable and the potential benefits for patients are considerable. In some cases, changing the question (eg. comparing techniques that enable greater malignant glioma resection such as fluorescent-guided surgery to standard techniques instead of comparing biopsy to resection) and increasing multi-center collaboration may be a viable path to generating Level I evidence. In others, well designed Level II trials may have to suffice. In any event, it is clear that outcomes will continue to shape clinical management for patients with brain tumors.
Outcomes after traumatic brain injury (TBI) are measured by more than survival rates. They are measured by an assessment of patients’ functional, cognitive, emotional and psychosocial status throughout the rehabilitation process and during long-term care and recovery. A major confounding factor in outcomes assessment after TBI is that the patient population is not comparable to society at large and does not have a normative dataset for comparison. Fortunately, several tools/scales are validated for and employed in the evaluation of outcomes from traumatic brain injury. They include the Glasgow Outcome Scale (GOS), Glasgow Outcome Scale Extended Version (GOS-E), Disability Rating Scale (DRS) and Functional Independence Measure (FIM™).

The Glasgow Coma Scale (led by Sir Graham Teasdale and Brian Jennett) was developed to monitor and grade closed head injuries. It was further developed into an outcomes tool as the Glasgow Outcome Scale (GOS) (Lancet 1:480-484, 1975). GOS classifies patient outcomes into five broad categories: dead, vegetative state, severe disability, moderate disability, and good recovery (Table 1).

It was further validated and extended into the GOS-E which involves a structured interview and includes 8 categories of outcomes: dead, vegetative state, lower severe disability, upper severe disability, lower moderate disability, upper moderate disability, lower good recovery, and upper good recovery (Table 2). This version of the GOS was noted to be more sensitive to changes in functional status between 3- and 6-month follow-up time points. This has important implications for clinical trials in that the GOS-E is likely to be more sensitive than the GOS to interventions, particularly in measuring changes over time.

The DRS was developed as an instrument for assessing impairment, disability and handicap in individuals with brain injury. The DRS assesses arousal and awareness within motor response, eye opening, and verbal response measured by using a slight modification of the Glasgow Coma Scale. Disability is evaluated through a patient’s cognitive ability in the areas of feeding, grooming, and toileting. Issues of handicap are measured through examining the patient’s level of independent functioning and their employability (for a complete description, visit tbims.org/combi/drs/drssyl.html). The DRS was originally developed more to track patient improvement after TBI than to place patients in a major outcomes group.

The FIM™ is a functional outcomes measure that evaluates six domains: self-care, mobility, locomotion, bowel/bladder control, communication and social cognition. (FIM™ is a registered trademark of Uniform Data System for Medical Rehabilitation, www.udsmr.org.) Each item is based on a seven point scale, ranging from complete dependence in a category to complete independence. Although it has been found to be less sensitive in monitoring longer term outcomes, it is the most widely used measure of functional outcomes in clinical rehabilitation settings and has been validated in the TBI population. The FIM™ is often used to assess progress through repeated measurements during an inpatient rehabilitation stay.

Other outcomes tests are administered to TBI patients to evaluate common emo-
tional, psychological, and social sequelae after injury. Emotional outcomes, such as the onset of depression, can be measured through tools such as the Patient Health Questionnaire (PHQ-9), Beck Depression Inventory, and subscales of the Center for Epidemiologic Studies (CES-D), and Neurobehavioral Functioning Inventory (NFI-D). The Brief Symptom Inventory is an assessment measure which can screen for psychological problems in patients. Social outcomes can be assessed through the evaluation of a patient’s integration into the community, social supervision needs and basic outcomes such as return to work or driving. The Community Integration Questionnaire (CIQ) is used to assess areas of home integration, social integration and productive activities.

Quality of life is another important social outcome that can be assessed in TBI patients through the Deiner Satisfaction with Life Scale, which measures the patient’s satisfaction with work and family life after a brain injury, and the Modified Perceived Quality of Life Questionnaire, which measures the patient’s satisfaction with his or her functioning in areas of physical, cognitive, and emotional domains.

Outcomes testing in brain injury patients can be challenging due to the wide range of cognitive and physical deficits that occur across a spectrum of injury severity levels. Injury severity ranges from a patient with “mild” TBI but persistent cognitive deficits to the severe TBI patient in a persistent vegetative state. Many neuropsychological assessments are inappropriate in severe TBI due to patients’ devastating cognitive and physical limitations. Current research is being conducted with assessment tools that may be utilized for these more severely injured patients. For example, the Severe Impairment Battery (SIB) is an assessment tool primarily used in patients with dementia. The SIB is being conducted to evaluate the severely impaired brain injury patient in areas such as orientation, language, memory, attention, visuospatial ability, construction, praxis, and social interaction. These domains are similar to the areas of outcomes evaluated in less severe patients through standardized neuropsychological assessments; however, the SIB is designed such that a patient can be evaluated in all domains regardless of whether he or she is experiencing a severe limitation in one or more cognitive areas.

Outcomes testing and measurement for patients with brain injury are commonly integrated into research studies within the fields of neurosurgery and rehabilitation. As mentioned above, it is impossible to establish a pre-injury baseline of function in patients; however, preliminary functional and neuropsychological outcomes can be obtained in the acute inpatient and rehabilitation settings to serve as a baseline for ongoing evaluations at established time points post injury. In this way, outcomes can be measured in patients to assess future changes during rehabilitation procedures and therapy. The effectiveness of novel therapies, procedures, and medications can be assessed through the serial evaluation of patient’s neuropsychological and functional outcomes assessments. Outcomes assessment can also be utilized in neuro-imaging research. Imaging tools, such as PET, SPECT, and fMRI are being used in the research setting in conjunction with neuropsychological assessment tools to evaluate areas of pathophysiology and corresponding neuropsychological outcomes.

Outcomes measures and assessments are utilized in various manners for patients who have sustained a brain injury. During the rehabilitation and recovery period, outcomes are collected to help establish a comprehensive rehabilitation plan for patients in conjunction with medical treatments, and physical, occupational, and speech therapies. Outcomes assessments can be administered after the rehabilitation phase at selected time intervals to monitor patient progress in functional and cognitive outcomes; assess emotional, psychological, and social alterations that can occur after injury; and guide re-integration and return to work decisions. In the research setting, outcomes evaluation is used to assess novel therapies and medications, as well as in conjunction with neuro-imaging techniques. Measuring functional, cognitive, and psychosocial outcomes in patients who have sustained a brain injury is an essential tool for the neurosurgeon to assist in patient rehabilitation and care. CNSQ

Further Reading


WWW.CNS.ORG
Two main groups benchmark performance for medical care in children: The National Association of Children’s Hospitals and Related Institutions (NACHRI)/The National Association of Children’s Hospitals (NACH), and Child Health Corporation of America (CHCA) (a business alliance of Children’s Hospitals). Data collected by these organizations include financial information, the numbers and types of operative cases, and staffing ratios.

While these data are important for hospital administration, more relevant to the physician are the organizations’ various benchmarking projects which are implemented not only to save costs, but to improve patient care. Most of these projects are not directly related to pediatric neurosurgery, but rather focus on more prevalent conditions such as asthma or catheter infections. In the past, NACHRI has investigated the cost and length of stay for craniotomies across Children’s Hospitals. While this information may relate to quality of care, it is certainly more significant for meeting the needs of hospital administrators trying to cut costs.

In August 2002, the Centers for Medicare and Medicaid Services (CMS), the Centers for Disease Control and Prevention (CDC), and a number of medical and surgical societies implemented the Surgical Infection Prevention Project (SIPP). This initiative was created to mitigate the finding that overuse, underuse, improper timing, and misuse of antimicrobials occurs in 25-50% of operative procedures. As a result, there is now an on-going CHCA Surgical Infection Prevention Collaborative which is particularly relevant to pediatric neurosurgery.

The Collaborative is investigating (among other surgical procedures) infections in spinal surgery with instrumentation and ventriculoperitoneal shunt procedures at participating Children’s Hospitals across the United States. The purpose of these investigations is to improve the quality of care delivered to pediatric patients undergoing surgery through the application of evidence-based surgical infection prevention processes in the peri-operative period. The focus is on surgical prophylactic antibiotics (selection, timing, intra-operative re-dosing, and post-operative discontinuation), appropriate hair removal, and skin antisepsis before surgery and basic surgical infection prevention strategies.

These studies are motivated by concerns about finances and quality of care. The financial impact of failing to prevent surgical site infections (SSIs) in the pediatric population was addressed by Sparling et al. (Qual Manag Health Care 16:219-225, 2007). Using a matched cohort design, researchers found that in 16 patients with SSIs compared...
to 16 patients without SSIs, the length of stay (LOS) was increased by 10.6 days and the cost increased by $27,288 per patient. Given that an estimated 40-60% of SSIs are preventable, this is an area where an implemented program could have significant positive impact on both medical and fiscal outcomes. With respect to spinal surgery, Labbé et al. (Infect Control Hosp Epidemiol 24:591-595, 2003) found that pediatric patients with myelodysplasia are at higher risk for SSIs after a spinal fusion and emphasize the importance of optimal antibiotic prophylaxis to minimize that risk.

Texas Children’s Hospital participated in the CHCA Surgical Infection Prevention Collaborative project and has successfully changed its practices in pediatric neurosurgery. This change came about due to strong physician leadership and significant collaborative effort from medical personnel and departments, including Quality and Outcomes Management, surgeons, residents, fellows, peri-operative nursing, anesthesiology, OR educator, and wound/ostomy nurses.

Most important was a change in culture. Step-by-step practices were instituted, such as mandating using clippers to cut hair by removing razors (which can increase staph aureus infection by causing small skin lacerations) from the operating room. Because the timing of prophylactic antibiotics (PA) is critical (with the optimal initial dose given in the 30-60 minute window prior to the initial incision), there was additional monitoring for spinal cases requiring more lines and positioning. Normothermia, also important in reducing SSIs, was addressed by pre-warming rooms and the use of french fry lights, Bair-Paws Gowns for spinal patients, and Bair Huggers for all ventriculoperitoneal shunts.

Skin preparation for patients was performed using Chlora-prep, with the application proceeding like painting a fence (up and down and then side to side- over the same surgical area until it is “clean”) and the Chlora-prep being allowed to dry for 2-3 minutes. Finally, wound/ostomy nurses were trained in dressing care for spinal patients so the patients would receive standardized postoperative wound care. While the project is ongoing, the trend was toward a reduction in shunt infections. (As the starting shunt infection rate was quite low, it is not statistically significant at this time and a reduction in spine infections is minimally statistically significant at this point in the project.)

The experiences of the Collaborative implemented at Texas Children’s Hospital provide vital information for neurosurgeons. A collaborative effort with all personnel involved and strong physician leadership with administrative support (both from a managerial and a financial standpoint) is essential for a successful project. We, as pediatric neurosurgeons, must take a leadership role in selecting and implementing these types of outcomes projects as our future depends upon our active and effective participation.

In the past, the purpose of looking at outcomes was to continually improve the quality of care that is delivered to patients. Now, outcomes are being sought as a means to grade physicians and hospitals with an aim to impact reimbursement for medical services. One has only to look at the recent trend for Medicare to refuse to reimburse for the cost of treating avoidable complications such as urinary tract infections to see how the results of these outcomes will be applied. It is therefore imperative that pediatric neurosurgeons become the leaders in selecting meaningful outcomes to investigate in multi-center collaborative studies. CNSQ
OUTCOMES STUDIES IN FUNCTIONAL NEUROSURGERY

Surgical treatment of functional brain disorders has evolved over the last 20 years. A unique set of circumstances has mandated rigorous outcomes studies in order to justify the use of new and costly technologies. The transition from ablation to neuromodulation has enhanced the safety of surgery, improved clinical outcomes, and expanded the potential indications for surgical intervention. At the same time the use of implantable devices has dramatically increased surgical costs and has added two additional layers of regulatory bureaucracy: the Food and Drug Administration (FDA), which oversees device approvals, and the Centers for Medicare and Medicaid Services (CMS), which determines for what therapies the government (and in turn private insurers) will reimburse. Furthermore, because physicians other than neurosurgeons traditionally treat patients with movement and other functional disorders outcomes studies have been necessary to illustrate to referring physicians the need for these procedures. In this section, I will highlight the key studies that moved this field forward, and point out some lessons learned.

Deep Brain Stimulation
The advent of chronic electrical brain stimulation as an alternative to thalamotomy and pallidotomy has revolutionized the treatment of Essential Tremor, Parkinson’s disease and torsion dystonia. While deep brain stimulation (DBS) technology was rapidly adopted in Europe and Canada, acceptance has been slow in the U.S. Well-designed outcome studies have been required in order to gain regulatory approval as well as acceptance among the neurology community and patient advocacy groups.

Tremor
The FDA granted its first approval for the Activa™ deep brain stimulation (DBS) system in 1997, specifically for unilateral use in the thalamus to treat Parkinsonian and Essential Tremor (www.fda.gov/cdrh/pdfp960009.pdf). This approval was based on a multi-center U.S. and European trial. In both instances, researchers employed validated tremor rating scales to document tremor severity with and without stimulation. The use of evaluators who were blinded to the patients’ stimulation status set an important standard for future DBS trials. The results of both trials demonstrated conclusively that thalamic DBS effectively suppresses medically refractory Parkinsonian and Essential Tremors; however, only for patients with ET does tremor suppression result in significant reductions in disability and medication requirements. A later randomized prospective trial, published in the New England Journal of Medicine (NEJM), demonstrated the superior safety of thalamic DBS, as compared to thalamotomy, in patients with refractory tremor.

Parkinson’s Disease
In 2002, the FDA approved Medtronic’s Activa™ system for use in the globus pallidus internus (GPI) and subthalamic nucleus (STN) for the treatment of advanced Parkinson’s disease (www.fda.gov/cdrh/pdf/P960009S007b.pdf). In some ways, this approval may be seen as a step backwards as it was based on prospective open-label data with no placebo control. This multi-national, multi-center trial included four U.S. centers. Patients were treated with stimulation at either the GPI (N = 47) or STN (N=92). The previously validated Unified Parkinson’s Disease Rating Scale (UPDRS) motor sub-score (UPDRSIII) served as the primary outcome measure. The study demonstrated significant improvements in the UPDRSIII scores 12 months after surgery. Stimulation at either target increased motor “on time” by more than 6 hours, on average, and reduced levodopa-induced dyskinesia significantly. The results were deemed so impressive in this group of advanced PD patients that approval was granted despite a 7.5% incidence of intracerebral hemorrhage and a 10.6% incidence of infection.

Dystonia
The FDA granted a Humanitarian Device Exemption (HDE) for stimulation of the GPI with the Activa™ device for the treatment of primary dystonia in 2003 (www.fda.gov/cdrh/pdf2/H020007b.pdf). This approval was based on a meta-analysis of 201 dystonia patients whose responses to DBS were reported in 34 publications. The FDA felt that this largely retrospective, open-label data in patients with primary dystonia was so robust and so superior to all other treatments available to this small patient population that compelling the performance of a true pivotal trial represented an unfair burden.

Since then a beautifully executed German multi-center trial has generated Class I evidence of the efficacy of pallidal DBS for primary dystonia. Patients were randomized to
either therapeutic or sham stimulation for the first three months following device implantation. Blinded raters evaluated the patients employing the Burke-Fahn-Marsden Dystonia Rating Scale at baseline, and after three months of treatment, therapeutic stimulation was given to members of the control group. A true crossover study was correctly deemed to be unethical. The study demonstrated a clinically and statistically significant reduction in dystonia severity in those receiving therapeutic stimulation. Minimal improvement was noted in the control group, which then went on to demonstrate improvements similar to those seen in the treatment group over the subsequent three months.

**Vagus Nerve Stimulation**

Chronic intermittent electrical stimulation of the left vagus nerve was developed by Cyberonics. Thus far, two indications for vagus nerve stimulation (VNS) have been FDA approved.

**Refractory Epilepsy**

The FDA approved VNS for epilepsy in 1997, based on two pivotal multi-center U.S. trials, termed E03 and E05. In the E05 study 254 patients aged 13-60 with intractable partial epilepsy underwent surgical implantation of a left VNS system. Two weeks after surgery patients were randomized into high or low stimulation groups (the latter serving as a sham stimulation control), an important milestone for neuromodulation outcome studies. The high stimulation group experienced a 28% reduction in seizure frequency as compared to a 15% reduction in the low stimulation group (P = 0.04). A total of 310 patients completed the E03 and E05 trials with a mean seizure reduction of 25-30% as compared to baseline.

**Refractory Depression**

The approval of VNS for refractory depression has been somewhat more controversial. Following compelling open-label results, which suggested a significant and durable response rate as measured by psychiatrists employing validated rating scales, Cyberonics instituted two pivotal trials termed D02 and D04. The former was a prospective, double-blind, sham stimulation-controlled trial similar to E05. The 24-item Hamilton Depression Rating Scale (HD24) served as the primary outcome measure; the three-month follow-up was the primary endpoint. Response was defined as a 50% reduction in the HD24; remission, an absolute HD24 <10. At the primary endpoint, the response rate was 15% in those randomized to therapeutic stimulation and 10% in the control group (P=0.25). Only one of the secondary outcomes measures demonstrated a statistically significant difference between therapeutic and sham stimulation.

Despite this failure to demonstrate improvement that was statistically superior to placebo, the FDA granted approval on the basis of the D04 study in which the response rates of patients receiving VNS for 1 year were compared to a control group who received the best medical therapy, a far less rigorous standard scientifically. Despite FDA approval, the CMS has not approved payment for this intervention, presumably because of the failed placebo-controlled trial, a factor that has also slowed acceptance of this therapy among the psychiatric community.

**Motor Cortical Stimulation**

The cortex is an attractive target for neuromodulation for the simple reason that it may be readily and safely accessed. In fact, one may stimulate the cortex epidurally so that cortical stimulators may be implanted with minimal risk as compared to DBS. To date, cortical stimulation has yet to be approved for any indication though studies examining its use for the treatment of tinnitus, atypical facial pain and chronic depression are underway. Stimulation employed to enhance recovery following stroke is the most developed potential application.

Northstar Neuroscience, Inc. (Seattle, WA) is developing a product based on the scientific observation that sub-threshold cortical stimulation can enhance neuronal sprouting following injury. Three clinical trials have been performed. The Phase I and II trials were randomized controlled prospective trials, which documented both the safety and probable efficacy of sub-threshold motor cortical stimulation (MCS) during rehabilitation for stroke-related deficits. Combined, these two studies (32 patients) demonstrated that epidural stimulation overlying the thumb region of the motor cortex (as delineated with fMRI) during rehabilitation was safe and yielded clinically meaningful improvements in motor function that are superior to physical therapy alone (data from NorthStar Neuroscience, Inc.). Based on these two studies, a pivotal U.S. trial was instituted encompassing 151 subjects, where 100 subjects underwent surgical implantation of the stimulating device and 51 control subjects received rehabilitation alone. The upper extremity Fugl-Meyer (UEFM) score, a validated rating scale of stroke related disability, will serve as the primary outcome measure. Study enrollment was completed in May 2007 and results are expected in 2008.

**Transplantation, Growth Factor Infusion, and Gene Therapy Trials for Parkinson’s Disease**

The neurochemical and anatomic specificity that characterize the neuronal degeneration underlying Parkinson’s disease make it an attractive target for biological modifying therapies. To date, none of the above-mentioned approaches has been shown conclusively to halt the progression or reverse the symptoms of PD.

Both embryonic mesencephalic transplantation and direct putaminal infusions of glial-derived neurotrophic factor (GDNF) showed great promise in open-label uncontrolled studies of small numbers of patients. However, in subsequent sham-surgery controlled trials, these therapies failed to yield clinical improvements that were statistically superior to placebo. These experiences demonstrated the scientific and ethical necessity of conducting sham surgery controlled trials in PD. In addition, these and other studies have set the

Despite FDA approval, the CMS has not approved payment for this intervention, presumably because of the failed placebo-controlled trial, a factor that has also slowed acceptance of this study among the psychiatric community.
In evaluating the future of outcomes studies and the further development of functional neurosurgery, we need to address two flaws in our current thinking.

standard that only 12-month outcomes are clinically valid in PD; 3-month improvements do not justify the risk or expense of surgical intervention and may simply be the result of placebo responses.

These study design principles are presently being applied to ongoing studies of one cellular transplantation therapy and at least two gene therapies. Spheramine (Titan Pharmaceuticals, San Francisco, CA) consists of levodopa-producing retinal pigmented epithelial cells mounted on gelatin micro carriers. Spheramine is implanted stereotactically into the putamen, enhancing local levodopa levels, which achieve steady-state production in a few months. Preliminary open-label data suggest that Spheramine improves motor function in PD patients and reduces dyskinesiae. Moreover, because levodopa production is continuous, it is hoped that Spheramine will smooth out the motor fluctuations that characterize advanced PD. Following a successful, open-label, Phase I trial, an international multi-center, sham surgery trial was completed in 2007. Results of this study are due in mid-2008, allowing for all study enrollees to reach the 12 month follow-up evaluation.

Neurologix, Inc. (Fort Lee, NJ) has completed a successful Phase I trial of their Adeno-Associated Virus (AAV) mediated in vivo transfer of the glutamic acid decarboxylase (GAD) gene into the Subthalamic Nucleus (STN). The therapy is designed to enhance the production of the inhibitory neurotransmitter GABA within the STN, by increasing the local expression of GAD, which catalyzes the rate-limiting step in GABA production. In turn, the increase in GABA will suppress the STN hyperactivity that is observed in Parkinson’s disease and is believed to account for many of PD’s motor symptoms. The Phase I study of 12 patients, who received one of three different doses, was designed primarily to evaluate the therapy’s safety. Nevertheless, improvements in motor function were also demonstrated and the likelihood of improvement seemed to correlate with dose.

Cere-120 (Ceregene Inc., San Diego, CA) is an AAV mediated in vivo transfer of the gene encoding Nurturin (NTN), a neurotrophic factor known to promote growth and enhance the health of dopaminergic cells. As such, Cere-120 is the first therapy designed not only to improve PD symptoms in the short-term, but also to slow progression of the disease thereafter. Cere-120 is injected into the putamen bilaterally, enhancing local production of the growth factor, which is then taken up by the nigral nerve terminals and transported retrograde to the cell bodies. A phase I trial of Cere-120 was completed in 2006 and included two doses. Open-label evaluations in these 14 patients again demonstrated that in vivo AAV-mediated gene transfer is safe and suggest improvements in motor function that approximate those achieved with DBS (Ceregene, Inc., data submitted for publication). Results are expected by the end of 2008.

Conclusion

The rebirth and expansion of functional neurosurgery has depended on well-designed outcomes studies that established the safety and efficacy of fledging neuromodulatory technologies. Thus far, DBS has played the starring role, supplanting neuroablative for the treatment of medically refractory movement disorders. Emerging applications of neuromodulation (particularly for psychiatric disorders) will demand careful study employing even more rigorous methods than were required to gain prior approvals. Adoption of cell and gene based therapies will also depend on stringent short- and long-term outcome studies in order to prove efficacy and overcome preconceived fears of oncogenesis and encephalitis among the public and broader medical community.

A number of lessons have been learned over the last 10 years, not the least of which is that placebo effect must be accounted for in surgical trials as it is in pivotal trials of new medications. The use of sham surgery controls is not only preferable but ethically mandated in order to prevent large numbers of patients from undergoing invasive and expensive procedures that are ineffective. Clinical efficacy must be measured with validated objective rating scales, preferably by blinded raters, so that the response to new procedures may be reliably compared to currently approved therapies and also in order to gain acceptance for surgical intervention among referring physicians and third party payers.

In evaluating the future of outcomes studies and the further development of functional neurosurgery, we need to address two flaws in our current thinking. First, while randomized, prospective, double-blind trials remain the gold standard for demonstrating clinical efficacy, the FDA and CMS must also be flexible enough to accept (when circumstances dictate) robust open-label results that might impair researchers’ ability to recruit patients for sham-controlled trials. In some circumstances, it may not be ethically or logistically possible to perform a sham-controlled trial if open-label results for a devastating illness are compelling and patients have means other than a trial to access the therapy.

In the final analysis, the American people have the greatest interest in having these studies performed in as scientifically and ethically sound manner as possible, and deserve access to the results of such studies, regardless of outcome, for the purposes of determining how best to spend their healthcare dollars. Moreover, the prohibitive costs of performing outcomes studies in the U.S. forces American device manufacturers to outsource many trials to Europe and Canada to the detriment of U.S. research facilities. Greater NIH involvement in the funding of device-related clinical research will help to reverse this trend.

Healthcare expenditures in the U.S. exceeded $2 trillion in 2007 and continue to rise at twice the rate of inflation. With the prospect of an expanded federal healthcare program ever more likely and a population that continues to age, robust clinical evidence including cost-effectiveness analyses will be required to justify payment for new medical technologies. Consequently, the performance of scientifically and ethically sound outcomes studies will be critical to the future development of functional neurosurgery.
Surgery of the Human Cerebrum
A Three-Part Supplement
Michael L.J. Apuzzo, Editor

June 2008 Part 3 Topics
Psycho-affective Disorders and Pain
3D Microsurgical and Tractographic Anatomy
Intraventricular Tumors
Extra Axial Lesions
Basal Lesions
Giant Aneurysms
Revascularization
Endovascular Techniques
Advanced Methodologies

Giovanni Broggi, MD
Albert L. Rhoton, Jr., MD
M. Gazi Yaşargil, MD
Peter M. Black, MD, PhD
Madjid Samii, MD, PhD
Robert F. Spetzler, MD
Laligam N. Sekhar, MD
L. Nelson Hopkins, MD
Michael L.J. Apuzzo, MD

For more information, contact the Editorial Office
Ph: (323) 442-3001 • Fax: (323) 442-3002 • Email: neurosurgery-journal@hsc.usc.edu

Neurosurgery
Degenerative Spondylolisthesis with Lumbar Spinal Stenosis: an Evaluation of the SPORT Results

Peter D. Angevine, MD, MPH
New York, NY

Michael G. Kaiser, MD
New York, NY

Paul C. McCormick, MD, MPH
New York, NY

In May 2007 the New England Journal of Medicine published the results of the Spine Patient Outcome Research Trial (SPORT), comparing the outcomes of the non-operative and operative treatment of patients with symptomatic lumbar spinal stenosis and degenerative spondylolisthesis (N Engl J Med 356:2257-2270, 2007). Initiated in March 2000, SPORT was intended to investigate the effectiveness of surgical versus non-surgical treatments for three common low back disorders: herniated lumbar disc (HLD), spinal stenosis (SS) and lumbar spinal stenosis with degenerative spondylolisthesis (DS). This highly publicized National Institutes of Health (NIH)-funded clinical trial was conducted at 13 centers in 11 different states and cost an estimated 21 million dollars. Dr. Steven Katz, Director of the National Institute of Arthritis and Musculoskeletal and Skin Diseases (NIAMS), said before the study began “...based on this trial, we shall, for the first time, have scientific evidence regarding the relative effectiveness of surgery versus non-surgical treatment of these commonly diagnosed lumbar spine conditions.”

Similar to the SPORT HLD study (JAMA 296:2451-2459, 2006; JAMA 296:2441-2450, 2006) (the results of which were published 6 months prior) the DS trial was based on a simple “either-or” treatment comparison. The result was a trial with limited accord with the real-world management of patients with DS. The trial was also compromised by the high proportion of patients that did not undergo their randomly assigned treatment. The magnitude of crossover seen in the DS trial diminishes the strength of the randomized clinical trial (RCT) design by uncoupling the treatment received from the treatment assigned. Intent-to-treat analysis no longer provides a valid estimate of the effect size of the investigational intervention. Analysis by treatment received (“as treated” analysis), may yield a more accurate estimate of treatment effect but is potentially subject to significant bias due to the non-random segregation of subjects to the different treatments. Presumably to address these concerns, the authors of the DS study manuscript included the results of both the ITT and the as-treated analyses. These results, the implications of their differences, and other aspects of the SPORT DS trial which are of importance to neurosurgeons are reviewed.

Summary of the SPORT Results
Patients presenting with at least 12 weeks of neurogenic claudication with or without associated neurologic deficits and radiographic confirmation of both spinal stenosis and a single level degenerative spondylolisthesis who were considered to be surgical candidates were eligible to enroll in the SPORT DS trial. Initially patients were offered enrollment in the randomized arm of the study or a parallel observational arm in which care would be determined in consultation with his or her physician.

Patients who were treated non-operatively received “usual care;” including, at least, active physical therapy, nonsteroidal anti-inflammatory agents (NSAIDs), and education or counseling regarding home exercise. Surgical treatment was also not rigidly dictated; acceptable procedures included a posterior decompressive laminectomy without fusion, a decompressive laminectomy with arthrodesis, and a decompressive laminectomy with instrumentation and arthrodesis.

Patients were evaluated at 6 weeks and 3, 6, 12, and 24 months following enrollment. The primary outcomes measures were the bodily pain and physical function scores of the SF-36 General Health Survey and the overall score on the modified Oswestry Disability Index (ODI). Secondary outcomes consisted of patient self-reported satisfaction and improvement and the Stenosis and Low Back Pain Botherlessness Scales.

A sample size of 150 patients in each treatment arm of the randomized cohort, based on a two-sided t-test, was calculated to be necessary for an 85% probability of detecting a 10 point difference on either primary outcome instrument if, in fact, such a difference existed. This calculation accounted for an anticipated 20% lost to follow-up but did not account for noncompliance with treatment assignment leading to crossover between treatment groups.

Of the more than 1100 patients who were screened, 892 were con-
It is essential for neurosurgeons who treat patients with degenerative lumbar disorders to understand the design, analysis, and results of the SPORT study as they were reported, the limitations of the study, and the potential misinterpretations of the trial.

Considered eligible for participation in the study, 285 candidates declined to participate, 304 patients agreed to be randomized, 303 subjects participated in the observational cohort and 17% were lost to follow-up. Baseline demographic characteristics were nearly identical between the randomized and observational cohorts. Significantly worse pain and disability scores were observed, however, in those patients who ultimately underwent surgery in either the randomized or the observational cohort.

At six months, 53% of the participants in the randomized cohort assigned to undergo surgery had had surgery, while 38% of patients assigned to non-operative treatment had undergone surgery. Ultimately, 64% of patients in the surgical cohort and 49% of those in the non-operative group had surgery by the 2-year follow-up. The proportion of patients in the observational cohort who had surgery by the six month follow-up visit were 95 and 7% among those initially selecting operative and non-operative treatment, respectively; the corresponding proportions at the 2 year follow-up were 97 and 25%.

The intent-to-treat analysis demonstrated no statistically significant difference in the primary outcome measures. The as-treated outcomes for the combined randomized and observational cohorts revealed a statistically significant improvement following surgery for both primary and secondary outcome measures. The treatment effects at 2 years based on the as-treated analysis were as follows: 18.1 points (95% CI, 14.5 to 21.7) on the SF-36 bodily pain subscale, 18.3 points (95% CI, 14.6 to 21.9) on the SF-36 physical function subscale, and an improvement of 16.7 points (95% CI, 13.9 to 19.5) on the ODI. The as-treated analysis was similar in the randomized and observational cohorts. A post-hoc subgroup analysis revealed no difference in treatment effect based on sex, history of tobacco use, baseline symptom severity, duration of symptoms, treatment preference, number of stenotic levels, severity of stenosis on imaging and number of co-existing conditions.

Critical Analysis

It is essential for neurosurgeons who treat patients with degenerative lumbar disorders to understand the design, analysis, and results of the SPORT study as they were reported, the limitations of the study, and the potential misinterpretations of the trial.

Several specific areas of concern of the SPORT DS are outlined below. Beginning with the study’s fundamental hypothesis, the conception of the trial was flawed. A dichotomous “either-or” formulation of the treatment of lumbar stenosis with degenerative spondylolisthesis does not reflect the management of patients with this disorder. It is unclear, regardless of the specific results of the study, how the answer to the clinically irrelevant question “which single treatment is better” would be applied to the practice of medicine.

A second concern is the representativeness of the study population. It is likely, although not possible to know, that the cohort enrolled in the study (particularly in the randomized arm) differs from the overall population with DS. This makes the generalization of any study result to all patients tenuous.

A third area of concern is the analysis of the results; limitations of both the intent-to-treat and as-treated analyses are discussed. The ITT analysis is biased toward the null by the high number of crossovers in each treatment allocation group, while the as-treated analysis is likely biased as well. Finally, the heterogeneity of treatments in both the operative and non-operative cohorts leaves significant uncertainty regarding the outcomes of specific treatment strategies in individual patients.

Study Hypothesis

The null hypothesis for the study was that no difference in primary outcomes exists at the measured time points between surgical and non-surgical treatment. This dichotomous hypothesis ignores the actual treatment of patients with the disorder. Almost all patients are initially treated non-operatively. Those individuals who improve without surgery continue with that treatment; those patients who fail to improve sufficiently may chose to undergo an operation. Practitioners who treat these patients understand that non-operative treatment and spinal surgery are not interchangeable interventions that are equally appropriate in all patients who may be considered to be “surgical candidates.” Surgery is reserved as an option for those patients who, having tried non-operative treatment, continue to have significant pain or disability.

By structuring the study as a head-to-head comparison of surgery and non-operative treatment the investigators imply that there is a single one-size-fits-all “answer” for all patients with lumbar spinal ste-
nosis and degenerative spondylolisthesis. The result is that SPORT provided us with an expensive demonstration that treatment assignment independent of the patient’s symptom severity, response to non-operative treatment, and self-assessed quality of life is a non-viable management strategy. For instance, a moderately symptomatic patient who is a candidate for surgery but is improving with physical therapy and is reasonably satisfied with his or her quality of life would be unlikely to elect to undergo an operation, regardless of the results of the study. Similarly, a patient with a severely compromised quality of life who, despite aggressive non-operative treatment, is experiencing worsening symptoms is not likely to be satisfied with continuing that treatment and would probably wish to have surgery. It is not clear, therefore, what impact even a clear-cut result in favor of one treatment or the other would have on clinical practice.

The study hypothesis also did not account for the potential heterogeneity of responses to the treatment alternatives. Previous reports have demonstrated that patients with differing disease severity respond differently to the same treatment (Spine 25:556-562, 2000; Spine 30:936-943, 2005). Basing the primary analysis on a comparison of the mean improvement may obscure a significant benefit for a subgroup of patients. Lost in the comparison may be specific subgroups of patients, for example those with severe or worsening symptoms, that experience a greater than average benefit from one treatment. Failing to specify these subgroups a priori commits the investigators to post hoc analysis, which will likely be underpowered and subject to justified methodological criticism. Clinicians are subsequently left without the information that would be most valuable to them in the treatment of these patients: which patients are most likely to respond to each treatment.

**Study Population**

In order for the results of the study to be applicable to patients outside of the study who have symptomatic DS, the study population must be representative of the underlying population with the disorder. Patients who elect to participate in studies, randomized or not, are likely to differ from those who decline enrollment. These differences between the study population and the overall patient population may include measured and unmeasured factors associated with outcome from treatment. The net result is that it is difficult to know the correspondence between the results of the SPORT trial and the expected outcomes of patients outside of the study.

**Intent-to-Treat Analysis**

Intent-to-treat analysis, the inclusion of the results of all patients in the randomly constituted treatment cohorts regardless of the treatment actually received, is the gold standard method of analysis for RCTs. Only the initial groups established through random assignment may be assumed to be balanced, on average, for measured and unmeasured factors that may affect outcome of treatment. This analysis is based on the assumption, however, that a high proportion of patients receive the treatment to which they are assigned. Otherwise, the analysis de facto compares the effect of a program of treatment assignment, not of the treatment itself.

In the current study, 49% of patients in the non-operative cohort ultimately decided to undergo surgery, while 36% of patients assigned to surgery elected to forgo an operation. This high proportion of crossover reduces the apparent difference in effect between the treatments as the cohorts become similar in terms of treatment received (the one factor that is supposed to differ between them in the RCT design). In fact, the investigators did not account for any crossover in their study. An increase in the sample size would have increased the power of the study to return a statistically significant result, but would not have mitigated the dilution of effect created by the crossover.

In the setting of such a high crossover rate the analysis is biased toward the null, or there is a high likelihood of such a study finding a minimal effect size for the experimental treatment. The resulting
Although some potential sources of bias can be controlled for statistically, it is likely that there are unmeasured or unknown patient factors that may influence the results. These factors cannot be accounted for and will bias the results of any analysis other than intent-to-treat, which draws on the strength of random allocation to balance all potential biasing factors.

comparison is actually of the effect of a program of assigning patients to receive a treatment, not of the effect of the treatment itself. This important distinction was not discussed by the authors and is often overlooked by the casual reader.

As-treated Analysis
In contrast to the report of the herniated lumbar disc arm of SPORT, the DS paper reported the as-treated analysis of the combined randomized and observational cohorts. This approach, an attempt to overcome the problems with the intent-to-treat analysis outlined previously, is not without its own limitations. The allocation of patients to treatment by any method other than random assignment has the potential for significant bias, in which the perceived difference between outcomes from treatment is actually due to differences between the cohorts receiving the interventions. Although some potential sources of bias can be controlled for statistically, it is likely that there are unmeasured or unknown patient factors that may influence the results. These factors cannot be accounted for and will bias the results of any analysis other than intent-to-treat, which draws on the strength of random allocation to balance all potentially biasing factors.

In the presence of a high proportion of crossovers and with appropriate adjustments for baseline differences between cohorts, the as-treated analysis may, however, provide a more accurate estimate of the effect size than intent-to-treat analysis. However, the nature of the treatments is such that there still exists a significant potential source of bias against surgery since it is not possible for patients who have had surgery to crossover to the non-operative cohort. Patients who fail non-operative treatment, however, may crossover to surgery and their outcomes, good or poor, are pooled with the surgical group. Over time, then, the surgical cohort becomes a repository for the most severely affected patients and for those who are not satisfied with their responses to non-operative treatment. There is a reasonable likelihood that patients who are unlikely to improve with any treatment will end up in the surgical group. Averaging baseline pain and disability scores further favors the non-operative cohort, since the net effect of this adjustment is to diminish the magnitude of improvement observed following surgery and exaggerate the benefit of non-operative therapy because the surgical group had worse scores at study entry. This may be one reason for the smaller magnitude effect observed for surgery in this study compared to previously published reports.

Heterogeneity of Treatments
The variability of surgical and non-surgical treatments is an additional, but perhaps unavoidable, limitation of the SPORT study. Neither treatment method was standardized. Non-operative treatment was recommended to include at least active physical therapy, NSAIDs, and education or counseling regarding home exercise. Surgery was allowed to be laminectomy, laminectomy with arthrodesis, or laminectomy with instrumentation and arthrodesis, at the surgeon’s discretion. Advocates of particular non-operative treatment protocols will no doubt decry the lack of a standardized treatment protocol, and including three distinct procedures in the study introduces another variable into an already complicated situation. In fact, almost 75% of the procedures were laminectomies with an instrumented fusion. As the study was not designed a priori to determine differences in outcomes between different surgical procedures, the optimal surgical treatment will have to be determined in a separate investigation.

Conclusions
While the resource-intensive SPORT RCT of DS suffered from serious limitations in its design and analysis, it did establish two important facts.

First, it demonstrated that the treatment of patients with lumbar spinal stenosis, with DS (based on symptom severity and educated patient choice) leads to the appropriate use of surgery. Patients who select surgery, on average, receive a clinically and statistically significant benefit that exceeds the magnitude of the improvement of those patients who select non-operative treatment. This study therefore vindicates the contemporary strategy of evaluating and managing patients with degenerative spondylolisthesis and lumbar spinal stenosis that has evolved over time.

Secondly, when treatment alternatives for benign but painful conditions are freely available, sequentially applied (rather than interchangeable), and differ in their risk and recovery profiles, an RCT will be a practical impossibility even with the imprimatur and financial support of the NIH. Under these conditions, rational patient choice will tend to override treatment assignment. Patients will have had experience with the non-operative treatments and be able to judge their responses to that intervention. Based on these perceptions of their responses, patients will decide either to continue with the same care or to undergo an operation.

Based on the results of the as-treated analysis (which in the setting of significant patient crossover are expected to more accurately estimate the true benefit of each treatment), clinicians may inform their patients with lumbar stenosis and DS that those who undergo surgery have, on average, better outcomes up to 2 years afterward compared to those who undergo non-operative treatment. It is not, unfortunately, possible to provide an individual patient with a probability of improvement with either treatment based on his or her individual characteristics. For this most important and clinically relevant information, neurosurgeons will have to await the results of additional, more sophisticated studies. CNSQ
2008 DANDY ORATOR

Maya Angelou
Renowned Author, Actress, Poet

Maya Angelou is hailed as one of the great voices of contemporary black literature and as a remarkable Renaissance woman. She is a poet, historian, author, actress, playwright, civil-rights activist, producer and director. A mesmerizing vision of grace, swaying and stirring when she moves; Dr. Angelou captivates her audiences lyrically with vigor, fire and perception. She has the unique ability to shatter the opaque prisms of race and class between reader and subject throughout her books of poetry and her autobiographies.

Dr. Angelou has authored twelve best selling books and numerous magazine articles earning her Pulitzer Prize and National Book Award nominations. Dr. Angelou, poet, hit the bestsellers lists with I Know Why the Caged Bird Sings, a chronicle of her life up to age sixteen, which was published in 1970 with great critical and commercial success. In 1993, Angelou became the second poet in US History to have the honor of writing and reciting original work at the Presidential Inauguration. On the Pulse of Morning, at Bill Clinton’s presidential inauguration, was an occasion that gave her wide recognition for which she was awarded a Grammy award (best spoken word).

In the film industry, Dr. Angelou’s work in script writing and directing, has been groundbreaking for black women. She has written and produced several prize-winning documentaries and her screenplay Georgia, Georgia, was the first by a black woman to be filmed. She was also nominated for an Emmy Award for her acting in Roots. In theatre, she produced, directed and starred in Cabaret for Freedom in collaboration with Godfrey Cambridge; starred in Genet’s The Blacks at St Mark’s Playhouse; and adapted Sophocles Ajax, which premiered in Los Angeles in 1974.

Register Online by August 20, 2008!
www.cns.org
How much profit is the hospital you are affiliated with making on the cases you do? Many neurosurgeons want to know but may have a difficult time getting the figures. There are several reasons for this. Hospital administrators may not be anxious to reveal the amount you bring in because they may assume you will ask that some of it be reinvested in neurosurgery. Those in administration could also be concerned about losing a bargaining position once you know what their profits are. It is possible that they will give you excuses about where funds are going, citing the need to support other ventures that cannot make money. They may say they do not have the data you request because their systems do not track that information, or even that they do not make much money in the current insurance and reimbursement environment.

In my experience, none of these explanations is credible. Can you imagine that the CEO of a hospital, a $200 million business, does not know who is making money for the hospital and how much is coming in? The hospital’s Board of Directors would certainly need this information. The profits you generate may be going into the category called “overhead,” which is an expenditure account for the administration.

Here are some basic principles of hospital accounting. On profit and loss sheets, the first item listed is the Gross Revenues from your cases (the amount they charge the various insurers). Next is listed Net Revenues (the amount the hospital collects after the insurance companies pay them the negotiated rate). Finally, Direct Costs related to neurosurgery (such as equipment, nursing, etc.) are deducted.

The figure remaining is called the Contribution (to) Margin. This is how much money you actually generate for the hospital. A portion of this amount may go to overhead, called Indirect Costs, that includes support for other losing ventures, administrative salaries, advertising, security, etc. Once those who formulate the budget subtract that figure from the Contribution Margin, they are left with the Net Profit. This may be the figure you are given, and it may seem to indicate that you
You and you need them. They do not want you to take your business elsewhere.

Remember one last thing: you use the same logic every day that a businessperson uses. You confront a problem, get the facts (History and Physical Examination), look at a series of options to solve the problem (Differential Diagnosis), evaluate the possibilities (Test the various Differential Diagnoses), and then make a choice (Diagnose). You repeat this process many times a day and hundreds of times a year.

Your goal is to help people, and a businessperson’s goal is to make money. Yet in the best business situations, where people care about their customers, they are the most successful. Remember that for the hospital administrators you, not your patient, are the customer. You bring the patient to the hospital. You are a satisfied customer when you and your patients are happy.

Also, you can understand the concerns of operating within a budget because you do so in your office and at home. Hospitals are no different.

Our society rewards those who are winners. You are a winner for the hospital, and you want to be rewarded for what you do. The hospital administration does not want to lose you, but at the same time they may want to share with you as small a portion of your profits as possible.

The material in this article is an opinion. The views of the author do not reflect nor express any opinions, beliefs, or thoughts of the CNS or the CNSQ.

In order to facilitate discussion and opinions on all aspects of neurosurgery, we have devoted this portion of the CNSQ to you, the reader. Please send your responses to this article, as well as any other opinions or articles, to info@1cns.org.
It is with much sadness that we report the passing of Julius Goodman, MD, on Sunday, January 27, 2008. Neurosurgeons will recall him as leader, director, and guiding force of the Neurosurgery Oral Board Course, which now carries his name.

Dr. Goodman was a co-founder of the Indianapolis Neurosurgical Group in the 1970s, which grew to be the largest private practice neurosurgery group in the country. He had an active clinical practice and specialized in the treatment of pituitary disorders. As an innovator and educator, he accepted and utilized advancements in the field. For example, he was the first neurosurgeon in the Indianapolis region to perform a transsphenoidal approach to a pituitary tumor.

Dr. Goodman may be best remembered for developing and organizing the ABNS Oral preparatory course. He composed this course to assist neurosurgeons in preparing for the ABNS Oral boards. Attendees can distinctly recall being the target of specific and direct questioning on neurosurgical issues. Most can also remember the joy of being placed on the “hot seat,” as we were questioned on the pertinent anatomy, pathophysiology and treatments for carpal tunnel syndrome or other disease processes. Even if we answered the majority of these questions incorrectly, Dr. Goodman was able to restore examinees’ composure.

For his exceptional ABNS course that condensed the neurosurgery residency program into several days’ review, and for many other valuable contributions to our field, we join with his many appreciative patients and colleagues in expressing our gratitude to Dr. Goodman and our sincere condolences to his family. CNSQ
Editor’s Note: Spinal cord injury is a devastating event not only for the individual but for society as a whole. Recently this issue has been highlighted on national television by the injury and subsequent recovery of Kevin Everett of the Buffalo Bills. The media has attempted to analyze and sensationalize varying treatment protocols for this patient. In order to help clarify the course of treatment for Mr. Everett, the CNSQ has asked Kevin J. Gibbons, the neurosurgeon involved in this case, to review and comment on the patient’s condition and the care and treatment he received.

The Medical Care of Kevin Everett: Setting the Record Straight

In early September 2007, during the NFL Buffalo Bills’ home opening game, tight end Kevin Everett sustained a serious spinal cord injury. The public and serious nature of the injury, the treatments Everett reportedly received, and his significant recovery have attracted national attention.

The injury occurred on the opening kickoff of the second half when Everett tackled the kick returner for the Denver Broncos, Domenik Hixon. He remained facedown on the field after the collision and after several minutes it became apparent he was tetraplegic.

The subsequent care that he received was widely, sensationaly, and often incorrectly reported in the lay media (“Spinal cooling key to Everett’s recovery;” CBSNews.com, 9/13/2007). The medical media was similarly effusive in attributing his recovery to the use of moderate hypothermia (“On-field hypothermia saves Everett;” SpineUniversity.com).

For the edification of the neurosurgical community, here are the facts regarding the neurologic injury and subsequent medical care provided to Kevin Everett.

The Medical Team

Many individuals were involved in the care of Kevin Everett. Andrew Cappuccino, MD, a team physician for the Bills and orthopedic spine specialist, directed the prehospital care and was the attending physician of record. I was the consulting neurosurgeon and directed Everett’s care in the neurosurgical ICU after surgery in my role as intensivist and ICU director. Further, Dr. Cappuccino and I were co-surgeons for the decompression and stabilization of the fracture dislocation of the spine performed the evening after his injury. Neurosurgical residents and fellows from the SUNY Buffalo Department of Neurosurgery were involved throughout, and, as is always the case, the bulk of critical care was provided by skilled nurses and therapists.

Prehospital Care

Everett was injured at 1435 on September 9th, 2007. The Buffalo Bills’ medical staff, training staff (under the direction of Bud Carpenter), and the EMS team provided excellent and carefully planned immobilization of the injured player. Dr. Cappuccino directed the on-field care and assessment, determining that the player was tetraplegic but maintained a degree of crude touch sensation, and was breathing adequately despite a feeling of dyspnea. Within 15 minutes of injury, Everett was in an ambulance where he received intravenous methylprednisolone (NASCIS II doses) and an intravenous infusion of chilled saline was started. Approximately 2 liters of chilled saline were given prior to his arrival at the hospital. During this time he remained without any voluntary movement of his extremities, was awake and conversant, mildly dyspneic, and complained of a numb feeling throughout his body.

ER Evaluation

Everett was taken to Kaleida Health’s Millard Fillmore Gates Circle Hospital (a comprehensive stroke center and home of the SUNY Buffalo Neurosurgery Department) due to the availability of 24/7 CT and MRI and the fully equipped neurosurgical operating and neurointerventional suites. On arrival at 1520, his vital signs were a recorded temperature of 98.x °F (36.6 °C; the “x” superscript is not legible), pulse 67, BP 179/91, respirations 21, and oxygen saturation 100%.

Further neurologic assessment revealed normal level of consciousness, normal cranial nerve function, no voluntary movement of any extrem-
ity, crude touch pressure sensation on his trunk and extremities, and hyperreflexia with non-sustained clonus in his legs. NASCIS II steroid infusion continued. A forced vital capacity obtained preoperatively was 660 mL, less than 10% of predicted. He was determined to have a C4 ASIA B incomplete spinal cord injury, due to preserved sensation.

**Imaging**

A cervical lateral x-ray demonstrated a C3-4 subluxation of less than 50%, with kyphosis and a rotatory component. An immediate CT scan of the cervical spine demonstrated an unilateral jumped facet on the left at C3-4. The right C3-4 facet was somewhat flexed, with the inferior articulating process of C3 just short of a true perch, but without a fracture. An MRI further demonstrated the spinal deformity and ventral compression on the cord from the superior posterior lip of C4. MRI T2 imaging revealed cord hyperintense signal at the level of the injury without T1 imaging abnormalities. A left vertebral artery occlusion was noted.

**Surgery**

Everett was taken to the operating room at 1730 (3 hours from the injury) where an attempt at an awake closed reduction with Gardner–Wells tongs was performed. This improved his overall spinal alignment but was unable to reduce the left jumped facet and subsequent translation. Therefore an open reduction was performed. Surgery began at 1830 (4 hours from the injury) and the patient underwent an anterior discectomy with an open reduction performed by distraction (intervertebral body spreader) and manual manipulation of the tongs. This resulted in a complete reduction of the subluxation and normalization of the alignment. Intra-operatively, the posterior longitudinal ligament (PLL) was noted to be avulsed from C4, and the spine was reconstructed with a PEEK cage and anterior cervical plate. The posterior approach noted an extensive muscle hematoma, torn tendons and ligaments which were reconstructed with lateral mass fixation of C3-4. Due to the extent of the injury, laminectomies sparing the facets were performed at C3 and C4. An intraoperative ultrasound was performed and demonstrated the spinal cord to have its typical anatomic appearance, with maintained CSF spaces and pulsations. During surgery with passive cooling, his temperature ranged between 34.8 and 35.2 °C. Surgery was completed at 2135 (7 hours from injury).

**Postoperative Imaging**

After surgery, Everett remained intubated and sedated. CT and MRI were obtained within 4 hours after surgery, and demonstrated good decompression of the cord and decreased T2 hyperintense signal in the spinal cord.

**Postoperative Exam and Care**

A neurologic exam at 0500 the next day revealed trace 1/5 movement of his thigh adductors. Repeat exams over the next three hours revealed return of trace movement of hip flexors, quadriceps, and gastroc, and steady improvement in his adductors to 3/5. No voluntary movement in his arms was noted. His forced vital capacity improved to 1600 mL, albeit measured through a ventilator circuit.

Surface cooling was utilized in the postoperative period but was ineffective. His temperature quickly normalized to 98.4 °F (36.8 °C) by 0330 and then was rising above 99 °F despite surface cooling; shivering

From an initial anterior cord syndrome he progressively improved to a severe central cord syndrome, and has made dramatic improvement over the course of 4 months but remains with neurologic deficits in terms of his sensation, strength, coordination, and endurance.
Medical stories presented in the media are typically incomplete and often contain inaccuracies; thus, the interpretations that follow become exponentially more inaccurate.

was present. Between 0800 and 0900 an Alsius femoral cooling catheter was placed, and over several hours he was cooled to 92.3 °F (33.5 °C). Deep sedation and neuromuscular blockade prevented detailed neurologic exam until the next day. That day, Tuesday 9/11/07, he demonstrated return of function in his triceps 2-3/5, trace biceps 1/5, and antigravity 3/5 in his quadriceps. Rewarming was slowly begun, and he was extubated mid-day Wednesday, 9/12.

Hospital Course
Slow improvement in Everett’s leg function in all muscle groups and minimal return of arm movement (deltoids 0/5, biceps 1/5, triceps 3/5, FDI 1/5), opponens and abductor pollicis brevis 2/5) occurred during his 12-day stay in our hospital. His pulmonary capacity slowly improved but remained significantly impaired. A paralyzed left hemidiaphragm was diagnosed early in his hospital course. His hospital stay was further marked by septic episodes, pneumonia and a severe ileus. Continued intravascular cooling was utilized to maintain a normal body temperature for several days after moderate hypothermia was discontinued. He was in the neurosurgical ICU until transfer to the Texas Institute of Rehabilitation and Research in Houston.

Summary of Injury and Recovery
In summary, Kevin Everett sustained a severe cord injury, motor complete with some preserved sensation (ASIA B), without evidence of spinal shock in terms of either flaccidity or loss of sympathetic tone but with significant respiratory impairment. He was found to have a unilateral C3-4 facet-dislocation, with ventral cord impingement and kyphosis, but canal compromise was no greater than 50%. He underwent early assessment and diagnosis, early surgery, NASCIS II steroid treatment, and received hypothermic treatment with varying degrees of success (in terms of lowering body temperature) during the first 72 hours after injury. From an initial anterior cord syndrome he progressively improved to a severe central cord syndrome, and has made dramatic improvement over the course of 4 months but remains with neurologic deficits in terms of his sensation, strength, coordination, and endurance.

Discussion
There has been considerable attention drawn to this case through the national media and much of it credits hypothermia as the major cause of Everett’s recovery. I, however, believe the facts suggest otherwise.

Why did Everett Improve?
First, improvement was due to the nature of his injury. He had a unilateral facet dislocation, with less than 50% canal compromise, and kyphosis that, while adding to the angulation over the ventral spinal cord, actually improved the overall size of the spinal canal. The cord was not “transected” nor “crushed,” nor was there anterior-posterior scissoring action of the osseous elements (common with bilateral facet dislocations). This was a unilateral facet dislocation. A unilateral jumped facet is among the most favorable spinal dislocations. The majority of patients with this fracture pattern do not suffer spinal cord injury.

Second, he received prompt diagnostic evaluation and surgical correction of his cord compression, deformity, and instability in less than 5 hours from the injury.

Third, he received excellent prehospital care in terms of safe immobilization and transport. As most neurosurgeons and orthopedic spine surgeons acknowledge, in many spine injuries the neurologic damage is severe and irreversible from the point of injury, and in many other neck injuries even with fractures, the spinal column is relatively stable. In the case of Kevin Everett, we know now there clearly was reversible cord damage, and the ligamentous injuries he sustained were severe. His neck was grossly unstable, and this prevention of further movement, even of millimeters or a few degrees, may have prevented him from having an irreversible injury.

Fourth, he received high dose steroids in the minutes after his injury. Although still controversial, steroids may be of benefit in cases with reversible extrinsic cord compression, and the sooner they are given the better.

Finally, hypothermia may have been one of several agents of benefit in terms of the speed of his recovery. Although it is cited in the lay media as the most significant intervention in this case (Big Chill Saves NFL Player), I do not believe hypothermia was the critical intervention. First, the chilled saline infusion given en route to the hospital had a negligible effect on his temperature recorded on arrival in the ER. Most serious spinal cord injuries come in cooler without benefit of any efforts to cool them. His temperature during surgery was fairly typical for neurosurgical patients undergoing general anesthesia. His definite improvement in the 12 hours after surgery (with clear progressive return of motor function in his legs, doubling of his preoperative forced vital capacity, improvement in the hyperintense T2 cord signal changes on MRI) was well on its way before placement of a femoral cooling catheter system and the induction of moderate hypothermia.

Further study is warranted for victims of acute neurological injury and ischemia, whether direct trauma, spontaneous intracranial hemorrhage, or after cardiac arrest or stroke. Hypothermia may prove to be of use in some or all of these disease states. However, more study is warranted and funding required. One obstacle will be that very early interventions need some means of consent-exempt entry into rigorous clinical trials.

Medical knowledge and improvements in care can certainly be accomplished through the review of care provided to a single patient; the Kevin Everett case is one example. However, it is critical that the case be presented accurately and that all aspects of care be reviewed. This case as generally presented in the lay press sensationalizes one portion of the patient’s treatment (hypothermia) without careful review of its temporal relationship to the patient’s recovery, and ignores the other components of the patient’s care (including early diagnosis, immobilization, medications, and surgical intervention).

Medical stories presented in the media are typically incomplete and often contain inaccuracies; thus, the interpretations that follow become exponentially more inaccurate. The reported role of hypothermia in the recovery of Kevin Everett is an outstanding example. After reading this article, neurosurgeons may better appreciate the minor role this treatment played in his recovery in view of the nature of his injury and numerous other interventions. CNSQ
MEMBERSHIP NEWS

CNS Bylaws: Proposed Amendments

At the January 2008 meeting of the CNS Executive Committee (EC), two proposed CNS Bylaw amendments were discussed and approved for a vote by the membership. These proposals are to be reviewed by the CNS membership prior to voting at the CNS business meeting in Orlando in September 2008.

Proposed CNS Bylaws amendment #1:

Since the CNS no longer publishes Neurosurgery News and has replaced that publication with Congress Quarterly, the CNS Bylaws should reflect that change.

1. Proposed Amendment:

Article IV, Section 4; Article VII, Section 1(L); and Article VII, Section 9 shall be amended by replacing the words Neurosurgery News with Congress Quarterly.

Proposed CNS Bylaws amendment #2:

Since the Joint Committee on the Assessment of Quality has been dissolved and its responsibilities assumed by the Washington Committee (WC), Article VII, Section 1(T) should be eliminated from the CNS bylaws. Because the CNS bylaws already describe the appointment of individuals to the WC, no substitute language is required.

2. Proposed amendment:

Article VII, Section 1(T) shall be deleted.

CNS Membership: Applications in Progress

Manish Aghi
Hooman Azmi
Orhan Barlas
Sadok Ben Amor
Carl Bevering, III
Deidre Buckley
Geert Buyse
Ali Bydon
Satnam Chhabra
Carlos Codas
Cengiz Cokluk
Georgy Daneri
Amir Dehdashti
Rodrigo Diaz
Erin Dunbar
Sudesh Ebenezer
Daniel Fassett
Adriano Fernandes
Amanda Frakes
Alfredo Fuentes Davila
Ashok Hande
Jorge Jaramillo-de la Torre
Ryan Jewell
Christopher Koebbe
Yu-Hung Kuo
Herman Christopher Lawson
Dean Lin
Christer Lindquist
Darlene Lobel
Mitchell Maltenfort
Thomas Manning
Francisco Mery Munoz
Yzabel Michaud
Thomas Moriarty
Balamurugan Nadar
Malini Narayanan
Christopher Neumann
Misao Nishikawa
Juan Ortega-Barnett
Sun Paek
Michael Petr
Kevin Petrecca
Julie Pilitsis
Warren Roberts
Bello Shehu
Jun Shin
Prakash Singh
Michael Louis Smith
Michael Stanley
Sandeep Teja
Leonide Toussaint, III
Pablo Vasconez
Frederick Vincent
Luis Vintimilla
Bradley Wallace
Ziv Williams
A representative confocal microscopy image of the rodent spinal cord after transfection with adeno-associated virus (AAV) is shown above. This axial slice shows successful transfection of glial cells, which are green due to expression of green fluorescence protein (GFP). Neurons labeled with a neuron-specific marker (anti-neuronal nuclei) typically appear red, although a few appear yellow due to simultaneous transfection with AAV.