Methodology of the Guidelines for the Management of Acute Cervical Spine and Spinal Cord Injuries

Beverly C. Walters, MD, MSc, FRCSC

Division of Neurological Surgery, University of Alabama at Birmingham, Birmingham, Alabama, and Department of Neurosciences, Inova Health System, Falls Church, Virginia

Correspondence:
Beverly C. Walters, MD, MSc, FRCSC, FACS,
UAB Division of Neurological Surgery,
510 - 20th St S, FOT 1008,
Birmingham, AL, 35294-3410.
E-mail: bcwmd@uab.edu

KEY WORDS: Guidelines methodology, Practice guideline methodology, Practice parameter development

Neurosurgery 72:17–21, 2013 DOI: 10.1227/NEU.0b013e318276ed9a www.neurosurgery-online.com

The original publication of guidelines dealing with the care of acutely injured cervical spine and spinal cord patients occurred within a few years of organized neurosurgery embracing the concept of producing evidence-based recommendations.[1] Much has been learned as a result of the careful critical evaluation of the medical literature pertaining to neurosurgical patients, and the methodology used in formulating practice parameters (or recommendations) has undergone further change and development. The methodology used in this iteration of these recommendations is a product of contributions from many sources, including multiple guidelines produced by neurosurgery and other specialty organizations.

BACKGROUND OF METHODOLOGY FOR EVIDENCE-BASED RECOMMENDATIONS

In the 1990s, professional and governmental organizations such as the American Medical Association, the American College of Physicians, the Institute of Medicine, the former Agency for Health Care Policy and Research, and others recognized that clinical decision making had to be founded in scientific discovery and that the best clinical trials provided the best evidence for treatment. While not ignoring clinical experience, they recognized that expert opinion could result from conformity in practice as easily as from science, noting the superiority of the latter. In 1993, under the leadership of the American Association of Neurological Surgeons (AANS), a move was made away from consensus-based, potentially biased, "review criteria" to use of a more formalized system to classify and grade extant medical literature for creating recommended clinical maneuvers. This was heavily influenced by the American Academy of Neurology's approach to the same task, explained and proposed by Rosenberg and Greenberg.[2] In this paradigm, the recommendations regarding the overall process include the following, modified from recommendations by the Institute of Medicine [3]:

- Practice parameters should be developed in conjunction with physician organizations.
- Reliable methodologies that integrate relevant research findings and appropriate clinical expertise should be used to develop practice parameters.
- Practice parameters should be as comprehensive and specific as possible.
- Practice parameters should be based on current information.
- Practice parameters should be widely disseminated.

In the spirit of compliance with the above recommendations, the AANS, later joined by the Congress of Neurological Surgeons (CNS), moved forward with embracing the production of evidence-based practice recommendations, or parameters, under the generic rubric of "guidelines." Topics were chosen for their controversies, their weight in terms of burden of illness to society, and the (sometimes wide) variability in practice across the country. These topics were comprehensively specific, covering treatment, prognosis, clinical assessment, diagnosis, and, more recently, economic analysis and clinical decision-making. Recommendations were based on the most current information available and used careful methodology that focused on the quality of a given study's design, awarding more

ABBREVIATIONS: AANS, Association of Neurological Surgeons; **CNS,** Congress of Neurological Surgeons

TABLE 1. Classification of Evidence and Subsequent Recommendations by Woolf[16] and Based on Canadian Recommendations[15]

Canadian Task Force Classification of Recommendations and Study Designs: 1992 (Modified)	
Category	**Description**
Recommendations	
A	There is good evidence to support the recommendation
B	There is fair evidence to support the recommendation
C	There is poor evidence to support use, but recommendations can be made on other grounds
D	There is fair evidence to support the recommendation that the treatment NOT be used
E	There is good evidence to support the recommendation that the treatment NOT be used
Study design	
I	Evidence obtained from at least 1 properly designed randomized controlled trial
II-1	Evidence obtained from well-designed controlled trials without randomization
II-2	Evidence obtained from well-designed cohort or case-control analytic studies, preferably from > 1 center or research group
II-3	Evidence obtained from comparisons between times or places with or without intervention; dramatic results from uncontrolled experiments (such as the results of treatment with penicillin in the 1940s) could also be included in this category
III	Opinions of respected authorities, based on clinical experience, descriptive studies (such as case series), or reports of expert committees

weight to those studies with the least methodological flaws. The development of this methodology has its roots in critical evaluation of the medical literature and weighting that evidence in a way that the robustness of the evidence supporting the recommendations could be inferred by the nomenclature used in classifying the recommendations. The original and subsequent early guidelines produced in neurosurgery used a 3-tier classification system in which literature was qualified as **Class I**, **Class II**, and **Class III** medical evidence.[4] This reflected a decreasing certainty in the appropriateness of the conclusions of the literature and therefore the strength of the recommendations. In this classification, quality of evidence was indicated as follows:

- **Class I:** Evidence from 1 or more well-designed randomized controlled clinical trials, including overviews of such trials.
- **Class II:** Evidence from 1 or more well-designed comparative clinical studies, such as nonrandomized cohort studies, case-control studies, and other comparable studies, including less well-designed randomized controlled trials.
- **Class III:** Evidence from case series, comparative studies with historical controls, case reports, and expert opinion, as well as significantly flawed randomized controlled trials.

It is this designation that was used in the previous iteration of these guidelines, as well as several other neurosurgical guideline documents.[5-12] The levels of recommendations as used in the previous iteration of the "Guidelines for the Management of Acute Cervical Spine and Spinal Cord Injuries," which are derived from the classes of evidence listed above, are related to the certainty that a clinician has that the evidence is strong enough to support the recommendation(s) as follows:

- **Standards:** Reflection of a high degree of clinical certainty
- **Guidelines:** Reflection of a moderate degree of clinical certainty
- **Options**: Reflection of unclear clinical certainty

TABLE 2. Integrating Evidence Quality Appraisal With an Assessment of the Anticipated Balance Between Benefits and Harms if a Policy Is Carried Out Leads to Designation of a Policy as a Strong Recommendation, Recommendation, Option, or no Recommendation

AMERICAN ACADEMY OF PEDIATRICS CLASSIFYING RECOMMENDATIONS FOR CLINICAL PRACTICE GUIDELINES[3]		
Evidence Quality	**Preponderance of Benefit or Harm**	**Balance of Benefit and Harm**
A. Well-designed, randomized controlled trials or diagnostic studies on relevant populations	Strong	Option
B. Randomized controlled trials or diagnostic studies with minor limitations; overwhelmingly consistent evidence from observational studies	Recommendation	
C. Observational studies (case-control or cohort design)	Recommendation	
D. Expert opinion, case reports, reasoning from first principles	Option	No Recommendation
X. Exceptional situation in which validating studies cannot be performed and there is a clear preponderance of benefit or harm	Strong	
	Recommendation	
	Recommendation	

This 3-tiered system was originally suggested by David Eddy in 1990,[13] embraced by the American Academy of Neurology,[2] and subsequently became policy for the AANS/CNS Guidelines Committee.[14] However, during the years that professional societies have taken on the responsibility for practice recommendation development for their specialties, many different paradigms with different nomenclature have been suggested and used for guideline designation. For example, the Canadian Task Force on the Periodic Examination proposed a more extensive schema very early on in the history of guideline development.[15] This 5-tiered system was modified and further proposed by Steven Woolf, from the Office of Disease Prevention and Health Promotion in the Public Health Service in Washington, DC (Table 1).[16] A similar 5-tiered system is being used by the North American Spine Society.[17] The American Academy of Pediatrics proposed a 4-tiered system, shown in Table 2.[18] More simply, the American Thoracic Society generated recommendations for a 2-tiered system that indicated whether recommendations were "strong" or "weak."[19] In this context, the neurosurgical 3-tiered system appears to be appropriate and easy to implement and understand.

As noted in Table 2, the pediatricians added an "X" category for those situations when there is clear evidence that some action should (or should not) be taken, and no formal comparative study could (or should) be done. An example of an "X" category in neurosurgery would be a circumstance in which a patient with an acute intracranial epidural hematoma demonstrates a unilateral dilated pupil; no one would suggest randomizing that patient to a "no treatment" arm of a randomized controlled trial. The most that one could obtain is a case-control study of a population of patients, some of whom received treatment and some of whom did not (for whatever reason—delay in transport, unavailability of a neurosurgeon, failure of diagnosis, etc). This would provide Class II medical evidence but has never been carried out. This was the very struggle faced by the author group of the Guidelines for the Surgical Management of Traumatic Brain Injury.[10] In that publication, the group wrestled with the paucity of evidence that could make evacuation of an intracranial epidural hematoma in the scenario of impending brainstem compression a practice Standard, instead of being required to relegate it to the category of practice Option. Because this categorization would be completely inappropriate, the group decided to abandon the Standards, Guidelines, and Options nomenclature that had been adhered to for many years (including the previous iteration of these guidelines), switching instead to the categories of **Level I** (for Standards), **Level II** (for Guidelines) and **Level III** (for Options), with the same classes of evidence that had been used previously in the old nomenclature. It is this categorization that is being used in this iteration of the "Guidelines for the Management of Acute Cervical Spine and Spinal Cord Injuries," continuing with the 3-tier system that has always been used in neurological surgery, and continues to be consistent with the policy of the AANS/CNS.[12] There is an underlying belief in the simplicity of this system compared with one that may be more sensitive to nuances in quality of medical publications but more unwieldy in the context of evaluation of the neurosurgical literature (Table 3).[20]

GUIDELINES METHODOLOGY USED IN THE CURRENT VERSION OF THESE RECOMMENDATIONS

For the purposes of these guidelines, the author group has chosen to use a modification of the North American Spine Society criteria for evaluation of the medical literature[17] (Table 4). There are significant differences in the original North American Spine Society criteria and those being used in these guidelines. The obvious, and most notable, difference is that there are 3 classes of evidence, as described above, consistent with other neurosurgical guidelines, for

TABLE 3. Detailed Levels of Evidence Recommended by Oxford[4]

Oxford Center for Evidence-Based Medicine Levels of Evidence	
1a	Systematic reviews (with homogeneity) of randomized controlled trials Systematic review of randomized trials displaying worrisome heterogeneity
1b	Individual randomized controlled trials (with narrow confidence interval) Individual randomized controlled trials (with a wide confidence interval)
1c	All or none randomized controlled trials
2a	Systematic reviews (with homogeneity) of cohort studies Systematic reviews of cohort studies displaying worrisome heterogeneity
2b	Individual cohort study or low-quality randomized controlled trials (< 80% follow-up) Individual cohort study or low-quality randomized controlled trials (< 80% follow-up/wide confidence interval)
2c	"Outcomes" research; ecological studies
3a	Systematic review (with homogeneity) of case-control studies Systematic review of case-control studies with worrisome heterogeneity
3b	Individual case-control study
4	Case series (and poor-quality cohort and case-control studies)
5	Expert opinion without explicit critical appraisal or based on physiology, bench research, or "first principles"

TABLE 4. Rating Scheme for the Strength of the Evidence: Modified North American Spine Society Schema[5] to Conform to Neurosurgical Criteria as Previously Published and for Ease of Understanding and Implementation: Levels of Evidence for Primary Research Question[a]

Class	Therapeutic Studies: Investigating the Results of Treatment	Diagnostic Studies: Investigating a Diagnostic Test	Clinical Assessment: Studies of Reliability and Validity of Observations, Including Clinical Examination, Imaging Results, and Classifications
I	High-quality randomized controlled trial with statistically significant difference or no statistically significant difference but narrow confidence intervals	Testing of previously developed diagnostic criteria on consecutive patients (with universally applied reference "gold" standard)	Evidence provided by 1 or more well-designed clinical studies in which interobserver and intraobserver reliability is represented by a κ statistic ≥ 0.60 or an intraclass correlation coefficient of ≥ 0.70
	Systematic review[b] of Class I randomized controlled trials (and study results were homogeneous[c])	Systematic review[b] of Class I studies	
II	Lesser-quality randomized controlled trial (eg, < 80% follow-up, no blinding, or improper randomization)	Development of diagnostic criteria on consecutive patients (with universally applied reference "gold" standard)	Evidence provided by 1 or more well-designed clinical studies in which interobserver and intraobserver reliability is represented by a κ statistic of 0.40-0.60 or an intraclass correlation coefficient of 0.50-0.70
	Prospective[d] comparative study[e]	Systematic review[b] of Class II studies	
	Systematic review[b] of Class II studies or Class I studies with inconsistent results	Study of nonconsecutive patients; without consistently applied reference "gold" standard	
	Case-control study[g]	Systematic review[b] of Class III studies	
	Retrospective[f] comparative study[e]	Case-control study	
	Systematic review[b] of Class II studies		
III	Case series[h]	Poor reference standard	Evidence provided by 1 or more well-designed clinical studies in which interobserver and intraobserver reliability is represented by a κ statistic of < 0.40 or an intraclass correlation coefficient of < 0.50.
	Expert opinion	Expert opinion	

[a]A complete assessment of quality of individual studies requires critical appraisal of all aspects of the study design.
[b]A combination of results from 2 or more prior studies.
[c]Studies provided consistent results.
[d]Study was started before the first patient enrolled.
[e]Patients treated 1 way (eg, halo vest orthosis) compared with a group of patients treated in another way (eg, internal fixation) at the same institution.
[f]The study was started after the first patient was enrolled.
[g]Patients identified for the study on the basis of their outcome, called "cases" (eg, failed fusion), are compared with those who did not have outcome, called "controls" (eg, successful fusion).
[h]Patients treated 1 way with no comparison group of patients treated in another way.

ease of understanding. Second, the case-control study has been retained with its Class II designation, as in previous guidelines and because they are, in most other suggested guideline schema (eg, Tables 1-3), differentiated from case series, case reports, and expert opinion. The reasons are that there is a comparison group and that case-control studies can be strengthened with robust study design to be comparable to nonrandomized cohort studies.[21]

Attention must be brought to the fact that all of the above discussion has focused on therapeutic effectiveness as studied in randomized controlled trials, comparative cohort studies, case-control studies, and case series. However, these guidelines also discuss diagnostic tests, which are largely imaging studies of various kinds, as well as clinical assessment such as neurological scoring and classification of injury. Investigation of these aspects of clinical study does not use the same trial designs as those used for therapeutic effectiveness. The methodology for evaluating these 2 types of studies are outlined in detail in the last iteration of these guidelines,[1] including rationale for the rating schema, and remain relatively unchanged. However, because there have been new and improved studies in clinical assessment and because these studies use different statistical evaluations, the older recommended criteria have been broadened, as depicted in Table 4.

PROCESS FOR GUIDELINES DEVELOPMENT

The current author group was selected for its expertise in spinal surgery (both neurosurgical and orthopedic), neurotrauma,

clinical epidemiology, and, in several cases, prior experience with guideline development. The topics chosen for inclusion in this iteration of these guidelines are contemporary and pertinent to the assessment, evaluation, care, and treatment of patients with acute cervical spine and/or spinal cord injuries. A National Library of Medicine (PubMed) search of the literature published from 1966 through 2011 was accomplished using the search terms defined in each guideline manuscript. The search was limited to human subjects and included English language literature for all but one of the chapters. Additional articles were found through the reference lists in the articles found, as well as from other sources known to the authors. Articles were rejected on the basis of irrelevance to the clinical question at hand. Case reports were included if there was insufficient material from case series. On occasion, the assessed quality of the study design was so contentious and the conclusions so uncertain that we assigned a lower medical evidence classification than might have been expected without such a detailed review. In every way, adherence to the Institute of Medicine's criteria for searching, assembling, evaluating, and weighing the available medical evidence and linking it to the strength of the recommendations presented in this document was carried out.

Selected articles were carefully reviewed by the authors. Evidentiary tables were created that reflected the strengths and weaknesses of each article. Articles that did not achieve immediate consensus among the author group were discussed extensively until a consensus was reached. Very few contributions required extensive discussion. Most articles were easily designated as containing Class I, II, or III medical evidence using the criteria set forth by the author group at the initiation of the literature evaluation process as described above.

SUMMARY

Efforts on the part of neurosurgical specialty societies to remain involved and active in the development of practice recommendations are commendable and completely necessary. The development of practice policies for control of healthcare costs is not a new movement and has been gaining momentum over the last 20 years. In 1990, David Eddy described the (then) recent changes in the view and use of practice policies[22]:

> . . .Practice policies now are being designed explicitly as instruments for quality assurance, pre-certification, utilization review, accreditation, coverage, and cost containment.. . .But the greatest concern pertains to control. It is not stretching things too far to say that whoever controls practice policies controls medicine. That control used to lie exclusively, if diffusely, within the medical profession. However, as policies are designed and used as management tools, control could shift outside the profession.. . . As non-physician organizations develop policies to use as management tools, physician groups must race to develop their own policies, lest they lose control.

Spinal surgeons, including neurosurgeons and orthopedic surgeons, must continue to wrestle with defining the best treatment possible for their patients. This includes generating the best possible evidence to support treatment paradigms and summarizing that evidence periodically in evidence-based recommendations. This publication is an example of the latter; it remains an unavoidable responsibility of spinal specialists to pursue the former.

Disclosure

The author has no personal financial or institutional interest in any of the drugs, materials, or devices described in this article.

REFERENCES

1. Hadley M, Walters B, Grabb P, et al. Guidelines for the management of acute cervical spine and spinal cord injuries. *Clin Neurosurg.* 2002;49:407-498.
2. Rosenberg J, Greenberg MK. Practice parameters: strategies for survival into the nineties. *Neurology*. 1992;42(5):1110-1115.
3. Field M, Lohr K. *Clinical Practice Guidelines: Directions for a New Program—Committee to Advise the Public Health Service on Clinical Practice Guidelines: Institute of Medicine*. Washington, DC: National Academy Press; 1990.
4. Bullock R, Chesnut RM, Clifton G, et al. Guidelines for the management of severe head injury: Brain Trauma Foundation. *Eur J Emerg Med.* 1996;3(2):109-127.
5. Shaw E, Bernstein M, Recht L, et al. Practice parameters in adults with suspected or known supratentorial non-optic pathway low-grade glioma. *Neurosurg Focus.* 1998;4(6):e10.
6. Bullock R, Chesnut R, Clifton G, et al. Guidelines for the management of severe traumatic brain injury. *J Neurotrauma.* 2000;17(6/7):100.
7. Aarabi B, Alden T, Chesnut R, et al. Guidelines for the management of penetrating brain injury. *J Trauma.* 2000;51(2):82.
8. Gabriel E, Ghajar J, Jagoda A, Pons P, Scalea T, Walters B. Guidelines for prehospital management of traumatic brain injury: methodology: guideline development rationale and process. *J Neurotrauma.* 2002;19(1):3.
9. Resnick DK, Choudhri TF, Dailey AT, et al. Guidelines for the performance of fusion procedures for degenerative disease of the lumbar spine, part 1: introduction and methodology. *J Neurosurg Spine.* 2005;2(6):637-638.
10. Bullock MR, Chesnut R, Ghajar J, et al. Surgical management of severe traumatic brain injury: surgical management of acute epidural hematomas. *Neurosurgery*. 2006;58(3 suppl):10.
11. Warden D, Gordon B, McAllister T, et al. Guidelines for the pharmacologic treatment of neurobehavioral sequelae of traumatic brain injury. *J Neurotrauma.* 2006;23(10):33.
12. Robinson PD, Kalkanis SN, Linskey ME, Santaguida PL. Methodology used to develop the AANS/CNS management of brain metastases evidence-based clinical practice parameter guidelines. *J Neurooncology.* 2010;96(1):6.
13. Eddy DM. Clinical decision making: from theory to practice: designing a practice policy: standards, guidelines, and options. *JAMA.* 1990;263(22):3077, 3081, 3084.
14. Walters BC. Clinical practice parameter development in neurosurgery. In: Bean JR, ed. *Neurosurgery in Transition: The Socioeconomic Transformation of Neurological Surgery*. Baltimore, MD: Williams & Wilkins; 1998:99-111.
15. Canadian Task Force on the Periodic Health Examination. The periodic health examination. *Can Med Assoc J.* 1979;121:1193-1254.
16. Woolf SH. Practice guidelines, a new reality in medicine, II: methods of developing guidelines. *Arch Intern Med.* 1992;152(5):946-952.
17. North American Spine Society. Diagnosis and treatment of degenerative lumbar spinal stenosis: rating scheme for the strength of the evidence. http://www.guideline.gov/content.aspx?id=11306#Section405.
18. American Academy of Pediatrics Steering Committee on Quality Improvement and Management. Classifying recommendations for clinical practice guidelines. *Pediatrics*. 2004;114(3):874-877.
19. Guyatt GH, Oxman AD, Vist GE, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ*. 2008;336(7650):924-926.
20. Oxford Center for Evidence-Based Medicine. Oxford Centre for Evidence-Based Medicine levels of evidence. 2001. www.cebm.net/index.aspx?o=1025.
21. Hu X, Wright JG, McLeod RS, Lossing A, Walters BC. Observational studies as alternatives to randomized clinical trials in surgical clinical research. *Surgery.* 1996;119(4):3.
22. Eddy DM. Clinical decision making: from theory to practice: practice policies—what are they? *JAMA.* 1990;263(6):877-878, 880.