FEATURED ARTICLE

COMPARATIVE EFFECTIVENESS RESEARCH – WHAT DOES IT MEAN FOR YOU AND YOUR PRACTICE?

State of Neurosurgical Evidence
Most published clinical studies in neurosurgery are retrospective. Prospective studies including randomized trials are expensive and difficult to execute. Even when RCTs are performed, they do not necessarily provide definitive evidence that guides clinical practice. Organized neurosurgery, through the AANS/CNS Guidelines Committee, has emphasized the need for evidence-based practice by endorsing evidence-based guidelines on a variety of topics. Despite this initiative, many of the guidelines have demonstrated that there is little evidence to guide practice on major practical matters such as whether an unruptured aneurysm should be clipped or coiled, or whether a lumbar spinal deformity with nerve compression should be decompressed alone or decompressed with instrumented fusion. Many ask whether acoustic neuroma should be treated with surgery vs. stereotactic radiosurgery. Often, finding the answer requires careful long-term follow-up. Large administrative claims databases in their current form are unable to answer these questions because they lack the necessary patient identifiers to provide risk-adjusted outcome assessment and they lack sufficient follow-up outcome data to make meaningful conclusions.

Are More RCTs the Answer?
The RCT remains the gold standard for evaluating the clinical benefit of any particular intervention primarily because it reduces bias from known or unknown confounders. RCTs, for example, have established the clear benefit of carotid endarterectomy over medical treatment for hemodynamically significant carotid stenosis. Traditional RCTs might be unlikely to provide answers for many current questions that involve neurosurgeons because the comparative treatment options are not always testing a medical hypothesis. Rather, they are comparing two valid treatment options whose outcome depends more upon patient-specific factors and technical factors that can be highly variable. Randomized studies generally require homogeneous clinical populations that are often a very select group, and therefore results from these carefully defined populations might limit generalizability. Patient crossover has also been shown to limit the power of the RCT, as shown in the SPORT trials where up to 40% of patients did not comply with their randomized assignment. In the SPORT trials, the treatment groups as treated were not comparable at baseline. Most clinicians feel that SPORT did not provide data to guide clinical practice. Perhaps we should ask if CER is the answer for evaluating neurosurgical problems where patient heterogeneity is the norm.

Modern Prospective Registries and Comparative Effectiveness
CER in contrast to the traditional RCT aims to assess the actual effectiveness of a medical test or intervention. That is, by design, it includes all or most patients with a particular condition to be studied, and it therefore might improve generalizability. The heterogeneity among patient populations might limit the power of CER studies to detect meaningful differences, although CER studies that show non-inferiority might be equally valuable as those that show superiority—again, because the results represent actual practice and therefore actual effectiveness.

Advanced Technology and Changing Practice Patterns
What might CER look like for a neurosurgical problem and how might it work? Let’s consider ruptured intracranial aneurysms (Figure 1). The International subarachnoid aneurysm trial (ISAT) randomized 2,143 patients at 42 sites in an effort to compare microsurgical clipping to endovascular coiling for the treatment of ruptured intracranial aneurysms. Endovascular coiling was more likely to result in independent survival at one year compared to microsurgical clipping, although the risk of late re-bleeding was higher in the endovascular cohort. As endovascular techniques continue to advance, the re-bleeding risk might reduce. As practice patterns change, the population of patients treated with microsurgical clipping might change and as a result the complication rate may also change. Further RCTs comparing clipping and coiling are...
unlikely to be performed for these and other reasons. Comparative effectiveness studies might provide useful data if performed correctly. As practice patterns change, outcomes assessments including return to work status, survival and re-bleeding rates might provide meaningful status reports for neurosurgeons. Prospective multicenter database studies might provide useful information comparing the effectiveness of coils versus clipping, but also might provide data on the overall outcomes of different sites that may practice differently, with differential utilization of clipping versus coiling of aneurysms. Doing this successfully will require organization, infrastructure, risk adjustment and modern statistical methodology to adjust for confounders. The Obama administration appears to recognize this and has put in place many grant mechanisms to build such an infrastructure that will permit scalable outcomes assessment from the majority of practitioners.

Clinical Heterogeneity and Patient-Reported Outcomes

Let’s now consider a spinal disorder. In any condition where there is significant heterogeneity, it is critical to define a population in which there is clinical equipoise between two or more interventions before conducting a comparative study. Cervical spondylotic myelopathy was designated by the IOM as among the top research priorities for CER. For over 50 years, debate has raged over the approach (Figure 2). There are several options, ventral decompression and fusion, dorsal decompression with fusion, and dorsal decompression without fusion (laminectomy alone or laminoplasty). The heterogeneity of the clinical population has complicated the interpretation of study results in the past including the recent prospective AO Foundation studies. Recent efforts have defined an appropriate and significant population where equipoise between ventral and dorsal approaches appear to exist. This equipoise population now serves as the basis for a comparative effectiveness trial being considered by NIAMS to compare ventral fusion to dorsal fusion. Since patient-reported quality of life outcomes, re-operation rates and complications including CS palsy and pseudoarthrosis will all play a role in defining overall success in these patient populations, a careful study with dedicated study coordinators will be required to ensure the completeness and integrity of the data. A simple claims-based registry would be unlikely to provide answers.

Implications of Comparative Effectiveness Research

One of the major differences in CER compared with other clinical research is that many efforts in this realm will need to involve many stakeholders and will not necessarily be purely investigator-initiated. This means that the agenda of those who design CER studies might be different from physician-scientists. CER could be designed to limit access to expensive health care, CER might be designed to promote changes in healthcare policy.

Although the US government does not plan to make payment decisions based on CER alone, no one can deny the enormous economic pressures to reduce health care costs. If CER demonstrated that two medical interventions were similar in terms of safety and effectiveness but that one treatment was ten times the cost of another, it would be difficult to justify paying for the more expensive treatment as a general rule.

Neurosurgeons cannot afford to sit on the sidelines waiting for CER to define what should be done for back pain, for cervical spondylotic myelopathy, or for cerebrovascular disease. As researchers, we must lead the process for evaluating, treating, and monitoring the conditions that we treat. As neurosurgeons, we must organize ourselves in order to make certain that policy makers and purchasers of health care become familiar with our unique perspectives on which procedure is best for which patients and why.

References: