

WJO 5<sup>th</sup> Anniversary Special Issues (8): Spine**Research in spinal surgery: Evaluation and practice of evidence-based medicine**

Mark E Oppenlander, Christopher M Maulucci, George M Ghobrial, James S Harrop

Mark E Oppenlander, Christopher M Maulucci, George M Ghobrial, James S Harrop, Department of Neurological Surgery, Thomas Jefferson University Hospital, Philadelphia, PA 19107, United States

Mark E Oppenlander, Division of Neurological Surgery, Barrow Neurological Institute, St. Joseph Hospital and Medical Center, Phoenix, AZ 85003, United States

Author contributions: Oppenlander ME designed the research, contributed to data review, drafted the paper; Maulucci CM and Ghobrial GM contributed to data review and drafting of paper; Harrop JS was the consultant supervisor, contributed to review and approval of paper.

Correspondence to: James S Harrop, MD, Professor of Neurological Surgery, Department of Neurological Surgery, Thomas Jefferson University, 909 Walnut Street Second Floor, Philadelphia, PA 19107, United States. [james.harrop@jefferson.edu](mailto:james.harrop@jefferson.edu)  
Telephone: +1-215-9557959 Fax: +1-215-5037007

Received: September 29, 2013 Revised: December 19, 2013

Accepted: January 13, 2014

Published online: April 18, 2014

**Abstract**

Evidence-based medicine (EBM) is a common concept among medical practitioners, yet unique challenges arise when EBM is applied to spinal surgery. Due to the relative rarity of certain spinal disorders, and a lack of management equipoise, randomized controlled trials may be difficult to execute. Despite this, responsibility rests with spinal surgeons to design high quality studies in order to justify certain treatment modalities. The authors therefore review the tenets of implementing evidence-based research, through the lens of spinal disorders. The process of EBM begins with asking the correct question. An appropriate study is then designed based on the research question. Understanding study designs allows the spinal surgeon to assess the level of evidence provided. Validated outcome measurements allow clinicians to communicate the success of treatment strategies, and will increase the quality of a given study design. Importantly,

one must recognize that the randomized controlled trial is not always the optimal study design for a given research question. Rather, prospective observational cohort studies may be more appropriate in certain circumstances, and would provide superior generalizability. Despite the challenges involved with EBM, it is the future of medicine. These issues surrounding EBM are important for spinal surgeons, as well as health policy makers and editorial boards, to have familiarity.

© 2014 Baishideng Publishing Group Co., Limited. All rights reserved.

**Key words:** Evidence-based medicine; Spinal surgery; Trial design; Research; Methodology

**Core tip:** This paper highlights the intricacies of spinal research. The difficulties of conducting high quality research in spinal surgery are discussed, but the tools for success are outlined. Specifically, the tenets of implementing evidence-based research are provided, along with a discussion of validated outcome measures which will increase the quality of a given study design. Importantly, the randomized controlled trial should not always be considered the best study design for a given research question, and observational cohort studies may be more appropriate in certain circumstances. Ultimately, spinal surgeons are responsible for evidence-based research to justify treatment paradigms.

Oppenlander ME, Maulucci CM, Ghobrial GM, Harrop JS. Research in spinal surgery: Evaluation and practice of evidence-based medicine. *World J Orthop* 2014; 5(2): 89-93 Available from: URL: <http://www.wjgnet.com/2218-5836/full/v5/i2/89.htm>  
DOI: <http://dx.doi.org/10.5312/wjo.v5.i2.89>

**INTRODUCTION**

The concept of evidence-based medicine (EBM) assumes

that current medical research, along with individual clinician judgment, can optimally guide clinical decision making to result in the best possible patient outcomes<sup>[1,2]</sup>. While EBM requires the use of best available evidence, multiple challenges may arise in its practical application to spinal surgery. For instance, rare disorders result in small patient numbers and subsequent lower quality data. At the other extreme, randomized controlled trials (RCTs) attempt to generate high quality evidence, yet are hindered by expense and difficulties in study recruitment and conduct.

Despite the challenges involved with EBM, it is the future of medicine<sup>[3,4]</sup>. If spine surgeons do not want poor-quality studies to dictate and limit their clinical decisions, then responsibility rests with this group of practitioners to design high quality studies to justify certain treatment modalities.

This review therefore highlights the tenets of implementing evidence-based research, through the lens of spinal disorders. Techniques of conducting and evaluating EBM are first discussed, followed by a review of pertinent outcome measures in spinal surgery. It is the authors' goal that these basic tools will provide a basis of EBM for the practicing spinal surgeon.

## TECHNIQUES OF EBM

### **Asking the correct question**

Before a study is designed, a research question must be asked. The importance of the research question lies in the fact that it dictates a study's design. Often, the RCT is considered the gold standard of evidence-based medicine, yet the research question may exclude the RCT from feasibility or utility. For example, a question of superiority in treatment protocols, where each treatment has equipoise for the surgeon and patient, is suitable for a RCT. However, a question may best be answered with a prospective cohort design if there are subjective treatment preferences among surgeons, the presence of significant selection biases, or poor generalizability<sup>[3,5]</sup>.

A well-designed and focused research question will not only dictate the study design, but will also aid in literature searches. Instead of turning up hundreds to thousands of citations, a well-defined question will limit the pertinent literature to a manageable number that allows a focused interpretation.

In addition, the research question will permeate through a central theme of the research manuscript. The question should be stated in the introduction of a manuscript, and contain the intervention of study and cohort of interest. Returning to the research question throughout the manuscript will allow reporting of more concise results and a more pertinent discussion.

### **Designing the study**

Once a research question is proposed and a formal review of the literature is performed, the study design is implemented. Table 1 highlights the advantages and disadvantages of various study designs. A case series tracks

patients with a known pathology given similar treatment, and allows assessment of a clinical course based on that treatment. Case series are often retrospective but may be prospective. They are often confounded by selection bias, limiting elucidations of causal relationships. Case series may be improved, however, with well-defined selection criteria and the use of validated outcome measures. For rare spinal disorders, a case series may be the best available evidence.

A case-control study design is a type of observational study, wherein two patient groups with differing outcomes are identified and compared for a supposed causal attribute. They are retrospective in nature and relatively inexpensive. Because of their retrospective nature, however, there is difficulty in obtaining reliable information about a patient's exposure over time. This effectively hinders the ability to make claims of causation.

A cohort study is also observational in nature. It follows a group of patients without a disease in order to determine risks of contracting that condition, or compares two treatment options. A cohort study may be retrospective or prospective in nature. It is beneficial for identifying the natural history of a disease, the risk factors of a disease, or the impact of an intervention. Cohort studies are more expensive and time consuming than a case-control design or a case series, with strict inclusion and exclusion criteria. Prospective cohort studies in particular are considered to yield the most reliable results in observational studies.

A RCT is considered the gold standard for a clinical trial. It is often used to test the effectiveness of a medical or surgical intervention within a well-defined patient population. The intervention is provided to the patient based on a process of randomization, in a blinded or unblinded manner. The RCT offers reliable evidence because it reduces bias and spurious causality. Nonetheless, RCTs are prone to high cost, administrative difficulties, and limited recruitment.

Spine surgeons in particular may struggle with obtaining a high quality RCT because of high crossover rates and low patient recruitment. In addition, it is relatively difficult for surgeons to design a study that randomizes patients to interventions that are typically used in sequence<sup>[5]</sup>. For example, a RCT may be designed to compare operative *vs* non-operative treatment of neck pain. Most surgeons, however, consider failure of non-operative measures as an indication for surgery. Therefore, a patient in this study would need to accept being randomized to a non-operative treatment modality that s/he has already failed. If the concept of equipoise is used in an attempt to circumvent this problem, then surgeons may be relegated to operating on patients without clear operative indications. In this example, a supposed state of equipoise could lead to a surgeon operating on a patient with neck pain who has not failed conservative measures, further confounding results and perhaps leading to poorer surgical outcomes.

Because practical and ethical reasons may prevent the initiation of a RCT, strong observational alternatives are

Table 1 Types of study design

Design type	Advantages	Disadvantages
Case series	Suitable for rare diseases or new treatments	No comparison group Retrospective nature
Case control	Small sample size Short duration	Presence of confounding Retrospective nature
Cohort studies	Evaluates risk factors Compares two treatments May be prospective	Presence of confounding
Randomized controlled trials	Prospective in nature Reduce confounding and bias	Limited generalizability Potential for low recruitment and high crossover High cost and administrative oversight
Systematic review	Provides summation of available literature	Dependent on quality of individual studies

Adapted from Fisher *et al*<sup>[5]</sup>.

Table 2 Levels of evidence

Evidence level	Therapeutic studies: Evaluating results of treatment	Prognostic studies: Evaluating outcome of disease
I	RCT Systematic review of level I RCTs	Prospective study (> 80% follow-up) Systematic review of level I studies
II	Prospective cohort study Poor quality RCT ( <i>e.g.</i> , < 80% follow-up) Systematic review of level II studies	Retrospective study Systematic review of level II studies
III	Case control study Retrospective cohort study Systematic review of level III studies	
IV	Case series	Case series
V	Expert opinion	Expert opinion

Adapted from Wright *et al*<sup>[8]</sup>. RCT: Randomized controlled trial.

needed in spinal research<sup>[2]</sup>. This notion would circumvent the impossibility of randomizing every component of intervention.

Systematic, evidence-based literature reviews provide a summation of the available literature on a topic. This type of study is valuable as a synopsis of previously-reported data, aiding understanding of outcomes, safety, risk factors, and impact of spinal surgery intervention<sup>[6]</sup>. Systemic reviews should be transparent, so that data is presented in an unbiased manner, thus allowing the surgeon to make independent conclusions based on the data. A quantitative synthesis of high quality data is termed a meta-analysis, which may be useful when pooling studies which are individually under-powered to find conclusive results<sup>[7]</sup>.

### Assessing the level of evidence

Based on study design, the level of evidence for an intervention can be assessed (Table 2)<sup>[8]</sup>. RCTs are categorized as level I or II. Cohort studies are level II or III. Case-control studies are level III, and case series are level IV. Expert opinion is considered level V. The level of evidence correlates with certainty of risks and benefits of a given intervention, so that higher levels of evidence (and thus higher quality studies) provide more certainty in their conclusions, and therefore stronger recommendations for treatment.

Although a majority of studies in spinal surgery are of levels III and IV, a select number of studies are of higher level evidence<sup>[9,10]</sup>. Certainly the level of evidence, however, is not the final answer in evaluating the literature. The lack of RCTs in spinal surgery research reflects the complexities and limitations of this study design. In addition, the current system of analyzing levels of evidence ignores whether the study asked the correct question or examined the relevant patient population<sup>[3]</sup>.

## OUTCOME MEASUREMENTS IN SPINE RESEARCH

The use of standardized outcome measurements is important for conducting evidence-based research. The quantification of patient symptoms, ability to perform activities of daily living, and overall health status is necessary to track patient progress as well as to conduct clinical studies. Outcome measurement tools allow clinicians to communicate, in a standardized manner, the success of treatment strategies. However, there is no standardized set of clinical outcome measures for all spine patients<sup>[11]</sup>. Outcome measurements for those with cervical pathology differ from those with lumbar pathology, for example. The questionnaires given to the patient must be carefully selected so as to elicit the most pertinent information in

**Table 3 Outcome tools in spinal surgery**

Topic	Tool	Notes
Pain	VAS	May be used for generalized or localized pain
Disability	ODI, NDI	Evaluates multiple life experiences The NDI is an adaptation of ODI for patients with neck disability
Myelopathy	JOA, mJOA	Evaluates motor function, sensation, and bladder function
Quality of life	SRS-22, EQ-5D-5L, SF-36	SRS-22 developed for patients with spinal deformity

VAS: Visual analog scale; NDI: Neck disability index; ODI: Oswestry disability index; JOA: Japanese Orthopaedic Association; mJOA: Modified JOA; SRS-22: Scoliosis Research Society-22; SF-36: Short Form-36.

the most efficient manner. It is important to recognize that providing an excessive number of questionnaires will decrease patient compliance. The aim of this section is to highlight the outcome measurement tools commonly used in spinal surgery research (Table 3).

The Oswestry disability index (ODI) is the most common questionnaire utilized to evaluate the physical symptoms of patients with low back pain, with an emphasis on quality of life<sup>[12]</sup>. This questionnaire evaluates ten categories: pain, personal care, lifting, walking, sitting, standing, sleeping, sex life, social life, and traveling. There are six answers available per question with point values of zero to five; the maximum score is fifty. A quantification of patient disability may be calculated by dividing the point total by fifty then multiplying by one hundred percent. Those with 0%-20% disability are considered minimally disabled. A score of 21%-40% is moderate disability, 41%-60% is severe, 61%-80% is crippled. Those with scores of 81%-100% are bed bound. A change of 4 points is the minimum difference that can be considered clinically significant. A 15 point change, though, is what is considered significant for patients undergoing spinal fusion.

The neck disability index (NDI) represents a modification of the ODI for patients with cervical spine pain<sup>[13]</sup>. The questions elicit information about activities such as concentration and reading which can be affected by cervical pain. Soft tissue injury can also lead to headache, which is also evaluated by the NDI. The scoring system is the same as that of the ODI.

The visual analog scale is a measurement instrument which quantifies patient subjective pain<sup>[14]</sup>. It consists of a 10 centimeter line with one end representing no pain and the other end the worst pain possible. The patient indicates where on the line his or her pain is in relation to these two extremes. This outcome measure can be used to quantify generalized pain or any specific type of pain (back, leg, *etc.*).

Patients with cervical myelopathy may suffer from a constellation of disabling symptoms, but pain may be a relatively minor issue. The Japanese Orthopaedic Association (JOA) scale is an objective assessment of upper and lower extremity motor function, sensation, and bladder function<sup>[15]</sup>. The highest possible score is 17. The JOA is specific for patients who utilize chopsticks to feed themselves. For those who do not, the modified JOA (mJOA)

has been developed<sup>[16,17]</sup>; the questionnaire has replaced the word “chopsticks” with “knife and fork”. The mJOA is therefore more often used in the United States compared to the JOA. In addition, the mJOA involves a highest score of 18, rather than the JOA’s high score of 17<sup>[15,17]</sup>.

Those with spinal deformities, in particular idiopathic scoliosis, have a slightly different set of concerns and health issues than those with degenerative conditions. These patients are typically adolescents or young adults. The Scoliosis Research Society-22 (SRS-22) questionnaire targets 5 domains: physical function, pain, self image, mental health, and satisfaction with management of scoliosis<sup>[18]</sup>. Each of the 22 questions contains 5 answers with point values from 1 to 5, with 5 being the best. The mean scores from each of the 5 categories are averaged to produce a single value. Studies indicate that significant point differences for the SRS-22 are: pain 0.6, function 0.8, self image 0.5, mental health 0.4, average sum score 0.5, and raw sum score 6.8.

The EQ-5D-5L is a questionnaire that investigates patient quality of life<sup>[19]</sup>. It consists of 2 forms. The first assesses 5 dimensions: mobility, self-care, usual activities, pain/discomfort, and anxiety/depression. Each dimension is associated with 5 statements; the patient selects the statement that most correlates with their condition. No point values are assigned to each statement. The second form consists of a 20 cm vertical line with endpoints labeled “the best health you can imagine” and “the worst health you can imagine.” Patients are asked to indicate where on the line they believe their present state of health to be. Given that no numerical score is calculated from the 5 questions, the data can be presented in a variety of formats.

Similar to the EQ-5D-5L is the Short Form-36 (SF-36)<sup>[20,21]</sup>. This is a health survey analyzing 2 general domains: physical health and mental health. There are 36 questions and 5 possible responses per question. Physical health is divided into physical functioning, physical role functioning, bodily pain, and general health. Mental health is divided into vitality, social functioning, emotional role functioning, and mental condition. The SF-36 consists of eight scaled scores, which are the weighted sums of the questions in each section. Each scale is directly transformed into a 0-100 scale on the assumption that each question carries equal weight.

Ultimately, an outcome assessment tool must be reliable, reproducible, specific to the outcome of interest, yet brief enough to promote compliance. For these purposes, an array of well-validated standardized questionnaires is available.

## CONCLUSION

Evidence-based research in spinal surgery has received a growing amount of attention, not only from surgeons and scientists, but from government regulators and the lay press. With continued pressure to produce high quality evidence for the success of spinal interventions, one must recognize that the RCT is not always the optimal study design for a given research question. Rather, prospective observational cohort studies may be more appropriate in certain circumstances, and would provide superior generalizability. In addition, case series and case-control study designs have their own utility, particularly in studying rare diseases and new treatment options. The use of validated outcome measurements will increase the quality of a given study design. Finally, evaluating spinal research with levels of evidence, I - V, allows for an objective measurement of study quality, yet this system does not account for whether the correct question was asked or if the correct patient population was studied. These issues surrounding EBM are important for spinal surgeons, as well as health policy makers and editorial boards, to have familiarity.

## REFERENCES

- 1 **Fisher CG**, Vaccaro AR. Improving the best available evidence component of evidence-based medicine: it's all in the question! *Spine* (Phila Pa 1976) 2013; **38**: E28-E29 [PMID: 23026868 DOI: 10.1097/BRS.0b013e318275cbd7]
- 2 **Croft P**, Malmivaara A, van Tulder M. The pros and cons of evidence-based medicine. *Spine* (Phila Pa 1976) 2011; **36**: E1121-E1125 [PMID: 21629165 DOI: 10.1097/BRS.0b013e318223ae4c]
- 3 **Glassman SD**, Branch CL. Evidence-based medicine: raising the bar. *Spine J* 2007; **7**: 513-515 [PMID: 17602890 DOI: 10.1016/j.spinee.2007.04.011]
- 4 **Vaccaro AR**, Fisher CG. "Help, I need guidance in how to manage my patients but I am so confused": evidence-based medicine to the rescue? *Spine* (Phila Pa 1976) 2012; **37**: E873-E874 [PMID: 22751142 DOI: 10.1097/BRS.0b013e3182454dd3]
- 5 **Fisher CG**, Wood KB. Introduction to and techniques of evidence-based medicine. *Spine* (Phila Pa 1976) 2007; **32**: S66-S72 [PMID: 17728684 DOI: 10.1097/BRS.0b013e318145308d]
- 6 **Dettori JR**, Norvell DC, Dekutoski M, Fisher C, Chapman JR. Methods for the systematic reviews on patient safety during spine surgery. *Spine* (Phila Pa 1976) 2010; **35**: S22-S27 [PMID: 20407348 DOI: 10.1097/BRS.0b013e3181d70494]
- 7 **Jacobs WC**, Kruyt MC, Verbout AJ, Oner FC. Spine surgery research: on and beyond current strategies. *Spine J* 2012; **12**: 706-713 [PMID: 23021034 DOI: 10.1016/j.spinee.2012.08.424]
- 8 **Wright JG**, Swiontkowski MF, Heckman JD. Introducing levels of evidence to the journal. *J Bone Joint Surg Am* 2003; **85-A**: 1-3 [PMID: 12533564]
- 9 **Weinstein JN**, Tosteson TD, Lurie JD, Tosteson AN, Hanscom B, Skinner JS, Abdu WA, Hilibrand AS, Boden SD, Deyo RA. Surgical vs nonoperative treatment for lumbar disk herniation: the Spine Patient Outcomes Research Trial (SPORT): a randomized trial. *JAMA* 2006; **296**: 2441-2450 [PMID: 17119140 DOI: 10.1001/jama.296.20.2441]
- 10 **Allen RT**, Rihn JA, Glassman SD, Currier B, Albert TJ, Phillips FM. An evidence-based approach to spine surgery. *Am J Med Qual* 2009; **24**: 15S-24S [PMID: 19890181 DOI: 10.1177/1062860609348743]
- 11 **Deyo RA**, Battie M, Beurskens AJ, Bombardier C, Croft P, Koes B, Malmivaara A, Roland M, Von Korff M, Waddell G. Outcome measures for low back pain research. A proposal for standardized use. *Spine* (Phila Pa 1976) 1998; **23**: 2003-2013 [PMID: 9779535 DOI: 10.1097/00007632-199809150-00018]
- 12 **Fairbank JC**, Pynsent PB. The Oswestry Disability Index. *Spine* (Phila Pa 1976) 2000; **25**: 2940-2952; discussion 2952 [PMID: 11074683]
- 13 **Vernon H**, Mior S. The Neck Disability Index: a study of reliability and validity. *J Manipulative Physiol Ther* 1991; **14**: 409-415 [PMID: 1834753]
- 14 **Scott J**, Huskisson EC. Graphic representation of pain. *Pain* 1976; **2**: 175-184 [PMID: 1026900 DOI: 10.1016/0304-3959(76)90113-5]
- 15 **Yonenobu K**, Abumi K, Nagata K, Taketomi E, Ueyama K. Interobserver and intraobserver reliability of the Japanese orthopaedic association scoring system for evaluation of cervical compression myelopathy. *Spine* (Phila Pa 1976) 2001; **26**: 1890-1894; discussion 1895 [PMID: 11568701 DOI: 10.1097/00007632-200109010-00014]
- 16 **Chiles BW**, Leonard MA, Choudhri HF, Cooper PR. Cervical spondylotic myelopathy: patterns of neurological deficit and recovery after anterior cervical decompression. *Neurosurgery* 1999; **44**: 762-769; discussion 762-769 [PMID: 10201301 DOI: 10.1097/00006123-199904000-00041]
- 17 **Benzel EC**, Lancon J, Kesterson L, Hadden T. Cervical laminectomy and dentate ligament section for cervical spondylotic myelopathy. *J Spinal Disord* 1991; **4**: 286-295 [PMID: 1802159 DOI: 10.1097/00002517-199109000-00005]
- 18 **Bagó J**, Pérez-Gruoso FJ, Les E, Hernández P, Pellisé F. Minimal important differences of the SRS-22 Patient Questionnaire following surgical treatment of idiopathic scoliosis. *Eur Spine J* 2009; **18**: 1898-1904 [PMID: 19533179 DOI: 10.1007/s00586-009-1066-x]
- 19 **Brooks R**. EuroQol: the current state of play. *Health Policy* 1996; **37**: 53-72 [PMID: 10158943 DOI: 10.1016/0168-8510(96)00822-6]
- 20 **Brazier J**, Jones N, Kind P. Testing the validity of the Euroqol and comparing it with the SF-36 health survey questionnaire. *Qual Life Res* 1993; **2**: 169-180 [PMID: 8401453 DOI: 10.1007/BF00435221]
- 21 **Haley SM**, McHorney CA, Ware JE. Evaluation of the MOS SF-36 physical functioning scale (PF-10): I. Unidimensionality and reproducibility of the Rasch item scale. *J Clin Epidemiol* 1994; **47**: 671-684 [PMID: 7722580 DOI: 10.1016/0895-4356(94)90215-1]

**P- Reviewers:** Kaneyama S, Kasai Y, Park P, Teli MGA  
**S- Editor:** Zhai HH **L- Editor:** A **E- Editor:** Lu YJ





百世登

**Baishideng**®

Published by **Baishideng Publishing Group Co., Limited**

Flat C, 23/F., Lucky Plaza, 315-321 Lockhart Road,

Wan Chai, Hong Kong, China

Fax: +852-65557188

Telephone: +852-31779906

E-mail: [bpgoffice@wjgnet.com](mailto:bpgoffice@wjgnet.com)

<http://www.wjgnet.com>

