Editorial: Embracing Rasch Analysis for Enhanced Spine Surgery Outcomes––The Outsider's Viewpoint

Igor Elman

Int J Spine Surg published online 4 November 2024
https://www.ijssurgery.com/content/early/2024/10/28/8671

The International Journal of Spine Surgery
2397 Waterbury Circle, Suite 1,
Aurora, IL 60504, Phone: +1-630-375-1432

International Journal of Spine Surgery, Vol. 00, No. 0, 2024, pp. 1–5
https://doi.org/10.14444/8671

Editorial: Embracing Rasch Analysis for Enhanced Spine Surgery Outcomes—The Outsider's Viewpoint

IGOR ELMAN, MD[1]

[1]*Department of Psychiatry, Cambridge Health Alliance/Harvard Medical School, Cambridge, MA, USA*

As an experienced clinical investigator, I was invited by Dr. Morgan Lorio, Co-President of the International Society for the Advancement of Spine Surgery (ISASS), to review and provide commentary on the articles in this special issue of the *International Journal of Spine Surgery* (*IJSS*). This issue focuses on the intricate relationship between a surgeon's experience, clinical judgment, and skill level with a chosen procedure and how these factors influence clinical outcomes as perceived by patients. Some may question how a clinician researcher from a different subspecialty and nonsurgeon is qualified to address these critical surgical outcome issues. My role as medical director at the Community Mental Health Center, coupled with my academic appointment at Harvard Medical School and leadership position at the Providence VA Medical Center's Substance Abuse Treatment Program, has endowed me with significant insights into advancing evidence-based treatments for complex painful conditions. These conditions are often associated with chemical and behavioral addictions, psychosis, and posttraumatic stress disorder, and so we must frequently treat patients concurrently suffering from psychiatric and chronic pain syndromes.

My theoretical knowledge and hands-on experience in various clinical trial research methodologies have been honed through my involvement in several National Institutes of Health–funded and other sponsored clinical trials. These trials encompass drug trials, protocol trials, and other properly designed studies, including double blinding, control groups, and prospective randomization. Having served as chair of 2 academic departments of psychiatry, I continuously applied my expertise in these methodologies to ensure the reliability and validity of trial outcomes, thereby contributing substantially to the fields of psychiatry and other areas of behavioral health. These areas often intersect substantially with interventional and surgical spinal pain management. Therefore, I present my opinions on the authors' research as an unbiased outsider, recognizing the fundamental similarities in clinical research, whether medical or surgical.

OVERCOMING LIMITATIONS OF TRADITIONAL CLINICAL TRIALS

Clinical trials in spine surgery face substantial limitations, often leading to the dismissal of innovative therapies. Challenges include randomization issues, crossover problems, and the difficulty of blinding in surgical trials.[1,2] Therefore, most randomized clinical trials in spine surgery have turned into well-controlled observation cohort studies with an intent-to-treat analysis. The fast-moving nature of surgical technologies can quickly render ongoing trials outdated. Moreover, systematic literature reviews suggest that well-designed prospective observational cohort studies could provide higher-grade evidence than poorly executed randomized trials.[3] This phenomenon has been coined the "glass ceiling" effect in outcome research in surgery.[4,5] It has been described in detail, along with a newly updated pyramid of clinical evidence examination.[6] However, the most relevant confounding factor in any surgical randomized clinical trial is the inability to randomize the surgeon's skill level.

Despite these limitations, Grade A clinical evidence can still be generated through observational studies if the reported outcomes are consistent and reproducible. This was demonstrated in the cases of arthroscopic knee surgery for degenerative osteoarthrosis or meniscal tears,[7,8] vertebroplasty for osteoporotic vertebral fractures,[9] and subacromial decompression for shoulder impingement.[10] Although these procedures did not withstand the scrutiny of sham-controlled randomized trials,[5] they are now well-established standards. Traditional systematic reviews and meta-analyses of extracted and processed data have long been considered the pinnacle of clinical evidence,[11–13] a concept still supported by the American Medical Association.[14] A new approach suggests that these types of investigations be

Figure. The concept of "evidence-based medicine" (EBM) was introduced by Dr. Gordon Guyatt at McMaster University.[16] The foundational work, also promoted by his adviser, Dr. David Sackett,[17,18] has since seen the term EBM become widespread in the medical field. Despite its common use, a deep comprehension of its full meaning is not as pervasive. EBM was originally crafted to blend 3 essential elements—best research evidence, clinical expertise, and patient values—as articulated by Sackett. This tripartite foundation, although sometimes neglected in conversations about EBM, is crucial to its application in health care. *Source:* Ekhtiari et al, "Surgeon Intuition: Fact or Fiction?" *Arthroscopy*. 2023;39(11):2269–2270.[19] Licensed under CC License CC BY 4.0.

used as a magnifying tool to examine available clinical evidence in greater detail rather than considering them the ultimate high standard.[15] This perspective allows for the validation of new protocols, technologies, or procedures without the necessity of achieving these traditional benchmarks as the highest-grade evidence that translates into changes in treatment recommendations.

Traditional systematic processes of literature review and committee work are laborious and resource-intensive, often outpaced by rapid advancements in spine surgery. The concept of "evidence-based medicine" (EBM), introduced by Dr. Gordon Guyatt at McMaster University,[16] originally blended best research evidence, clinical expertise, and patient values (Figure). In spine surgery, clinical evidence and guideline development face challenges due to the complexity of spinal disorders and the diversity of potential treatments.

UP-TO-DATE ANALYSIS IN SPINE SURGERY WITH RASCH ANALYSIS

Guideline development can be compromised by various biases, such as hindsight bias, where researchers may unintentionally affirm preconceived notions of clinical outcomes for specific surgical procedures, particularly when the outcomes are already known.[20,21] It can also be swayed by industry funding or other conflicts of interest, leading to skewed recommendations.[22,23] One of the biggest challenges in developing clinical guidelines is the bureaucratic process involved, which can be tedious and time-consuming. The clinical study of innovations often takes years to become established in the literature. As a result, clinical guidelines, which typically rely on retrospective analysis and exclude forward-looking perspectives, struggle to keep pace with rapid advancements in innovation and technology. While frequent updates are ideal, they are largely impractical.

The Rasch analysis approach of surgeon experience and clinical outcomes offers a more up-to-date rapid and timely assessment of technology innovations.[24] Nearly 25 years ago, it gained traction when Dr. Robert Florin, a neurosurgeon and member of the American Medical Association Relative Value Update Committee, used it to identify misvalued Current Procedural Terminology codes by analyzing total or intraservice work through a small-group panel comparison within code families. The method's aim is to pinpoint statistical outliers that are misaligned or compressed regarding physician work effort using a simpler approach than magnitude estimation. This technique is Rasch-paired comparison, which merges educational research methods and statistical regression, reorders families of codes, and converts the new scoring system into total work Relative Value Units. After the first 5-year review of the Medicare fee schedule, Dr. Florin undertook a systematic study to identify ranking anomalies among families of codes and address concerns from many surgeons that the resource-based relative value scales were compressed in certain families and possibly across specialties. The Relative Value Update Committee approach for generating Relative Value Unit weights for new procedures involves a form of paired comparison, where surveyed physicians report estimated time and complexity relative to 2 reference procedures closest to the new code.

This approach, called Rasch measurement analysis, was first developed by Georg Rasch in the 1950s to 1970s.[24,25] It has been used in various disciplines, including education, health outcomes research, physiology, psychophysics, writing performance, mathematics, marketing, physics, and ethical valuation. Paired comparison is a subanalysis under the broader polytomous Rasch analysis employed by the authors of this special issue. Rasch methods are based on psychometric literature dealing with measurement models, specifically:

$$Lni = \frac{Bn}{Di},$$

where *Lni* is the Rasch score, *Bn* represents the ability level of the *n-th* person (eg, test taker or rater), and *Di* is the difficulty of the *i-th* test item.[26] The measure incorporates both the abilities of the test taker or rater and the difficulty of the test items (ie, a specific spine surgery. In the context of physicians rating the work involved between 2 procedures, the Rasch ratio is reversed, with higher difficulty positively influencing the numerator and the physician's ability to perform difficult procedures in the denominator. Applying this concept to polytomous Rasch analysis, where multiple categories of procedure difficulty are considered, the model assesses varying levels of difficulty and the corresponding physician experience or confidence in obtaining favorable clinical outcomes. Higher difficulty levels increase the numerator, indicating greater complexity, while the denominator reflects the physician's experience and skill level. This adjustment in the Rasch model would impact clinical outcomes by providing a more detailed and granular understanding of how physician experience influences the effectiveness and efficiency of different procedures. By accurately weighing the difficulty of procedures against physician experience, the model can better predict which physicians are more likely to succeed with complex procedures, ultimately leading to improved patient care and outcomes by identifying high-value procedures and centers of excellence. This approach ensures that clinical guidelines and recommendations are based on a comprehensive assessment of physician capabilities across a range of procedural complexities, enhancing the quality and reliability of spine care delivery.

The researchers contributing to this webinar series and special issue, led by Dr. Kai-Uwe Lewandrowski, conducted a comprehensive review integrating insights from 4 ISASS webinars, resulting in recommendations for best clinical practices in endoscopic spine surgery. This research highlights the limitations of traditional surgical trials and amalgamates surgeons' experiences with cutting-edge techniques. Data from 3639 surgeons globally were analyzed using the polytomous Rasch model, ensuring a robust statistical assessment of endorsements and educational impacts. This approach focused on operative nuances and experience-based outcomes, with bias detection performed using the differential item functioning test.

The ISASS webinars (accessible at https://isass.org/education/isass-past-insights-webinars/) provided a dynamic platform for discussing advances in endoscopic spine surgery, identifying high-value procedures such as:

- percutaneous interlaminar endoscopic decompression for lateral canal stenosis
- transforaminal debridement of low-grade de generative spondylolisthesis
- transforaminal full-endoscopic interbody fusion for hard disc herniation
- endoscopic standalone lumbar interbody fusion
- endoscopic debridement of spondylolytic spondylolisthesis
- posterior cervical foraminotomy for herniated disc and bony stenosis

These procedures were identified through higher-intensity endorsement transformation from the pre- to postwebinar survey, with a shift to higher mean logit locations of test items both with unbiased and orderly threshold progression.

IDENTIFYING HIGH-VALUE PROCEDURES

Using Rasch analysis to evaluate the interplay between surgeon skill and favorable clinical outcomes in endoscopic spine surgery offers a comprehensive way to assess and enhance surgical quality by identifying high-value procedures. High-value procedures are more cost-effective in the long run because of lower revision rates and longer-lasting clinical benefits. By demonstrating how surgeon competence directly affects clinical success and health care value, this approach provides a data-driven foundation for advancing surgical practices and health care policies.

A COMPREHENSIVE APPROACH TO ENHANCING SPINE SURGERY

The ISASS webinar series significantly impacted surgeons' education and contributed to identifying high-value endoscopic spine surgery practices that may serve as cornerstones for surgeon training standards, policies, and guidelines development. I encourage the spine surgery and public health fields to recognize the importance of surgeons' experiences alongside patient expectations, adding another layer to EBM in conjunction with clinical trials.

By leveraging Rasch analysis, health care systems can maximize the value delivered to patients while minimizing unnecessary costs and improving overall treatment efficacy. Through this analysis, health care systems can set benchmarks for surgical competence, tailor training programs to address identified skill gaps, and prioritize resources toward the most impactful

training techniques, ensuring a better allocation of health care resources and promoting a more sustainable health care system.

Integrating surgeon experience research, patient expectations, and clinical evidence grading within spine surgery practices indeed seems to be a complex endeavor. Controlled clinical trials, while not a panacea, are indispensable for rigorous validation. However, methods like Rasch analysis can complement these trials by providing valuable insights from real-world surgical experiences, thus enriching the evidence base.

The research by the webinars' team of surgeons addresses some deficiencies in traditional clinical trial research, providing rapid and efficient insights. About 150 years ago, another physician who was exasperated with the health care system, Anton Chekhov, wrote to his brother:

> *In descriptions of Nature, one must seize on small details, grouping them so that when the reader closes his eyes he gets a picture. For instance, you will have a moonlit night if you write that on the mill dam a piece of glass from a broken bottle glittered like a bright little star, and the black shadow of a dog or a wolf rolled past like a ball.*[27]

The practice of formulating medical hypotheses has regrettably been overshadowed by an overemphasis on clinical trials, often neglecting the other 2 critical components of EBM: physician experience and patient values. Many innovations begin with personal observations, which can constitute Grade A clinical evidence if consistently validated. Relying exclusively on clinical trials for every innovation can distort the evidence discussion, potentially limiting surgeons' autonomy and restricting patients' access to high-quality care.

I am thrilled about the authors' intriguing ideas to seek more practical solutions for real-world challenges in surgical trials. Ultimately, a balanced discussion recognizing the complementary strengths of Rasch analysis of surgeons' experience, skill, and their assessment of clinical outcomes in conjunction with clinical trials can provide a more realistic view, fostering advancements in spine surgery outcomes through the integration of innovative methodologies with established research practices.

REFERENCES

1. Weinstein JN, Tosteson TD, Lurie JD, et al. Surgical vs nonoperative treatment for lumbar disk herniation: the spine patient outcomes research trial (SPORT): a randomized trial. *JAMA*. 2006;296(20):2441–2450. doi:10.1001/jama.296.20.2441
2. Abdu WA, Sacks OA, Tosteson ANA, et al. Long-term results of surgery compared with nonoperative treatment for lumbar degenerative spondylolisthesis in the Spine Patient Outcomes Research Trial (SPORT). *Spine*. 2018;43(23):1619–1630. doi:10.1097/BRS.0000000000002682
3. Burns PB, Rohrich RJ, Chung KC. The levels of evidence and their role in evidence-based medicine. *Plast Reconstr Surg*. 2011;128(1):305–310. doi:10.1097/PRS.0b013e318219c171
4. Phillips FM, Cheng I, Rampersaud YR, et al. Breaking through the "glass ceiling" of minimally invasive spine surgery. *Spine*. 2016;41 Suppl 8:S39–S43. doi:10.1097/BRS.0000000000001482
5. Solheim O. Randomized controlled trials in surgery and the glass ceiling effect. *Acta Neurochir*. 2019;161(4):623–625. doi:10.1007/s00701-019-03850-3
6. Lewandrowski K-U, León JFR, Dowling Á, et al. Breaking through the glass ceiling effect of high-grade clinical evidence creation in orthopaedics & trauma. *Rev Colomb Ortop Traumatol*. 2022;36(4):215–228. doi:10.1016/j.rccot.2022.10.003
7. Moseley JB, O'Malley K, Petersen NJ, et al. A controlled trial of arthroscopic surgery for osteoarthritis of the knee. *N Engl J Med*. 2002;347:81–88.
8. Sihvonen R, Paavola M, Malmivaara A, et al. Arthroscopic partial meniscectomy versus sham surgery for a degenerative meniscal tear. *N Engl J Med*. 2013;369:2515–2524.
9. Buchbinder R, Osborne RH, Ebeling PR, et al. A randomized trial of vertebroplasty for painful osteoporotic vertebral fractures. *N Engl J Med*. 2009;361:557–568.
10. Paavola M, Malmivaara A, Taimela S, et al. Subacromial decompression versus diagnostic arthroscopy for shoulder impingement: randomised, placebo surgery controlled clinical trial. *BMJ*. 2018;362.
11. Tomlin G, Borgetto B. Research pyramid: a new evidence-based practice model for occupational therapy. *Am J Occup Ther*. 2011;65(2):189–196. doi:10.5014/ajot.2011.000828
12. Rosner AL. Evidence-based medicine: revisiting the pyramid of priorities. *J Bodyw Mov Ther*. 2012;16(1):42–49. doi:10.1016/j.jbmt.2011.05.003
13. Paez A. The "architect analogy" of evidence-based practice: reconsidering the role of clinical expertise and clinician experience in evidence-based health care. *J Evid Based Med*. 2018;11(4):219–226. doi:10.1111/jebm.12321
14. Murad MH, Montori VM, Ioannidis JPA, et al. How to read a systematic review and meta-analysis and apply the results to patient care: users' guides to the medical literature. *JAMA*. 2014;312(2):171–179. doi:10.1001/jama.2014.5559
15. Murad MH, Asi N, Alsawas M, Alahdab F. New evidence pyramid. *Evid Based Med*. 2016;21(4):125–127. doi:10.1136/ebmed-2016-110401
16. Guyatt GH. Evidence-based medicine. *ACP J Club*. 1991;114(2). doi:10.7326/ACPJC-1991-114-2-A16
17. Sackett DL. How to read clinical journals: I. Why to read them and how to start reading them critically. *Can Med Assoc J*. 1981;124:555–558.
18. Sackett DL. Rules of evidence and clinical recommendations on the use of antithrombotic agents. *Chest*. 1989;95(2 Suppl):2S–4S.
19. Ekhtiari S, Khanduja V. Surgeon intuition: fact or fiction? *Arthroscopy*. 2023;39(11):2269–2270. doi:10.1016/j.arthro.2023.08.002

20. Henriksen K, Kaplan H. Hindsight bias, outcome knowledge and adaptive learning. *Qual Saf Health Care*. 2003;12 Suppl 2(Suppl 2):ii46–ii50. doi:10.1136/qhc.12.suppl_2.ii46

21. Zwaan L, Monteiro S, Sherbino J, Ilgen J, Howey B, Norman G. Is bias in the eye of the beholder? a vignette study to assess recognition of cognitive biases in clinical case workups. *BMJ Qual Saf*. 2017;26(2):104–110. doi:10.1136/bmjqs-2015-005014

22. Murayama A, Kamamoto S, Murata N, et al. Evaluation of financial conflicts of interest and quality of evidence in Japanese gastroenterology clinical practice guidelines. *J Gastroenterol Hepatol*. 2023;38:565–573.

23. Stegeman B, Schep A, Kuijpers T, Hofstede S. GRADE notes 3: two approaches to assess industry sponsorship bias used by two Dutch guidelines organizations. *J Clin Epidemiol*. 2023;164:9–14.

24. Boone WJ. Rasch analysis for instrument development: why, when, and how? *CBE Life Sci Educ*. 2016;15(4). doi:10.1187/cbe.16-04-0148

25. Aryadoust V, Tan HAH, Ng LY. A scientometric review of Rasch measurement: the rise and progress of a specialty. *Front Psychol*. 2019;10:2197.

26. Andrich D. An elaboration of guttman scaling with Rasch models for measurement. *Sociol Methodol*. 1985;15:33. doi:10.2307/270846

27. Chekhov AP. *The Unknown Chekhov: Stories and Other Writings Hitherto Untranslated*. New York, USA: Noonday Press; 1954.

Funding: Lange MedTech provided support for the publication of this special issue to ISASS. The author received no support for the writing or publication of this article.

Declaration of Conflicting Interests: The author reports no conflicts of interest in this work.

Corresponding Author: Igor Elman, Cambridge Health Alliance/Harvard Medical School Cambridge, 1493 Cambridge St, Cambridge, MA 02139, USA; dr.igorelman@gmail.com