



ORIGINAL

## Breaking through the glass ceiling effect of high-grade clinical evidence creation in orthopaedics & trauma



Kai-Uwe Lewandrowski<sup>a,\*</sup>, Jorge Felipe Ramírez León<sup>b</sup>, Álvaro Dowling<sup>c</sup>, Manuel Rodríguez García<sup>d</sup>, José Gabriel Rugeles<sup>e</sup>, Carolina Ramírez<sup>f</sup>, Alfonso García<sup>g</sup>, Jose Valerio<sup>h</sup>, Paulo Sérgio Teixeira de Carvalho<sup>i</sup>, Luis Miguel Duchén Rodríguez<sup>j</sup>, Jaime Moyano<sup>k</sup>, Mario Herrera<sup>l</sup>, Nicolás Prada<sup>m</sup>, Mauricio Zuluaga<sup>n</sup>, Anthony Yeung<sup>o</sup>

<sup>a</sup> Center for Advanced Spine Care of Southern Arizona and Surgical Institute of Tucson, Tucson AZ, USA, Full Professor Department of Orthopaedics, Fundación Universitaria Sanitas, Bogotá, D.C., Colombia

<sup>b</sup> Minimally Invasive Spine Center, Bogotá, D.C., Colombia, Reina Sofía Clinic, Bogotá, D.C., Colombia, Department of Orthopaedics, Fundación Universitaria Sanitas, Bogotá, D.C., Colombia

<sup>c</sup> Orthopaedic Spine Surgeon, Director of Endoscopic Spine Clinic, Santiago, Chile

<sup>d</sup> Spine Clinic, The American-British Cowdray Medical Center I.A.P, Campus Santa Fe, México City, México

<sup>e</sup> Cecimin, Bogotá, Colombia

<sup>f</sup> Cecimin, Bogotá, Colombia

<sup>g</sup> Espalda Saludable, Hospital Angeles Tijuana, Tijuana, B. C. México

<sup>h</sup> Neurosurgery Chair Palmetto steward General Hospital Co-Director Miami Neuroscience Center at Larkin, FL USA

<sup>i</sup> Pain and Spine Minimally Invasive Surgery Service at Gaffre e Guinle University Hospital, Rio de Janeiro, Brazil

<sup>j</sup> President of the Chapter of the Vertebral Column of the Latin American Federation of Neurosurgery Societies, Bolivian Spine Association, Public University of El Alto and Center for Neurological Diseases, La Paz, Bolivia

<sup>k</sup> Ortopedia y Traumatología, Ecuador

<sup>l</sup> Clínica Colombia. Fundación Universitaria Sanitas, Bogotá, D.C., Colombia

<sup>m</sup> Profesor Cirugía de Columna, Universidad Autónoma de Bucaramanga, Bucaramanga, Colombia

<sup>n</sup> Ortopedista y traumatólogo, Clínica Imbanaco. Cali, Colombia

<sup>o</sup> Desert Institute for Spine Care, Phoenix, AZ, USA, Executive Director, International Intradiscal Therapy Society (IITS.org)

Received 27 July 2022; accepted 13 October 2022

Available online 25 October 2022

### KEYWORDS

Randomized  
Controlled Trial;  
Clinical Trial  
Protocol;  
Epidemiologic  
Research Design;  
Cohort Studies;  
Statistics

### Abstract

**Introduction:** Randomization protocols in clinical studies of common orthopedic problems and traumatic injuries are challenging to orchestrate. The lack of high-grade clinical evidence from prospective randomized, double-blinded study design is often cited as a primary reason for rejecting proposed therapy advances in orthopedic surgery.

**Materials and Methods:** This position paper summarizes the clinical trial limitations in surgical subspecialties. We present a consensus statement on how practicing orthopedic surgeon can produce high-quality clinical evidence to affect changes in practice protocols.

\* Corresponding author.

E-mail address: [business@tucsonspine.com](mailto:business@tucsonspine.com) (K.-U. Lewandrowski).

**Results:** Our literature review revealed that classifications of level of evidence reporting varies between surgical subspecialties. Research in orthopedic- and musculoskeletal trauma care is primarily directed at the diagnosis, preferred treatment, and economic decision analysis, whereas other prognosis-based classifications are preferred in other areas such as plastic surgery. In orthopedics, controlled double-blinded randomization is rare and often unpractical or unethical for a placebo control purpose where patients may be harmed. Crossing over between study groups randomized surgical trials is common. Other obstacles in surgical trials range from lack of organizational and financial support, institutional review or ethics board approval, and registration requirements for clinical trials to insufficient time left outside an already busy clinical schedule to dedicate to this laborious uncompensated task.

**Conclusion:** Orthopedic surgery is as an experience- and skill-based subspecialty. Many innovations start with entrepreneur surgeons as reporting of opinions or retrospective cohort studies many of which suffer from bias. Prospective observational cohort studies with consistent results may offer higher-grade clinical evidence than poorly executed randomized trials.

© 2022 Sociedad Colombiana de Ortopedia y Traumatología. Published by Elsevier España, S.L.U. All rights reserved.

## PALABRAS CLAVE

Ensayo controlado aleatorizado;  
Protocolo de ensayo clínico;  
Diseño de investigaciones epidemiológicas;  
Estudios de cohortes;  
Estadísticas

## Rompiendo el efecto "techo de cristal" en la generación de evidencia clínica de alto nivel en ortopedia y traumatología

### Resumen

**Introducción:** En los estudios clínicos relacionados con problemas ortopédicos comunes y lesiones traumáticas, los métodos de aleatorización son difíciles de orquestar. La falta de evidencia clínica de alto nivel que tomen como fuente estudios prospectivos, aleatorizados y doble ciego, se cita a menudo como una razón principal para rechazar los avances terapéuticos propuestos en cirugía ortopédica.

**Materiales y métodos:** Este documento de opinión resume las limitaciones de los ensayos clínicos en las subespecialidades quirúrgicas. Se presente un consenso acerca de cómo el cirujano ortopédico en ejercicio puede producir evidencia clínica de alta calidad y de esta forma realizar cambios en sus protocolos de práctica clínica.

**Resultados:** Esta revisión de la literatura reveló que las clasificaciones del nivel de evidencia varían entre las subespecialidades quirúrgicas. La investigación en la ortopedia y traumatología se dirige principalmente al diagnóstico, el tratamiento preferido y el análisis de decisiones económicas, mientras que otras clasificaciones de pronóstico son preferidas en otras áreas, como cirugía plástica. En ortopedia, los estudios doble ciego controlados son raros y, a menudo, pocos prácticos o, incluso, pocos éticos. El cruzamiento entre ensayos quirúrgicos aleatorios de grupos de estudio es más común. Otras dificultades en los ensayos quirúrgicos van desde: la falta de apoyo organizativo y financiero, aprobación institucional o comité de ética y requisitos de registro para los ensayos clínicos, y hasta tiempo insuficiente por fuera de un programa clínico, ya ocupado para dedicarle a esta laboriosa tarea no compensada.

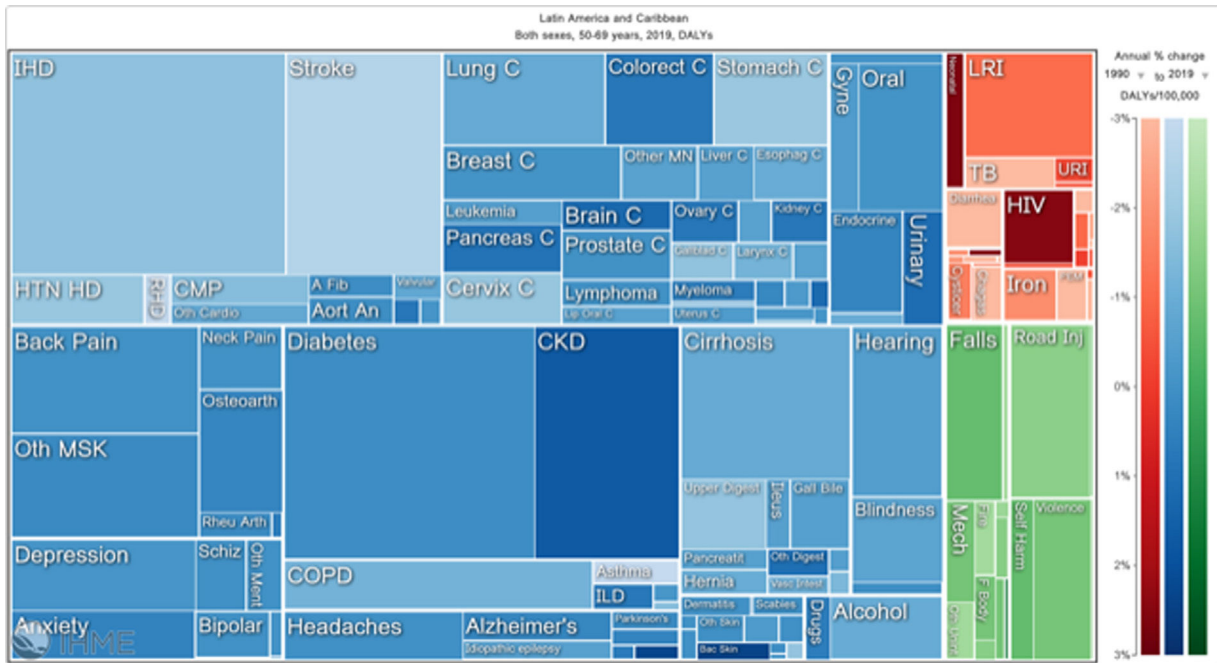
**Conclusión:** La cirugía ortopédica es una subespecialidad basada en la experiencia y la habilidad. Muchas innovaciones comienzan con cirujanos emprendedores reportando informes de opiniones o estudios de cohortes retrospectivos, muchos de los cuales tienen sesgo. Los estudios de cohortes observacionales prospectivos con resultados consistentes pueden ofrecer evidencia clínica de mayor grado que los ensayos aleatorios mal ejecutados.

© 2022 Sociedad Colombiana de Ortopedia y Traumatología. Publicado por Elsevier España, S.L.U. Todos los derechos reservados.

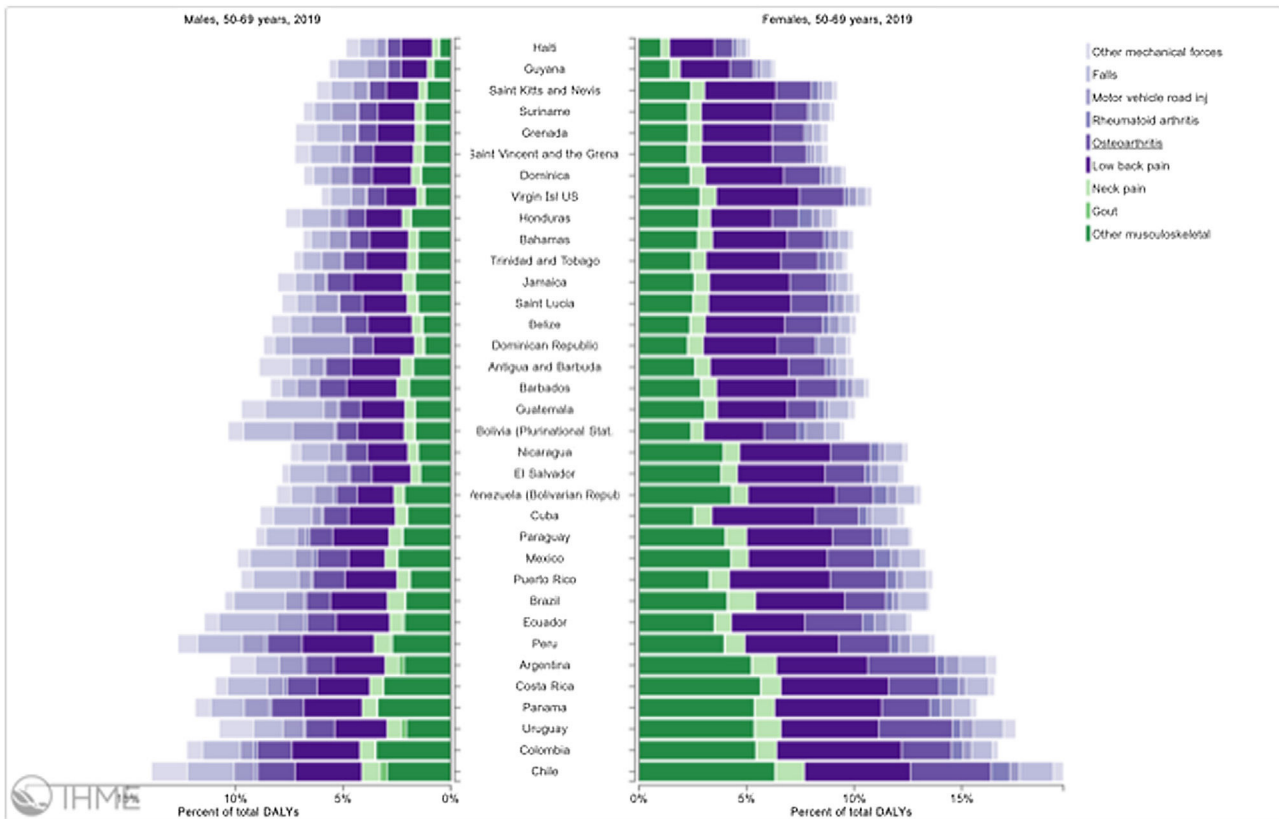
## Disease burden of musculoskeletal diseases

Orthopedic disorders and musculoskeletal trauma have a significant economic impact globally.<sup>1,2</sup> These conditions' acute and chronic nature are a major contributor to health care expenditure in Latin America and other industrialized

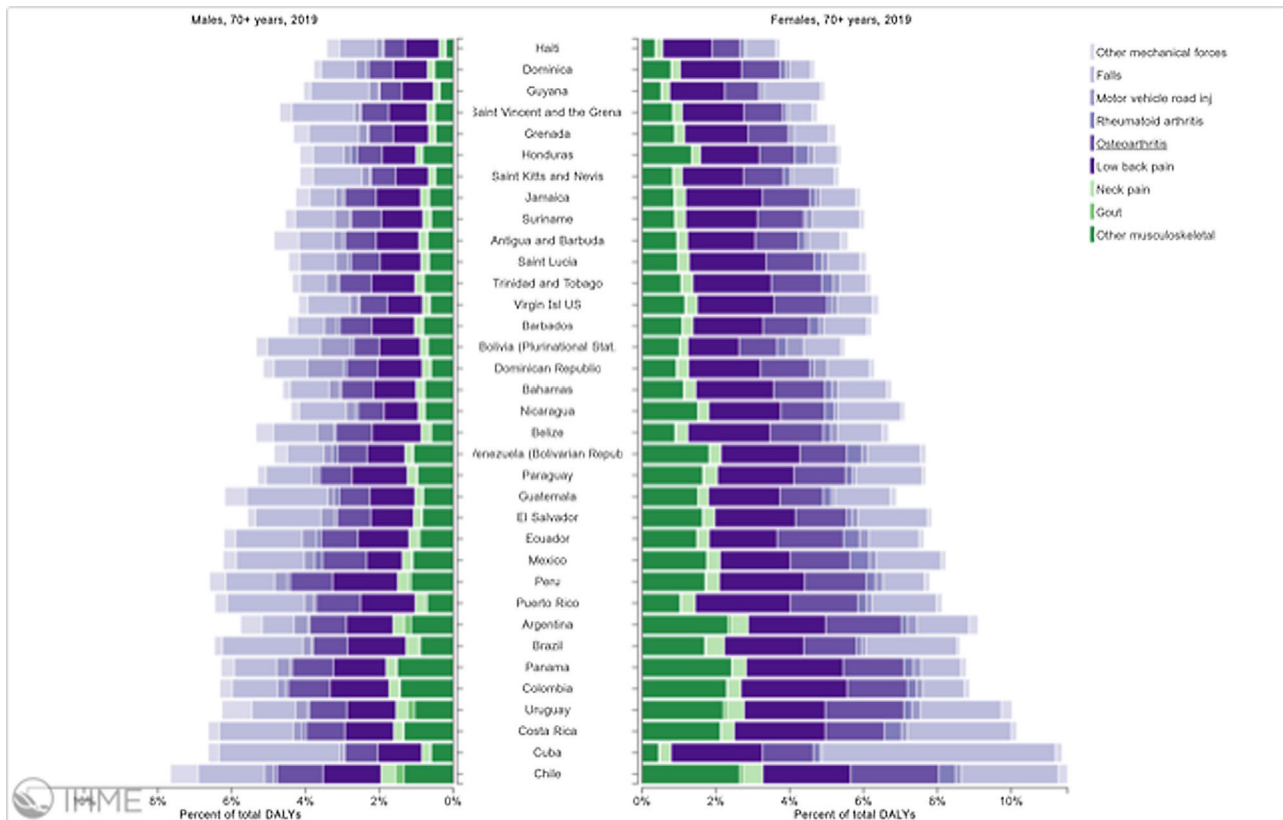
countries. Comprehensive care programs with effective collaborations between developed nations in the Americas are rare. The lack of resources to adequately manage trauma and diseases affecting the musculoskeletal system in developing countries and countries with emerging economies adds to the orthopedic disease burden. When measuring the



**Figure 1** Illustrative tree map of causes and disease burden expressed in years lived with disability (YLDs) in Latin America and the Caribbean in 2019. YLDs for Low back pain was 8.19% (6.3% - 10.19%), other musculoskeletal diseases 7.57% (5.54%-10.08%), osteoarthritis 4.12% (2.41% - 7.65%) neck pain 2.12% (1.36% - 3.22%), falls 2.59% (2.28 - 2.97), road injury 1.51% (1.37% - 1.67), exposure to mechanical forces 0.88% (0.69% - 1.14%), interpersonal violence 0.49% (0.42% - 0.56%), and other unintentional injuries 0.37% (0.3% - 0.47%): Source: Institute for Health Metrics Evaluation. Used with permission. All rights reserved.



**Figure 2** Pyramid plot of causes and disease burden expressed in years lived with disability (YLDs) by country in Latin America and the Caribbean in 2019. In the age group between 50 and 69 years, the combined YLDs for Low back pain, other musculoskeletal diseases, osteoarthritis, neck pain, falls, road injury, exposure to mechanical forces, interpersonal violence, and other unintentional injuries was 25.49% for males and 29.71%: Source: Institute for Health Metrics Evaluation. Used with permission. All rights reserved.



**Figure 3** Pyramid plot of causes and disease burden expressed in years lived with disability (YLDs) by country in Latin America and the Caribbean in 2019. In the age group over 70+ years, the combined YLDs for Low back pain, other musculoskeletal diseases, osteoarthritis, neck pain, falls, road injury, exposure to mechanical forces, interpersonal violence, and other unintentional injuries was 18.01% for males and 22.95%: Source: Institute for Health Metrics Evaluation. Used with permission. All rights reserved.

years lived with disability (YLDs) in Latin America and the Caribbean using the 2019 data of the Institute for Health Metrics Evaluation, a significant musculoskeletal disease burden becomes apparent with low back pain (8.19%; 6.3% - 10.19%) being the leading cause of disability (Figure 1). It is followed by osteoarthritis (4.12%; 2.41% - 7.65%), neck pain (2.12%; 1.36% - 3.22%), falls (2.59%; 2.28 - 2.97), road injury (1.51%; 1.37% - 1.67), exposure to mechanical forces (0.88%; 0.69% - 1.14%), interpersonal violence (0.49%; 0.42% - 0.56%), unintentional injuries (0.37%; 0.3% - 0.47%), and a group of other musculoskeletal diseases not included in the above (7.57%; 5.54%-10.08%; Source: Institute for Health Metrics Evaluation. Used with permission. All rights reserved). Breaking down the same YDL index by Latin American countries shows a higher orthopedic disease burden in the same categories between the ages of 50-69 years (Figure 2) for both men (YLDs: 25.49%) and women (YLDs: 29.71%) than above 70 years of age (18.01% for males and 22.95%; Figure 3). In the age group from 50 to 69, Colombia has the fourth highest disease burden measured in YLDs for rheumatoid and osteoarthritis, trauma, and falls after Argentina, Chile, and Uruguay (Figure 2). In the elderly (70+ years of age), the orthopedic disease burden is lower, and Colombia is preceded by Costa Rica, Mexico, Cuba, Brazil, Chile, Argentina, and Uruguay (Figure 3). Taking into account the population growth and a higher percentage of

elderly in Latin America with improved overall health due to a shift from communicable or infectious to manageable chronic diseases, such as ischemic heart disease and stroke, which were the top-ranked causes of YLDs in 2019, it is apparent that higher patient numbers should be expected in orthopedic clinics across Latin America many of which are already stretched to the brink of maximum capacity, particularly in the public health care sector. Considering that modern treatment of many orthopedic conditions and injuries is resource intense, a disproportionately greater health care expenditure and research and development investments are needed than in most other areas of public health to identify new, more effective therapies and surgeries. Orthopedic surgeons are at the heart of this debate on allocating resources most efficiently to achieve maximum reduction of musculoskeletal disease burden. The anticipation of these changes in demographics and an expanded emphasis on non-communicable diseases and injuries<sup>1</sup> require orthopedic surgeons to develop high-grade clinical evidence to avoid rationing of services pushed by decision-makers trying to mitigate cost explosion. This worst-case scenario would only make high-end modern orthopedic care affordable for the well-to-do, thereby increasing social disparities, which could become the target of counterproductive politicization strategies in the current environment of shifting political agendas in Latin America and the world over.

## Clinical trial standards

The 'good clinical practice' (GCP) standards for clinical trials were first formulated by the International Conference on Harmonization (ICH) in 1990 to protect human subjects participating in clinical trials.<sup>3</sup> The GCP guideline on human clinical trials is based on the ethical principles published in the Declaration of Helsinki in 1964<sup>4-7</sup> and its latest 2000 revision.<sup>8</sup> As an internationally accepted standard, GCP also promotes good science by formalizing all trial activities ranging from communication with the ethics or institutional review board committee, clinical study protocol, investigational product, randomization including blinding and unblinding procedures, informed consent, documentation, reporting, and trial staff responsibilities related to safely executing the trial protocol all while complying with the local law.

## The level of evidence

Various classifications for the level of clinical evidence have been adopted in numerous societies and journals. It was described initially by the Canadian Task Force on the Periodic Health Examination<sup>9</sup> in 1979 and expanded by Sackett in 1989 in an article on antithrombotic agents.<sup>10</sup> Since diverse specialties are interested in different clinical context, different types and level of clinical evidence is needed.<sup>11</sup> Research in orthopedic- and musculoskeletal trauma care may be directed at the diagnosis, preferred treatment, prognosis, and economic decision analysis.<sup>11</sup> For example, the grading system of the American Society of Plastic Surgeons (ASPS) is based on prognosis (Table 1).<sup>12-14</sup> A treatment-based level of evidence grading system has been proposed by the Centre for Evidence-Based Medicine (CEBM; Table 2).<sup>15</sup> It is apparent that quality of data assessment when assigning a level grading of the clinical evidence depends on the question asked – in this comparison, treatment versus prognosis.<sup>11</sup> Investigative studies on the natural history of an orthopedic condition cannot be analyzed with a randomized controlled trial (RCT) since no treatments are involved that can be compared. Hence, the highest grade of clinical evidence can only be derived from observational cohort studies or a systematic literature review of multiple cohort studies.<sup>11</sup> Many RCTs never play out as designed and generate lower quality data.<sup>16</sup> They are plagued by cross-over problems which hampered many of the anticipated comparisons in one of the largest orthopedic RCTs - the Spine Outcome research trial (SPORT).<sup>17,18</sup> The CEBM recommendations, therefore, assign the same level of evidence grading to poorly designed RCTs as cohort studies.<sup>11</sup>

## Practice recommendations

Modern clinical decision-making is guided by evidence-based medicine standards formulated as Practice Guidelines by many professional orthopedic subspecialty societies. Typically, a strong recommendation in favor of a diagnostic test or treatment is assigned when there is level I evidence and consistent evidence from Level II, III, and IV studies in support of it.<sup>11</sup> It is important to note that such a grading system for various clinical problems as the one proposed

by the American Society of Plastic Surgeons (Table 3)<sup>19</sup> and adopted by many societies such as the American Academy of Orthopedic Surgeons<sup>20,21</sup> or the North American Spine Society<sup>22-29</sup> does not degrade lower-level evidence when deciding recommendations if the results are consistent. In experienced-based surgical subspecialties, Level II, III, and IV studies, in fact, make up the majority of clinical studies forming the basis for the diagnosis-specific evidence-based clinical practice guidelines.

## The ceiling effect

Many innovations in surgery in general and in Orthopedics in particular have been implemented without clinical trials. Examples include the first heart transplant performed by Dr. Christiaan Barnard in 1967.<sup>30</sup> Since surgery is more skill- and experienced based than evidence based, many innovations over the last 200 years were implemented without any high-grade clinical evidence.<sup>31</sup> In fact, many examples of orthopedic innovation exist, such as arthroscopic knee surgery for degenerative osteoarthritis or meniscal tears,<sup>32,33</sup> vertebroplasty for osteoporotic vertebral fractures,<sup>34</sup> and subacromial decompression for shoulder impingement,<sup>35</sup> that did not withstand the scrutiny of sham-controlled randomized trial.<sup>16</sup> In fact, some of the major RCTs in orthopedic- and in spinal surgery failed to conclusively demonstrate the benefit of the intervention<sup>36</sup> in spite of cumulative and consistent low-grade evidence to the contrary.<sup>7,37-40</sup> It is obvious to the beholder that many surgical RCTs are far from the gold standard double-blinded prospective RCT and suffer from lower methodological quality.<sup>41</sup> The Jadad score is generally used to identify poor quality RCTs by assessing their methodological quality by scoring studies by randomization, masking, and patient retention (Table 4).<sup>42</sup> A maximum score of 5 can be assigned. Consolidated Standards for Reporting Trials (CONSORT) was started to standardize reporting criteria and improve the quality of clinical trials (Figure 4).<sup>43</sup> The CONSORT revised recommendations for improving the quality of reports of parallel-group randomized trials are illustrated in the flow diagram shown in Figure 4.<sup>44-46</sup> The obtainable level of evidence from a poor-quality RCT (Level 2B; Table 2) may not be better than a well-conducted prospective cohort study. As discussed in the following, the latter can ultimately be more informative, particularly when coupled with a durability or survival analysis of the desired treatment benefit.<sup>47-50</sup> In other words, breaking through this ceiling effect posed by the limitations of orthopedic study trials may prove nearly impossible considering the following problems with randomization in surgical patients.

## Randomization

Orthopedic RCTs are uncommon because several problems are hard to overcome. Orthopedic techniques and implants evolve and are constantly improved, making assessment over time difficult. The timing of an orthopedic RCT is not trivial. Early evaluation of a new implant or technique may be limited to a few surgeons who master it. On the contrary, the clinical equipoise may be lost if clinical trials are started too late. The technique may have become outdated when

**Table 1** Level of clinical evidence for prognostic studies (Adapted from Burns et al. and the American Society of Plastic Surgeons publications) <sup>11-14</sup>

Level	Type of evidence
I	High quality prospective cohort study with adequate power or systematic review of these studies
II	Lesser quality prospective cohort, retrospective cohort study, untreated controls from an RCT, or systematic review of these studies
III	Case-control study or systematic review of these studies
IV	Case series
V	Expert opinion; case report or clinical example; or evidence based on physiology, bench research or "first principles"

**Table 2** Level of clinical evidence for therapeutic studies (Adapted from Burns et al. and the Center for Evidence Based Medicine (CEBM) <sup>11,15</sup>

Level	Type of evidence
1A	Systematic review (with homogeneity) of RCTs
1B	Individual RCT (with narrow confidence intervals)
1C	All or none study
2A	Systematic review (with homogeneity) of cohort studies
2B	Individual Cohort study (including low quality RCT, e.g. < 80% follow-up)
2C	"Outcomes" research; Ecological studies
3A	Systematic review (with homogeneity) of case-control studies
3B	Individual Case-control study
4	Case series (and poor quality cohort and case-control study)
5	Expert opinion without explicit critical appraisal or based on physiology bench research or "first principles"

**Table 3** Grade Practice Recommendations (Adapted from Burns et al. and the American Society of Plastic Surgeons publications) <sup>11-14</sup>

Grade	Descriptor	Qualifying Evidence	Implications for Practice
A	Strong recommendation	Level I evidence or consistent findings from multiple studies of levels II, III, or IV	Clinicians should follow a strong recommendation unless a clear and compelling rationale for an alternative approach is present
B	Recommendation	Levels II, III, or IV evidence and findings are generally consistent	Generally, clinicians should follow a recommendation but should remain alert to new information and sensitive to patient preferences
C	Option	Levels II, III, or IV evidence, but findings are inconsistent	Clinicians should be flexible in their decision-making regarding appropriate practice, although they may set bounds on alternatives; patient preference should have a substantial influencing role
D	Option	Level V evidence: little or no systematic empirical evidence	Clinicians should consider all options in their decision making and be alert to new published evidence that clarifies the balance of benefit versus harm; patient preference should have a substantial influencing role

it was surpassed by newer technology. Orthopedic RCTs are also difficult to assess since the assumption that skill level and training among surgeons are equal is rarely true. Hence, standardization of interventions may be impossible. In addition, orthopedic surgeons may lose custody of their patients during essential periods of the postoperative recovery to physical therapy and many other subspecialties involved in

postoperative care that may impact clinical outcomes while being oblivious to the postoperative rehabilitation protocol prescribed by the study design.

Compared to drug trials, which typically start with Phase II RCTs to determine statistical power, endpoints, and adverse events before multicenter Phase III trials are begun, surgical RCTs rarely make it out of Phase II single-site study.

**Table 4** Jadad Scale (Adopted from Empirical Evidence of Associations Between Trial Quality and Effect Size)<sup>42</sup>

Dimension		Sub Score
Randomization	1. Was the study described as randomized (this includes the use of words such as randomly, random, and randomization)? = 1 point	Give 1 additional point if: For question 1, the method to generate the sequence of randomization was described and it was appropriate (table of random numbers, computer generated, etc.) Deduct 1 point if: For question 1, the method to generate the sequence of randomization was described and it was inappropriate (patients were allocated alternately, or according to date of birth, hospital number, etc.)
Blinding	2. Was the study described as double blind? = 1 point	Give 1 additional point: If for question 2 the method of double blinding was described and it was appropriate (identical placebo, active placebo, dummy, etc.) Deduct 1 point: If for question 2 the study was described as double blind but the method of blinding was inappropriate (e.g., comparison of tablet vs. injection with no double dummy)
Withdrawals and dropouts	3. Was there a description of withdrawals and dropouts? = 1 point	
TOTAL JADAD SCORE.		

There are several reasons for this observation. Carrying orthopedic RCTs out in consecutive stages similar to drug trials may prove logistically impossible. Orthopedic practices consecutively enroll patients, making choosing endpoints, sample size, and effect size very difficult, increasing the probability of higher numbers of false positive and false negative findings. Patients are also reluctant to sign up for risky and certainly irreversible surgeries if the orthopedic surgeon has no clear recommendation on the best treatment. Trials may also be underpowered with high false-negative findings because of significantly lower recruitment than drug trials, possibly restricting funding access. Patients may also resist randomization to a nonsurgical arm when they have already failed conservative care and were referred for a surgical consultation. Pre-inclusion and expectation bias may reduce the likelihood of demonstrating benefit in additional nonsurgical care. Cross-over is an incredibly difficult to manage problem when comparing surgery with non-operative care. As mentioned earlier, the SPORT study on surgical vs. non-operative treatment of many common degenerative lumbar spine problems had several flaws. In the study on lumbar disc herniation, 60% of patients were assigned to surgery had a discectomy at two years compared with 45% of patients initially randomized to non-operative treatment.<sup>17</sup> Thus, intention-to-treat analyses were much obscured, leaving the investigators with an observational rather than interventional analysis.

The most significant drawback to orthopedic RCTs is the inability to blind. While blinding reduces bias, as demonstrated by a recent systematic review of 250 RCTs,<sup>51</sup> which found considerable differences in treatment effects between double-blinded trials compared to open-label

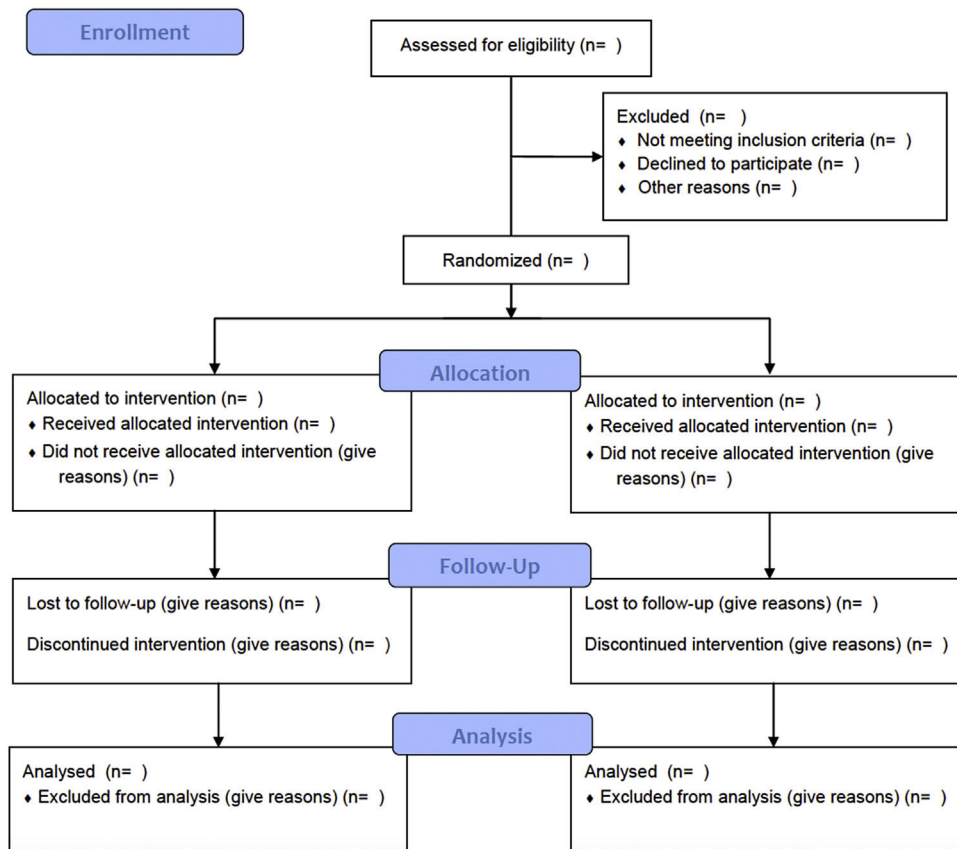
trials. It is nearly impossible in surgical trials since sham-controlled interventions are rarely feasible. At a minimum, the surgeon always knows what operation was performed. Placebo and nocebo effects between surgical and interventional treatments may vary, making the common use of patient self-reported outcome measures (PROMS) such as the visual analog score (VAS),<sup>52</sup> the Oswestry Disability Index (ODI),<sup>53,54</sup> and the modified Macnab criteria<sup>55</sup> particularly problematic due to their subjective nature. The minimally clinical important difference of these PROMs is procedure specific.<sup>56</sup> Short follow-up or different follow-up protocols may potentiate the latter problem.<sup>57</sup>

### Reducing bias by propensity scoring

Underestimation or overestimation of the true effects of orthopedic interventions is common when bias in RCTs is not adjusted for. The risk of bias (RoB) that is avoidable in trials has been assessed for surgical versus non-surgical RCTs.<sup>58</sup> The risk of bias should be diminished. On the patients' side, the risk of bias could be reduced by concealing the size of the incision by applying the same size dressing regardless of treatment, or by using similar postoperative rehabilitation protocols.<sup>59</sup> Further, an expertise-based setup might be used with multiple orthopedic surgeons performing the same procedure for the same group.<sup>58</sup> If it is not possible to blind key individuals, computer randomization and the use of a blinded secondary team of surgeons or surgical nurses is helpful when attempting to lower RoB with unblinded assessment being the most important factor when assessing the susceptibility of an outcome to bias.



## CONSORT 2010 Flow Diagram



**Figure 4** CONSORT flow diagram designed to standardize reporting criteria to improve the quality of clinical trials.

When randomization is impossible, propensity matching may reduce bias, particularly selection bias.<sup>60</sup> In observational studies, the treated and non-treated control groups could have considerable differences in covariates, leading to biased estimates of the treatment effects even when adjusted for by traditional covariance analysis. A propensity score is a conditional probability of being treated given the covariates. It can be used to reduce bias by balancing the covariates in the treatment groups. Estimating the propensity score requires modeling the distribution of the treatment indicator variable concerning the observed covariates. Such propensity scoring can reduce bias through matching, stratification, regression adjustment, or combinations of all three.

### Potential solutions

Given the listed limitations of randomized trials and the apparent ceiling effect of high-grade clinical evidence

development in orthopedic and trauma surgery, several workarounds seem practical to arrive at clinically meaningful outcomes assessments that translate into improved orthopedic patient care protocols. For example, pseudorandomization may be attempted by illustrating differences in preferred treatments between centers.<sup>61</sup> Statistical power may be enhanced by concurrent prospective data collection in separate registries at different institutions and by orchestrating parallel cohort studies. In addition, the separation, as mentioned earlier, of the surgical team from other study team administrators, outcome evaluators, and data analysts may mitigate the lack of blinding between patient and surgeon. One potential downside to RCTs is their limited generalizability because strict inclusion and exclusion criteria distort the incidence of clinical problems in routine practice. Cohort studies are better at capturing the clinical effectiveness of intervention in everyday patients. Others have proposed registration similar to RCTs with disclosure of funding for observational studies<sup>41</sup> before data is retrieved.



**Table 5** The STROBE checklist of items that should be addressed in reports of observational studies (Reproduced from Elm et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies) <sup>62</sup>

	Item	Recommendation	Reported on page
Title and abstract	1	(a) Indicate the study’s design with a commonly used term in the title or the abstract (b) Provide in the abstract an informative and balanced summary of what was done and what was found	
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	
Objectives	3	State specific objectives, including any prespecified hypotheses	
Methods			
Study design	4	Present key elements of study design early in the paper	
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	
Participants	6	(a) Cohort study—give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up Case-control study—give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls Cross-sectional study—give the eligibility criteria, and the sources and methods of selection of participants (b) Cohort study—for matched studies, give matching criteria and number of exposed and unexposed Case-control study—for matched studies, give matching criteria and the number of controls per case	
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	
Data sources/measurement	8*	For each variable of interest give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	
Bias	9	Describe any efforts to address potential sources of bias	
Study size	10	Explain how the study size was arrived at	
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen, and why	
Statistical methods	12	(a) Describe all statistical methods, including those used to control for confounding (b) Describe any methods used to examine subgroups and interactions (c) Explain how missing data were addressed (d) Cohort study—if applicable, explain how loss to follow-up was addressed Case-control study—if applicable, explain how matching of cases and controls was addressed Cross-sectional study—if applicable, describe analytical methods taking account of sampling strategy (e) Describe any sensitivity analyses	
Results			
Participants	13*	(a) Report the numbers of individuals at each stage of the study—e.g., numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analyzed (b) Give reasons for non-participation at each stage (c) Consider use of a flow diagram	
Descriptive data	14*	(a) Give characteristics of study participants (e.g., demographic, clinical, social) and information on exposures and potential confounders (b) Indicate the number of participants with missing data for each variable of interest (c) Cohort study—summarize follow-up time (e.g., average and total amount)	

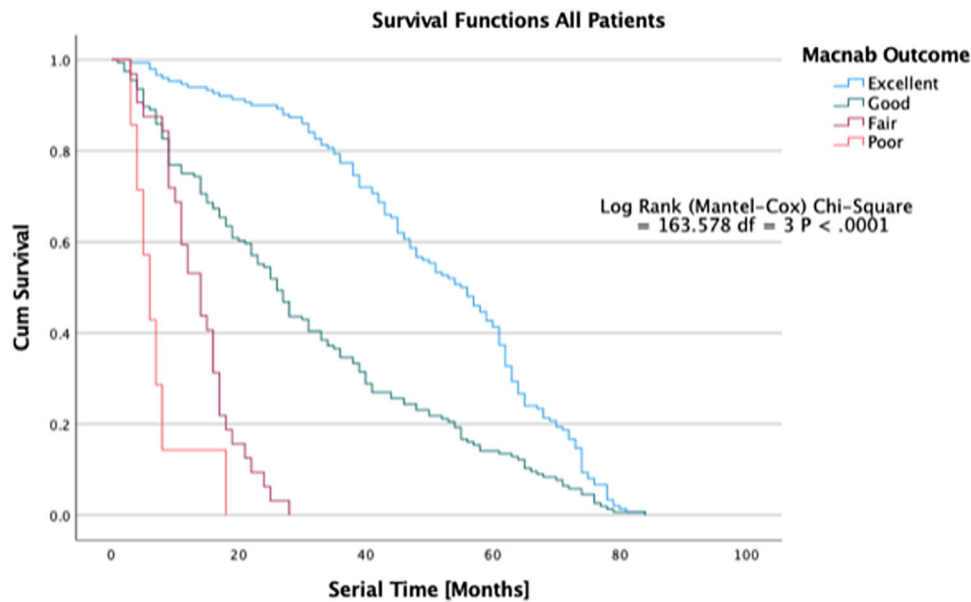
Table 5 (Continued)

	Item	Recommendation	Reported on page
Outcome data	15*	Cohort study—report numbers of outcome events or summary measures over time Case-control study—report numbers in each exposure category, or summary measures of exposure Cross-sectional study—report numbers of outcome events or summary measures	
Main results	16	(a) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (e.g., 95% confidence interval). Make clear which confounders were adjusted for and why they were included (b) Report category boundaries when continuous variables were categorized (c) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	
Other analyses	17	Report other analyses done—e.g., analyses of subgroups and interactions, and sensitivity analyses	
Discussion			
Key results	18	Summarize key results with reference to study objectives	
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	
Generalizability	21	Discuss the generalizability (external validity) of the study results	
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	

While this approach could limit and identify explorative data analyses, this increased bureaucracy of clinical research may also represent a significant hurdle for the individual surgeon clinical investigator who lacks funding and support. If implemented, such a rule would likely favor institutionalized researchers with access to such resources and marginalize the entrepreneurial clinical investigator in private practice or primary care settings. Centralized and institutionalized clinical studies in orthopedic- and trauma surgery may miss the mark on what is playing out in everyday patients because these tertiary care centers often have highly specialized niche practices. Instead of registration requirements, many journals rely on the STROBE checklist for observational studies,<sup>62</sup> whose requirements primarily deal with transparency of reporting (Table 5). Prospective authors are encouraged to give the requested information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies with separate versions of the checklist for cohort, case-control, and cross-sectional studies available on the STROBE website. An explanation and elaboration article published by Elm et al. in the *Lancet* in 2007 discusses each checklist item and gives methodological background and published examples of transparent reporting.<sup>62</sup>

Another way to demonstrate differences in treatment effects between orthopedic treatment study groups is to illustrate the durability of the treatment effect over time. Superior treatments may last much longer than inferior treatments as defined by the need for additional treatment

for the same condition – the serial time in Kaplan-Meier (K-M) curve construction of survival time probabilities listed in survival tables where survival times are sorted in an ascending manner beginning with the shortest serial times for each treatment group.<sup>63</sup> Patients are censored from the study if they drop out, are lost in follow-up, or in whom required data is not available. The cumulative probability of surviving (continued treatment benefit without additional treatment needed for the same condition) excluding censored events is plotted on the Y-axis of the K-M plot, allowing for the analysis of patient treatment intervals of varying duration. The difference between survival curves for the various treatment groups can be quantified for statistical significance using the log-rank test calculating the Chi-square ( $\chi^2$ ) for each event time in the treatment arms. The summed results for each group are typically added to derive the ultimate Chi-square to compare the full K-M curves between groups. The confidence intervals (95%) for the Likelihood ratios can be calculated using the “log method” according to Altman et al.,<sup>64</sup> While K-M curves do not represent the prognosis of the outcome – in fact, they deteriorate in accuracy with every patient being censored from the analysis – they provide an easy-to-understand graphic depiction of postoperative service utilization on an everyday patient level between different orthopedic treatments and help manage patients’ expectations regarding reoperations and functional outcomes in general. They are also helpful in illustrating to decision-makers on all sides of the health-care equation the clinical benefit of various treatments for



**Figure 5** Kaplan-Meier (K-M) Survival time in patients treated endoscopic, laminectomy decompression, and transforaminal lumbar interbody fusion treated for lumbar lateral canal and foraminal stenosis. The graphic depiction of the survival times illustrates to durability of the surgical treatment between patients with excellent, good, fair, and poor Macnab outcomes. The difference in clinical benefit between the three treatments in comparison to a control group have recently been published in a prospective cohort study of 412 patients. Lewandrowski et al., *J Personalized Med*, 2022; 12 (7).

common degenerative and traumatic orthopedic conditions, all aimed at reducing the disease burden in the Latin American populations with cost-effectiveness and good stewardship of the limited healthcare resources in mind.

### Example spinal endoscopy

The claim that targeted endoscopic treatment of common spinal pain generators produces higher patient satisfaction than traditional spine surgery used to lack support from statistically validated prognosticators of favorable outcomes with the endoscopic procedure for most surgical indications. In several previously published articles, the difficulties of orchestrating meaningful clinical outcome research studies were discussed.<sup>65–70</sup> Some of them were overcome with straightforward survival analysis of the benefit of endoscopic spinal surgery. Endoscopic spine surgery may be questioned for its medical necessity and effectiveness compared to other forms of spinal surgery. The lack of clinical evidence to warrant endoscopic decompression surgery of the spine is frequently called out. Escalating costs create pushback by payers and patients. Many outcome studies in endoscopic spine surgery are level III retrospective case series and cohort studies published by pioneers of the procedure, including several senior authors of this article.<sup>65–70</sup> Level I and II prospective randomized trials are few and far between.<sup>71–76</sup> In 2013, Birkenmeier et al. performed a meta-analysis of comparative controlled clinical trials on endoscopic and microsurgical standard procedures.<sup>77</sup> Two lumbar randomized controlled trials (RCTs)<sup>78,79</sup> and one controlled study (CS)<sup>80</sup> were finally identified as eligible for evaluation. Clinical outcomes were similar between the endoscopic and the microsurgical methods in these

trials. The endoscopic techniques were associated with fewer complications and a lower rate of revision surgeries requiring arthrodesis at a statistically significant level. In 2018, Kong et al. prospectively randomized patients (n = 40) with lumbar disc herniation and lateral recess stenosis to endoscopic lumbar discectomy or microsurgical laminotomy showing similar outcomes.<sup>81</sup> Another open-label randomized trial compared percutaneous transforaminal endoscopic discectomy to mid-line microendoscopic discectomy.<sup>82</sup> Outcome analysis of ODI, VAS for back and leg pain, Medical Outcomes Study 36-Item Short-Form Health Survey bodily pain and physical function (SF-36), and EuroQol Group’s EQ-5D were similar at one-year follow-up.

Does the question arise whether these traditional outcome tools, including the visual analog (VAS) leg and back pain score,<sup>52</sup> the Oswestry disability index (ODI),<sup>53,54</sup> Roland Morris score,<sup>83</sup> the short-form (SF) SF-12 and SF-36<sup>84,85</sup> are sensitive enough to detect the factors impacting patient satisfaction and clinical outcomes with endoscopic surgery? These traditional outcome tools do not adequately reflect the more favorable patient perception with the endoscopic surgery when compared to the types of translaminar surgeries it is trying to replace. They also do not measure the benefits of staging treatment options by targeting the predominant pain generator as opposed to treating all sources of pain currently suggested by imaged-based criteria.

Ultimately, the question of whether the endoscopic surgical treatment is more beneficial than other types of minimally invasive or open traditional spine surgery is complicated but can, at least in part can, be answered by examining the durability of the treatment benefit, defined as continued self-reporting of favorable PROM outcomes by patients without the utilization of any other service to manage the underlying degenerative disease process. A recent

article published by some of the authors in the Journal of Personalized Medicine illustrates the clinical outcomes of the protocol of treating predominant lumbar pain generators versus image-based criteria for traditional spine surgery employing K-M survival curves (Figure 5).<sup>67</sup> While this type of analysis disregards many covariates, it provides a graphic representation of postoperative recovery and utilization dynamic, allowing innovative surgeons to perform superiority studies which in the fast-moving field of orthopedic and trauma surgery may be the best way to stay ahead of technology advances before the clinical equipoise is lost.

## Conclusions

Outcome research in orthopaedic- and trauma surgery is hampered by randomization problems similar to many other surgical subspecialties. Cross-over is one major problem that limits well-intended RCTs to single-center trials similar to Phase II drug trials. This fast-moving field may quickly lose clinical equipoise. A glass ceiling effect does exist in surgical clinical trials where well-designed prospective case-controlled and cohort series may produce better clinical evidence than poorly executed RCTs. These seemingly lower grade clinical evidence studies may be of higher value to the every-day patient seeking the highest quality orthopedic care. This conclusion is relevant to orthopedic surgeons in Latin America in particular since limited time and resources to support meaningful outcome research are commonplace in the cash-strapped public health care systems. Durability analysis of competing treatments and technology advances is an easy-to-understand visual way to communicate treatment benefits to patients, surgeons, and payors.

## Disclaimer

The views expressed in this article represent those of the authors and no other entity or organization. The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

## Institutional Review Board Statement

In accordance with the Declaration of Helsinki, this editorial did not require approval by a Institutional Review Board.

## Informed Consent Statement

Informed consent was obtained from all subjects involved in the study.

## Data Availability Statement

The data presented in this study are available on request from the corresponding author.

## Funding

No funding was received to sponsor this work.

## Conflict of Interest

None. This manuscript is not meant for or intended to endorse any products or push any other agenda other than the clinical utilization data associated with the presented research. There were no funders involved in the design or conduction of this study, collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results. The authors declare no conflict of interest, and there was no personal circumstance or interest that may be perceived as inappropriately influencing the representation or interpretation of reported research results. This research was not compiled to enrich anyone. It was merely intended to highlight the common problems encountered by clinical investigators in Orthopedic Surgery and Trauma.

## Acknowledgments

None.

## References

1. Safiri S, Kolahi AA, Smith E, et al. Global, regional and national burden of osteoarthritis 1990-2017: a systematic analysis of the Global Burden of Disease Study 2017. *Ann Rheum Dis*. 2020;79:819–28.
2. Coury HJ. Time trends in ergonomic intervention research for improved musculoskeletal health and comfort in Latin America. *Appl Ergon*. 2005;36:249–52.
3. Rockhold FW. Industry perspectives on ICH guidelines. *Stat Med*. 2002;21:2949–57.
4. Malik AY, Foster C. The revised Declaration of Helsinki: cosmetic or real change? *J R Soc Med*. 2016;109:184–9.
5. Shrestha BM. The Declaration of Helsinki in relation to medical research: historical and current perspectives. *J Nepal Health Res Counc*. 2012;10:254–7.
6. Riis P. Thirty years of bioethics: the Helsinki Declaration 1964-2003. *New Rev Bioeth*. 2003;1:15–25.
7. Declaration of Helsinki. Recommendations guiding doctors in clinical research. Adopted by the World Medical Association in 1964. *Wis Med J* 1967;66:25-6.
8. Murano G. International Conference on Harmonization-critical discussion of the biotech'Specifications' document. *Curr Opin Biotechnol*. 2000;11:303–8.
9. The periodic health examination. Canadian Task Force on the Periodic Health Examination. *Can Med Assoc J* 1979;121:1193-254.
10. Sackett DL. Rules of evidence and clinical recommendations on the use of antithrombotic agents. *Chest*. 1989;95:2s–4s.
11. Burns PB, Rohrich RJ, Chung KC. The levels of evidence and their role in evidence-based medicine. *Plast Reconstr Surg*. 2011;128:305–10.
12. Ariyan S, Martin J, Lal A, et al. Antibiotic prophylaxis for preventing surgical-site infection in plastic surgery: an evidence-based consensus conference statement from the American Association of Plastic Surgeons. *Plast Reconstr Surg*. 2015;135:1723–39.
13. Sullivan D, Chung KC, Eaves FF3rd, et al. The Level of Evidence Pyramid: Indicating Levels of Evidence in Plastic and Reconstructive Surgery Articles. *Plast Reconstr Surg*. 2021;148:68s–71s.
14. Perdakis G, Dillingham C, Boukoulas S, et al. American Society of Plastic Surgeons Evidence-Based Clinical Practice

- Guideline Revision: Reduction Mammoplasty. *Plast Reconstr Surg.* 2022;149:392e–409e.
15. Centre for Evidence Based Medicine (CEBM). 2013.
  16. Solheim O. Randomized controlled trials in surgery and the glass ceiling effect. *Acta Neurochir (Wien).* 2019;161:623–5.
  17. 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;296:2441–50.
  18. 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 (Phila Pa 1976).* 2018;43:1619–30.
  19. From the American Society of Plastic Surgeons. Evidence-based clinical practice guidelines.
  20. Watters W, 3rd, Rethman MP, Hanson NB, et al. Prevention of Orthopaedic Implant Infection in Patients Undergoing Dental Procedures. *J Am Acad Orthop Surg.* 2013;21:180–9.
  21. Meiyappan KP, Cote MP, Bozic KJ, et al. Adherence to the American Academy of Orthopaedic Surgeons Clinical Practice Guidelines for Nonoperative Management of Knee Osteoarthritis. *J Arthroplasty.* 2020;35:347–52.
  22. Bono CM, Watters WC3rd, Heggeness MH, et al. An evidence-based clinical guideline for the use of antithrombotic therapies in spine surgery. *Spine J.* 2009;9:1046–51.
  23. Bono CM, Ghiselli G, Gilbert TJ, et al. An evidence-based clinical guideline for the diagnosis and treatment of cervical radiculopathy from degenerative disorders. *Spine J.* 2011;11:64–72.
  24. Kreiner DS, Shaffer WO, Baisden JL, et al. An evidence-based clinical guideline for the diagnosis and treatment of degenerative lumbar spinal stenosis (update). *Spine J.* 2013;13:734–43.
  25. Shaffer WO, Baisden JL, Fernand R, et al. An evidence-based clinical guideline for antibiotic prophylaxis in spine surgery. *Spine J.* 2013;13:1387–92.
  26. Kaiser MG, Eck JC, Groff MW, et al. Guideline update for the performance of fusion procedures for degenerative disease of the lumbar spine. Part 1: introduction and methodology. *J Neurosurg Spine.* 2014;21:2–6.
  27. Kreiner DS, Hwang SW, Easa JE, et al. An evidence-based clinical guideline for the diagnosis and treatment of lumbar disc herniation with radiculopathy. *Spine J.* 2014;14:180–91.
  28. Kreiner DS, Baisden J, Mazanec DJ, et al. Guideline summary review: an evidence-based clinical guideline for the diagnosis and treatment of adult isthmic spondylolisthesis. *Spine J.* 2016;16:1478–85.
  29. Matz PG, Meagher RJ, Lamer T, et al. Guideline summary review: An evidence-based clinical guideline for the diagnosis and treatment of degenerative lumbar spondylolisthesis. *Spine J.* 2016;16:439–48.
  30. Tan SY, Linskey K. Christiaan Barnard (1922-2001): First heart transplant surgeon. *Singapore Med J.* 2019;60:495–6.
  31. Gawande A. Two Hundred Years of Surgery. *New England Journal of Medicine.* 2012;366:1716–23.
  32. Moseley JB, O'Malley K, Petersen NJ, et al. A Controlled Trial of Arthroscopic Surgery for Osteoarthritis of the Knee. *New England Journal of Medicine.* 2002;347:81–8.
  33. Sihvonen R, Paavola M, Malmivaara A, et al. Arthroscopic Partial Meniscectomy versus Sham Surgery for a Degenerative Meniscal Tear. *New England Journal of Medicine.* 2013;369:2515–24.
  34. Buchbinder R, Osborne RH, Ebeling PR, et al. A randomized trial of vertebroplasty for painful osteoporotic vertebral fractures. *N Engl J Med.* 2009;361:557–68.
  35. 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:k2860.
  36. Probst P, Grummich K, Harnoss JC, et al. Placebo-Controlled Trials in Surgery: A Systematic Review and Meta-Analysis. *Medicine.* 2016;95:e3516.
  37. Weinstein JN, Tosteson TD, Lurie JD, et al. Surgical versus nonsurgical therapy for lumbar spinal stenosis. *N Engl J Med.* 2008;358:794–810.
  38. Weinstein JN, Lurie JD, Tosteson TD, et al. Surgical versus nonoperative treatment for lumbar disc herniation: four-year results for the Spine Patient Outcomes Research Trial (SPORT). *Spine (Phila Pa 1976).* 2008;33:2789–800.
  39. Pearson AM, Blood EA, Frymoyer JW, et al. SPORT lumbar intervertebral disk herniation and back pain: does treatment, location, or morphology matter? *Spine (Phila Pa 1976).* 2008;33:428–35.
  40. Weinstein JN, Lurie JD, Tosteson TD, et al. Surgical versus non-surgical treatment for lumbar degenerative spondylolisthesis. *N Engl J Med.* 2007;356:2257–70.
  41. Martin E, Muskens IS, Senders JT, et al. Randomized controlled trials comparing surgery to non-operative management in neurosurgery: a systematic review. *Acta Neurochirurgica.* 2019;161:627–34.
  42. Hempel S, Miles SMJNV, et al. Empirical Evidence of Associations Between Trial Quality and Effect Size [Internet]. Agency for Healthcare Research and Quality (US); 2011 Jun. [Table]. *Jadad Scale.* 2022. Available at: <https://www.ncbi.nlm.nih.gov/books/NBK56923/table/appendixes.app6.t1/>.
  43. Augestad KM, Berntsen G, Lassen K, et al. Standards for reporting randomized controlled trials in medical informatics: a systematic review of CONSORT adherence in RCTs on clinical decision support. *J Am Med Inform Assoc.* 2012;19:13–21.
  44. Moher D, Schulz KF, Altman D. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomized trials. *Jama.* 2001;285:1987–91.
  45. Moher D, Schulz KF, Altman DG. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomised trials. *Lancet.* 2001;357:1191–4.
  46. Moher D, Schulz KF, Altman DG. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomized trials. *Ann Intern Med.* 2001;134:657–62.
  47. Lewandrowski KU, de Carvalho PST, Calderaro AL, et al. Outcomes with transforaminal endoscopic versus percutaneous laser decompression for contained lumbar herniated disc: a survival analysis of treatment benefit. *J Spine Surg.* 2020;6: S99-S84.
  48. Lewandrowski KU, Dowling A, de Carvalho P, et al. Indication And Contraindication Of Endoscopic Transforaminal Lumbar Decompression. *World Neurosurg.* 2020.
  49. Lewandrowski KU, Yeung A. Meaningful outcome research to validate endoscopic treatment of common lumbar pain generators with durability analysis. *J Spine Surg.* 2020;6:S6–13.
  50. Lewandrowski KU, Tieber F, Hellinger S, et al. Durability of Endoscopes Used During Routine Lumbar Endoscopy: An Analysis of Use Patterns, Common Failure Modes. Impact on Patient Care, and Contingency Plans. *Int J Spine Surg.* 2021;15:1147–60.
  51. Schulz KF, Chalmers I, Hayes RJ, et al. Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *Jama.* 1995;273:408–12.
  52. Reed CC, Wolf WA, Cotton CC, et al. A visual analogue scale and a Likert scale are simple and responsive tools for assessing dysphagia in eosinophilic oesophagitis. *Aliment Pharmacol Ther.* 2017;45:1443–8.
  53. Fairbank J. Use of Oswestry Disability Index (ODI). *Spine (Phila Pa 1976).* 1995;20:1535–7.

54. Fairbank JC, Pynsent PB. The Oswestry Disability Index. *Spine (Phila Pa 1976)*. 2000;25:2940–52, discussion 52.
55. Macnab I. The surgery of lumbar disc degeneration. *Surg Annu*. 1976;8:447–80.
56. Lewandrowski KUPSTDECPDEC, et al. Minimal Clinically Important Difference in Patient-Reported Outcome Measures with the Transforaminal Endoscopic Decompression for Lateral Recess and Foraminal Stenosis. *Int J Spine Surg*. 2020;14:254–66.
57. Denis F, Basch E, Septans AL, et al. Two-Year Survival Comparing Web-Based Symptom Monitoring vs Routine Surveillance Following Treatment for Lung Cancer. *Jama*. 2019;321:306–7.
58. Barcot O, Boric M, Dosenovic S, et al. Adequacy of risk of bias assessment in surgical vs non-surgical trials in Cochrane reviews: a methodological study. *BMC Medical Research Methodology*. 2020;20:240.
59. Gurusamy KS, Gluud C, Nikolova D, et al. Assessment of risk of bias in randomized clinical trials in surgery. *British Journal of Surgery*. 2009;96:342–9.
60. Nerland US, Jakola AS, Solheim O, et al. Minimally invasive decompression versus open laminectomy for central stenosis of the lumbar spine: pragmatic comparative effectiveness study. *Bmj*. 2015;350:h1603.
61. Jakola AS, Skjulsvik AJ, Myrmet KS, et al. Surgical resection versus watchful waiting in low-grade gliomas. *Ann Oncol*. 2017;28:1942–8.
62. von Elm E, Altman DG, Egger M, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *The Lancet*. 2007;370:1453–7.
63. Kaplan E, Meier P. Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association*. 1958;53:457–81.
64. Altman D, Machin D, Bryant T, et al. *Statistics with confidence: confidence intervals and statistical guidelines*. John Wiley & Sons; 2013.
65. Yeung A, Lewandrowski KU. Five-year clinical outcomes with endoscopic transforaminal foraminoplasty for symptomatic degenerative conditions of the lumbar spine: a comparative study of inside-out versus outside-in techniques. *J Spine Surg*. 2020;6:S66–83.
66. Yeung A, Lewandrowski KU. Early and staged endoscopic management of common pain generators in the spine. *J Spine Surg*. 2020;6:S1–5.
67. Lewandrowski K-U, Abraham I, Ramírez León JF, et al. A Proposed Personalized Spine Care Protocol (SpineScreen) to Treat Visualized Pain Generators: An Illustrative Study Comparing Clinical Outcomes and Postoperative Reoperations between Targeted Endoscopic Lumbar Decompression Surgery, Minimally Invasive TLIF and Open Laminectomy. *Journal of Personalized Medicine*. 2022;12:1065.
68. Ramirez Leon JF, Ardila AS, Rugeles Ortiz JG, et al. Standalone lordotic endoscopic wedge lumbar interbody fusion (LEW-LIF) with a threaded cylindrical peek cage: report of two cases. *J Spine Surg*. 2020;6. S275-S84.
69. Ramirez Leon JF. The motivators to endoscopic spine surgery implementation in Latin America. *J Spine Surg*. 2020;6:S45–8.
70. Ramirez LF, Thisted R. Complications and demographic characteristics of patients undergoing lumbar discectomy in community hospitals. *Neurosurgery*. 1989;25, 226-230 discussion 30-1.
71. Ruetten S, Komp M, Merk H, et al. Full-endoscopic interlaminar and transforaminal lumbar discectomy versus conventional microsurgical technique: a prospective, randomized, controlled study. *Spine (Phila Pa 1976)*. 2008;33:931–9.
72. Osterman H, Seitsalo S, Karppinen J, et al. Effectiveness of microdiscectomy for lumbar disc herniation: a randomized controlled trial with 2 years of follow-up. *Spine (Phila Pa 1976)*. 2006;31:2409–14.
73. Knight MT, Goswami A, Patko JT, et al. Endoscopic foraminoplasty: a prospective study on 250 consecutive patients with independent evaluation. *J Clin Laser Med Surg*. 2001;19:73–81.
74. Sampath P, Bendebba M, Davis JD, et al. Outcome in patients with cervical radiculopathy Prospective, multicenter study with independent clinical review. *Spine (Phila Pa 1976)*. 1999;24:591–7.
75. Hermantin FU, Peters T, Quartararo L, et al. A prospective, randomized study comparing the results of open discectomy with those of video-assisted arthroscopic microdiscectomy. *J Bone Joint Surg Am*. 1999;81:958–65.
76. Choy DS. Percutaneous laser disc decompression (PLDD): twelve years' experience with 752 procedures in 518 patients. *J Clin Laser Med Surg*. 1998;16:325–31.
77. Birkenmaier C, Komp M, Leu HF, et al. The current state of endoscopic disc surgery: review of controlled studies comparing full-endoscopic procedures for disc herniations to standard procedures. *Pain Physician*. 2013;16:335–44.
78. Ruetten S, Komp M, Merk H, et al. Full-endoscopic cervical posterior foraminotomy for the operation of lateral disc herniations using 5.9-mm endoscopes: a prospective, randomized, controlled study. *Spine (Phila Pa 1976)*. 2008;33:940–8.
79. Ruetten S, Komp M, Merk H, et al. Recurrent lumbar disc herniation after conventional discectomy: a prospective, randomized study comparing full-endoscopic interlaminar and transforaminal versus microsurgical revision. *J Spinal Disord Tech*. 2009;22:122–9.
80. Lee DY, Shim CS, Ahn Y, et al. Comparison of percutaneous endoscopic lumbar discectomy and open lumbar microdiscectomy for recurrent disc herniation. *J Korean Neurosurg Soc*. 2009;46:515–21.
81. Kong L, Shang XF, Zhang WZ, et al. Percutaneous endoscopic lumbar discectomy and microsurgical laminotomy: A prospective, randomized controlled trial of patients with lumbar disc herniation and lateral recess stenosis. *Orthopade*. 2019;48:157–64.
82. Rong L. Erratum. Percutaneous transforaminal endoscopic discectomy compared with microendoscopic discectomy for lumbar disc herniation: 1-year results of an ongoing randomized controlled trial. *J Neurosurg Spine*. 2019:1–3.
83. Chiarotto A, Maxwell LJ, Terwee CB, et al. Roland-Morris Disability Questionnaire and Oswestry Disability Index: Which Has Better Measurement Properties for Measuring Physical Functioning in Nonspecific Low Back Pain? Systematic Review and Meta-Analysis. *Phys Ther*. 2016;96:1620–37.
84. Barber MD, Chen Z, Lukacz E, et al. Further validation of the short form versions of the Pelvic Floor Distress Inventory (PFDI) and Pelvic Floor Impact Questionnaire (PFIQ). *Neurourol Urodyn*. 2011;30:541–6.
85. Jenkinson C, Layte R. Development and testing of the UK SF-12 (short form health survey). *J Health Serv Res Policy*. 1997;2:14–8.