J Shoulder Elbow Surg (2019) 28, 2409–2417

Journal of Shoulder and Elbow Surgery

www.elsevier.com/locate/ymse

BASIC SCIENCE

The fragility of findings of randomized controlled trials in shoulder and elbow surgery

Joseph J. Ruzbarsky, MD[a,*], Ryan C. Rauck, MD[a], Joseph Manzi, BS[b], Sariah Khormaee, MD, PhD[a], Bridget Jivanelli, MLIS[a], Russell F. Warren, MD[a]

[a]*Department of Orthopaedics, Hospital for Special Surgery, New York, NY, USA*
[b]*Weill Cornell School of Medicine, New York, NY, USA*

Background: Considered the gold standard of study designs, randomized controlled trials' (RCTs) results shape clinical practice, effect policy, and influence reimbursement. The fragility index (FI) can be used to quantitate the relative robustness of RCT results, with higher scores indicating more stout results. Unfortunately, most RCTs in surgery have fragile results. The aim of this study was to report on the FI in addition to a qualitative assessment of recent RCTs within the field of shoulder and elbow surgery.

Methods: A systematic review was performed identifying recently published shoulder/elbow RCTs that included 1:1 allocated parallel study arms, dichotomous primary outcome variables, and statistical significance. The FI was calculated by sequentially modifying outcome groups by exchanging a nonevent in one group to an event until the *P* value for the outcome comparison, as calculated by the Fisher exact test, was increased above the .05 threshold.

Results: Thirty RCTs were included. The median FI was 4. Sixty percent trials had a FI of 2 or less. Fifty-three percent studies reported that participants were lost to follow-up. In 87.5% of these studies, the losses to follow-up exceeded their respective FIs. Only 53% of studies defined a primary outcome variable and 60% studies performed a prestudy power analysis.

Conclusions: The median FI reported in the recent shoulder/elbow literature is 4; however, a high proportion of included RCTs display significant methodological concerns. The FI is a useful adjunct to analyze RCT results, but careful analysis of trial methods should be employed in each circumstance before drawing conclusions.

Level of evidence: Basic Science Study; Statistics and Measurement Error

Keywords: Fragility index; fragility quotient; evidence-based medicine; randomized control trial; loss to follow-up; shoulder and elbow surgery

The fragility index (FI) is a metric that is gaining popularity throughout clinical medicine to evaluate the robustness, or lack thereof, of the results reported in randomized controlled trials (RCTs).[9,16,17,35,37-39,44] Its use is limited to scenarios comparing 2 outcome groups of dichotomous variables reporting a statistically significant (typically with

No IRB approval was required for this basic science study.

*Reprint requests: Joseph J. Ruzbarsky, MD, Hospital for Special Surgery, 535 E. 70th St, New York, NY 10021, USA.

E-mail address: ruzbarskyj@hss.edu (J.J. Ruzbarsky).

1058-2746/$ - see front matter
https://doi.org/10.1016/j.jse.2019.04.051

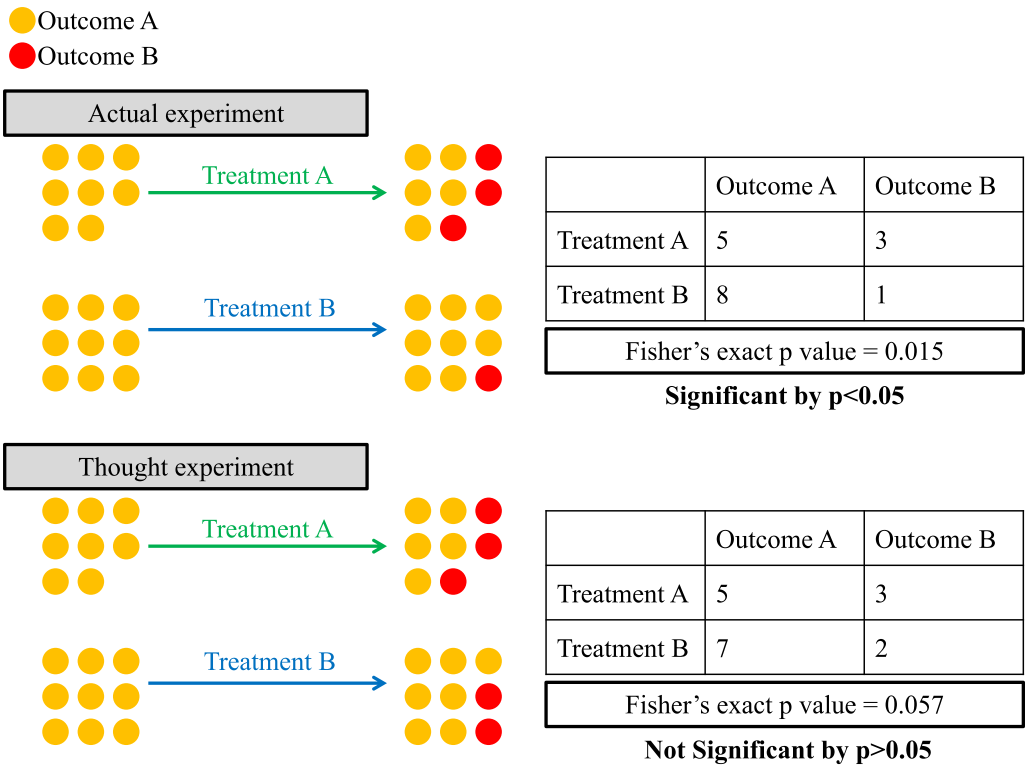

Figure 1 Graphical representation of the fragility index.

an α value of <0.05) result. An example would be a study that compares 2 operative techniques where the primary outcome variable of interest is the development of a nonunion (a dichotomous result of yes or no). The FI is calculated by sequentially modifying outcome groups by exchanging a nonevent in one group to an event in the other until the *P* value for the comparison, as calculated by the Fisher exact test, is increased above the .05 threshold. The higher the score the more stout the results.

The FI can be reported on a result-to-result basis, in conjunction with the *P* value and confidence intervals (CI) as a means of evaluating robustness of individual RCT's conclusions. To better clarify its meaning, consider a clinical trial result with a FI of 20. This value means that 20 patients in one group would have to be converted to the converse outcome group for that specific conclusion to lose its "statistical significance." On the contrary, a FI of 1 would only require a single patient to be changed, obviously a much more fragile result. The FI does not replace a *P* value, nor does it reinforce the *P* value of a "truth statistic"; but rather its integer value represents an easily interpretable metric to gauge the strength or fragility of a result. In addition to its use on a case-by-case basis for an individual trial's results, the FI has also been reported to summarize whole fields of medicine as a proxy for the quality of the RCT results within a subset of the literature. Among the currently published literature, surgical subspecialties have a tendency to report lower FIs than nonsurgical fields.[9,16,17,35,37,39,45] This discrepancy is likely closely related to the magnitude of the RCT and the complexity of the intervention being studied. Logistically, a pharmaceutical trial can be easier to execute than a surgical trial. Thus, some of the existing systematic reviews[17,36] that have reported on the FI for entire subspecialties, have called for multicenter collaboration in an effort to increase RCT size and potentially help in increasing the existing low FIs reported.

To date, no reports exist examining the FI within the field of shoulder and elbow surgery in its entirety. The purpose of this study was to systematically review the RCTs published in the last decade in order to determine the individual FIs of each study's primary outcome variable as well as to summarize the field as a whole. A secondary aim was to assess the overall quality of RCTs published including metrics such as risk of bias, losses to follow-up, and power analyses. We hypothesized that the FI would be consistent with other reports of FI within orthopedic subspecialties,[9,16,17,36] namely in the range of 1-3 and that there would be a high proportion of methodological deficiencies in RCT design.

Materials and methods

A systematic search of the MEDLINE via PubMed, Embase, and Cochrane databases for "shoulder," "elbow," "randomized control trial," and similar keywords from January 1, 2008, to January 1, 2018, was implemented. Supplementary Appendix A

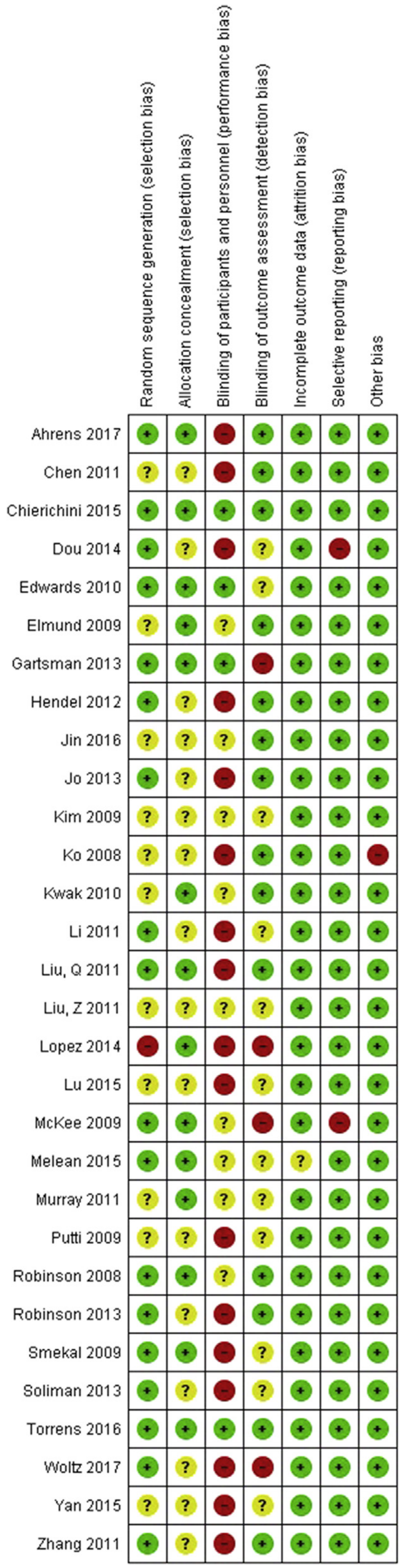

Figure 2 Cochrane risk of bias assessment table.

contains the unabridged search strategy. Covidence (Melbourne, Australia) software was used for organizing the systematic study analysis in addition to the inclusion, exclusion, and tracking of published works. Two independent reviewers screened titles, abstracts, and manuscripts to identify parallel arm RCTs allocating patients in a 1:1 ratio that were published in English, that included a dichotomous primary outcome variable, and that reported statistical significance. A third reviewer resolved any conflicts. Reasons for study exclusion included more than 2 outcome groups, lack of dichotomous variables, improper study design, lack of statistically significant results, secondary outcomes, nonsurgical RCTs, and pediatric shoulder and elbow RCTs. Variables extracted from each study included journal name, publication year, sample size for each group, unaccounted for losses to follow-up for each group, number of events for each group, reported P value, whether or not a prestudy power analysis was performed, and randomization parameters (ie, allocation method, method of concealment, etc.). The FI for each study was calculated as described by Walsh et al[45] using the FI calculator found at http://clincalc.com/Stats/FragilityIndex.aspx. This website allows for entering the proportion of events for each study arm then automates the process of FI calculation by converting one outcome to the other sequentially until the P value, as calculated by Fisher's exact test, increases above the .05 threshold (Fig. 1). The fragility quotient (FQ) was calculated by dividing the FI by the total sample size of the trial, providing a more pure metric of robustness that eliminates the variable of study size. Lower FQ values indicate stronger the results. Each included study was assessed using the Cochrane Risk of Bias Tool (Fig. 2). This systematic review was performed and reported according to Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines.

Results

The search strategy identified 2308 total articles for potential inclusion, of which 436 met inclusion criteria (Fig. 3). Authors, outcome assessed, type of comparison, number of patients enrolled, patients lost to follow-up, prestudy power calculation, FI, FQ, and risk of bias per the Cochrane Risk of Bias Tool are reported in Table I. *The Journal of Shoulder and Elbow Surgery* accounted for the greatest number of publications in this cohort, 8/30 (27%). The rest of the number of publications per journal are reported in Table II.

The mean trial size was 75.5 (range, 31-301) with a mean group size of 37.4 (range, 15-154), and the distribution of FIs ranged from 0 to 17 (Fig. 4). Excluding FIs of 0, the median FI was 4. The median FQ was 0.040 with a range from 0 to 0.239.

There were 9 publications with an FI of zero (Fig. 4), which means that when the Fisher's exact test was applied to the proportion of events to nonevents of both arms of this trial, the resulting P value output was greater than .05, thus precluding the calculation of the FI.

A total of 16 of 30 (53%) studies reported that participants were lost to follow-up with a mean reported loss to follow-up of 5.6 participants. In 14 of 16 (88%) of these studies, the losses to follow-up exceeded their respective FIs.

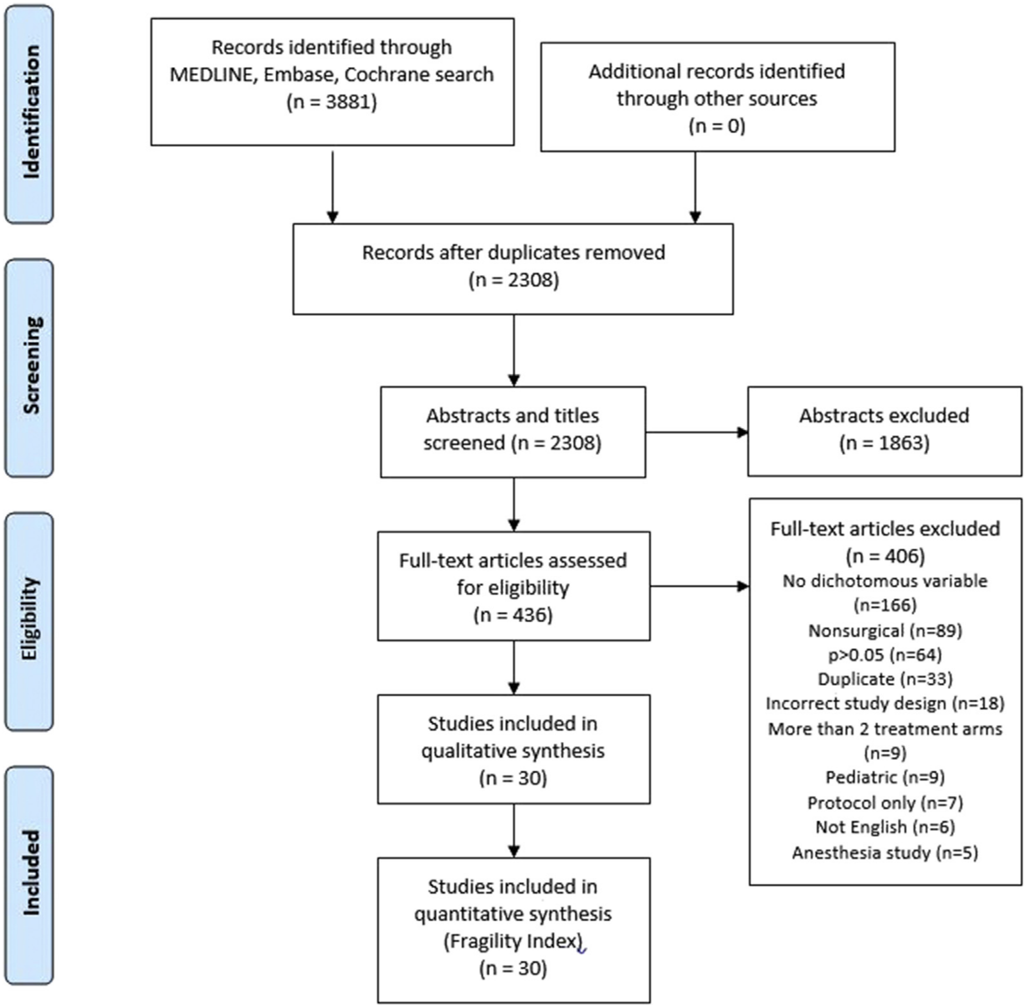

Figure 3 Consort diagram of study exclusion and selection.

A total of 8 of 30 (27%) studies were determined to have a high risk of bias (Fig. 2). Only 16 of 30 (53%) studies defined a primary outcome variable. Of 30 studies, 18 (60%) performed a prestudy power analysis. A total of 19 of 30 (63%) reported their randomization technique, which was most commonly computer generated (10/19).

Discussion

In a systematic search of RCTs published over the prior 10 years on the topic of shoulder and elbow surgery, 30 studies met criteria for FI calculation. Of these studies the median FI was 4, in other words, nullifying the reported statistical significance would require changing 4 patients from one outcome group to the converse. The median value of 4 is higher than any other reported subspecialty FIs within orthopedics,[9,16,17,36] potentially signifying more robust findings and stronger RCTs within the field of shoulder and elbow surgery.

In 2018 alone, *The Journal of Shoulder and Elbow Surgery* published 7 RCTs.[10,13,18,19,31,42,44] Given an increasing emphasis on evidence-based medicine,[2,3] which advocates for higher quality methodologies, clinical research has begun to move away from level IV cases series, with increasing numbers of RCTs performed throughout all orthopedic subspecialties. Although considered the gold standard for research, careful analysis of RCT methodologies and results remains essential so not to disproportionately value the conclusions of such studies over lower level of evidence investigations. Appropriate scrutiny requires attention to both study methods and design details including randomization technique, concealment, blinding, and methods of sample size determination in addition to the results: effect sizes, *P* values, and CIs. These considerations are especially important for the readers of surgical and orthopedic RCTs. Contrary to RCTs involving medications or noninvasive interventions, in most surgical intervention trials the overall number of patients included is often small due to cost and

Table I Included shoulder and elbow surgery trials with characteristics

Authors	Type of comparison	Outcome	Specific outcome assessed	Patients enrolled	Patients lost to follow-up	Risk of bias	Power calculation	Fragility index	Fragility quotient
Ahrens et al[1]	ORIF vs. nonoperative	Primary	Union at 9 mo	302	47	L	Y	7	0.023
McKee et al[28]	ORIF vs. TEA	Primary	Reoperation rate	40	2	U	Y	0	0
Hendel et al[12]	PSI vs. standard technique	Primary	Malposition	31	0	U	Y	4	0.129
Dou et al[6]	Hook plates vs. K-wire tension bands	NS	Recurrence rates of bone fracture or AC joint dislocation	100	0	H	N	1	0.010
Chen et al[4]	Radial head replacement vs. ORIF	NS	Broberg and Morrey scores	45	0	U	N	0	0
Chierichini et al[5]	Norepinephrine vs. epinephrine enriched irrigation fluid	Primary	Hypotensive and bradycardic events	119	0	L	Y	2	0.017
Gartsman et al[11]	Single vs. double row	Primary	Healing of rotator cuff	83	7	L	Y	1	0.012
Jin et al[14]	New depth gauge vs. traditional methods	NS	Screw penetration articular surface	40	0	H	N	1	0.025
Jo et al[15]	RCR with PRP vs. no PRP	Primary	Retear rate	38	10	U	Y	1	0.026
Kwak et al[21]	SCD vs. no SCD	Primary	Hypotension	50	0	U	N	2	0.040
Edwards et al[7]	Keeled vs. pegged glenoid	Primary	Glenoid lucency	45	7	U	Y	0	0
Elmlund et al[8]	PGA vs. PLLA suture anchors	Primary	Radiographic visibility of drill holes	35	5	U	Y	5	0.143
Kim et al[18]	Warmed vs. not warmed IVF	Primary	Hypothermia	46	0	H	Y	11	0.239
Liu Z et al[24]	PHILOS vs. MIIG	NS	Complications	62	NS	H	N	0	0
Liu Q-H et al[23]	Cable pin system vs. tension band	NS	Complications	50	0	L	N	1	0.020
Lopiz et al[25]	Curved vs. straight nail	NS	Symptoms of rotator cuff disease	52	2	U	Y	3	0.058
Lu et al[26]	Double screw vs. conventional tension band	NS	Complications	88	0	H	N	9	0.102
Murray et al[30]	Chlorohexidine vs. none	Primary	Positive culture	100	0	U	Y	7	0.070
Melean et al[29]	Nonoperative vs. ORIF	NS	Union at 12 wk	76	NS	U	Y	17	0.224
Robinson et al, 2013[33]	Nonoperative vs. ORIF	Primary	Nonunion	178	22	U	Y	6	0.034
Robinson et al, 2008[34]	Bankart repair vs. sham arthroscopy	Primary	Recurrent instability	84	4	L	Y	5	0.060
Soliman et al[41]	Biceps tenodesis vs. none	NS	Shoulder pain	37	8	U	Y	0	0
Torrens et al[43]	38 mm vs. 42 mm	Primary	Scapular notching	81	8	L	Y	7	0.086
Woltz et al[46]	ORIF vs. nonoperative	Primary	Secondary plate fixation after	154	6	U	Y	4	0.026

(*continued on next page*)

Table I Included shoulder and elbow surgery trials with characteristics (*continued*)

Authors	Type of comparison	Outcome	Specific outcome assessed	Patients enrolled	Patients lost to follow-up	Risk of bias	Power calculation	Fragility index	Fragility quotient
			nonunion						
Zhang et al[48]	Medial support screw vs. without	NS	Failure rate	68	4	L	N	1	0.015
Yan et al[47]	RH repair vs. RH replacement	NS	Complications	39	0	H	N	0	0
Ko et al[20]	MMLS vs. simple switch	NS	RCR failure	78	12	H	N	0	0
Li et al[22]	ORIF vs. IMN	Primary	Internal malrotation of 20+°	50	5	U	Y	2	0.04
Putti et al[32]	IMN vs. DCP	NS	Complications	34	0	H	N	0	0
Smekal et al[40]	Operative vs. nonoperative	NS	Delayed unions	60	8	U	N	0	0

ORIF, open reduction internal fixation; *L*, low risk for bias; *TEA*, total elbow replacement; *U*, unclear risk of bias; *PSI*, patient-specific instrumentation; *H*, high risk for bias; *RCR*, rotator cuff repair; *PRP*, platelet rich plasma; *SCD*, sequential compression device; *PGA*, polygluconate-B polymer; *PLLA*, poly-L-lactic acid polymer; *IVF*, intravenous fluid; *PHILOS*, proximal humeral internal locking system; *MIIG*, minimally invasive injectable graft; *RH*, radial head; *MMLS*, modified mattress locking stitch; *IMN*, intramedullary nail; *DCP*, dynamic compression plating; *NS*, not specified.

logistical concerns. Small magnitude trials have the potential to generate results that can be difficult to interpret and potentially overvalued, especially when "statistical significance" is reported. An increasingly used metric, the FI, has been introduced to supplement this analysis and provide an integer value of the number of patients that would need to change outcome groups in order to reverse a study's conclusions. Individually this study not only reports on the FI on a case-by-case basis for the trials included, but also summarizes the field as a whole, highlighting the more robust findings reported by this subspecialty compared with others.

Despite a relatively high median FI, the studies included still contained several limitations in RCT design and methodology. First, only 53% (16/30) of studies specified a primary outcome variable and, furthermore, only 60% (18/30) performed a prestudy power analysis to determine sample size. Although it remains possible that such determinations and calculations were implemented but not reported, the alternative is that many trials did not perform these basic requisites of RCT design and instead randomized a cohort of convenience while potentially tailoring their conclusions to the eventual results. A second limitation involves studies with a FI of 0. Thirty percent (9/30) studies meet this criterion. This circumstance frequently occurs when the χ^2 test is inappropriately used to compare outcomes between groups. Instead, when one of the groups contains less than 5 events, the case in all of the studies included in this analysis, the Fisher exact test is the more appropriate statistical test.[27] Five of 9 trials either report using the χ^2 test[4,7,23,28,32] or fail to report which statistical test was used.[20] In the remaining 4 trials[23,40,41,47] that

Table II Journal of included trials

Journal	n (%)
JSES	8 (27)
JBJS	5 (17)
Arthroscopy	4 (13)
JOT	2 (7)
AJSM	2 (7)
International Orthopaedics	2 (7)
Pakistan Journal of Medicine Science	1 (3)
Biomedical Research International	1 (3)
Journal of International Medical Research	1 (3)
Orthopedic Surgery	1 (3)
Acta Orthop Traumatol Turc	1 (3)
Injury	1 (3)
ANZ Journal of Surgery	1 (3)

JSES, Journal of Shoulder & Elbow Surgery; *JBJS*, Journal of Bone & Joint Surgery; *JOT*, Journal of Orthopedic Trauma; *AJSM*, American Journal of Sports Medicine.

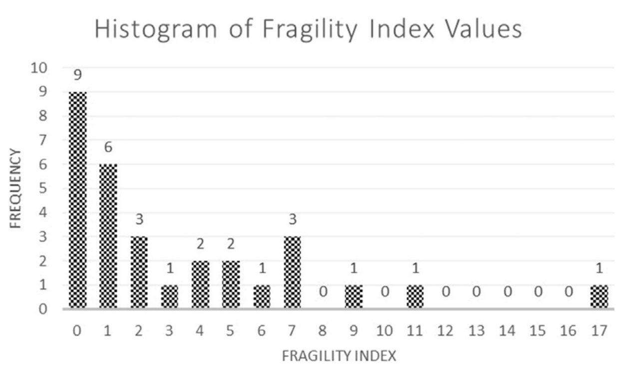

Figure 4 Histogram of fragility index values.

report an FI of 0, the Fisher exact test was used but likely miscalculated. Despite a relatively high median FI when compared with other orthopedic subspecialties, these substantial design and statistical flaws act as reminders that the gold standard of study designs can be compromised and that well performed lesser level of evidence studies, such as case-control studies, matched cohort studies, and meta-analyses, remain critical to scientific evidence. Meta-analyses, specifically, through the process of combining the results of several RCTs have the potential to amplify individual FIs. However, as evidenced by the high rate of methodological concerns in the present systematic review, caution should be used when applying this metric to meta-analyses as poor methodological quality studies with large magnitude FIs can disproportionately influence the overall value. Instead, this metric may be most useful in critical evaluation of individual studies basis alongside RCT methodologies, effect sizes, and CIs.

This study and type of systematic review has several limitations. Because calculating the FI requires dichotomous outcome variables and P values <.05, many studies, including those reporting continuous variables as the primary outcome measures must be eliminated from the analysis, as these variables are continuous in nature. As can be seen in Figure 1, 230 of 406 studies were excluded because of the above requirements alone. A second limitation includes the methodologic restriction of reports published in the last 10 years and those published in the English language. Because the reported FI of 4 is approximately similar to other similar studies reporting on orthopedic subspecialties, it is unlikely that expanding the search would significantly affect this value.

Conclusion

The median FI of RCTs focusing on shoulder and elbow surgery over the past decade is 4, indicating relatively robust findings when compared with RCTs published in other orthopedic subspecialties. The FI is a useful adjunct to other methods of quantitative analyses of results in addition to P values, CIs, and effect sizes. Because of its integer value and ease of calculation, FI facilitates easy interpretation of the robustness vs. fragility of results; however, it should not act as a proxy or attempt to replace P values or careful analysis of study design.

Disclaimer

The authors, their immediate families, and any research foundations with which they are affiliated have not received any financial payments or other benefits from any commercial entity related to the subject of this article.

Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jse.2019.04.051.

References

1. Ahrens PM, Garlick NI, Barber J, Tims EM, The Clavicle Trial Collaborative Group. The clavicle trial: a multicenter randomized controlled trial comparing operative with nonoperative treatment of displaced midshaft clavicle fractures. J Bone Jt Surg 2017;99:1345-54. https://doi.org/10.2106/JBJS.16.01112
2. Bernstein J. Evidence-based medicine. J Am Acad Orthop Surg 2004; 12:80-8.
3. Checketts JX, Scott JT, Meyer C, Horn J, Jones J, Vassar M. The robustness of trials that guide evidence-based orthopaedic surgery. J Bone Joint Surg 2018;100:e85. https://doi.org/10.2106/JBJS.17.01039
4. Chen X, Wang SC, Cao LH, Yang GQ, Li M, Su JC. Comparison between radial head replacement and open reduction and internal fixation in clinical treatment of unstable, multi-fragmented radial head fractures. Int Orthop 2011;35:1071-6. https://doi.org/10.1007/s00264-010-1107-4
5. Chierichini A, Frassanito L, Vergari A, Santoprete S, Chiarotti F, Saccomanno MF, et al. The effect of norepinephrine versus epinephrine in irrigation fluid on the incidence of hypotensive/bradycardic events during arthroscopic rotator cuff repair with interscalene block in the sitting position. Arthroscopy 2015;31:800-6. https://doi.org/10.1016/j.arthro.2015.02.030
6. Dou Q, Ren X. Clinical therapeutic effects of AO/ASIF clavicle hook plate on distal clavicle fractures and acromioclavicular joint dislocations. Pakistan J Med Sci 2014;30:868-71. https://doi.org/10.12669/pjms.304.5269
7. Edwards TB, Labriola JE, Stanley RJ, O'Connor DP, Elkousy HA, Gartsman GM. Radiographic and clinical comparison of pegged and keeled glenoid components using modern cementing techniques: midterm results of a prospective randomized study. J Shoulder Elbow Surg 2010;19:251-7. https://doi.org/10.1016/j.jse.2009.10.013
8. Elmlund AO, Kartus J, Rostgård-Christensen L, Sernert N, Magnusson L, Ejerhed L. A 7-year prospective, randomized, clinical, and radiographic study after arthroscopic Bankart reconstruction using 2 different types of absorbable tack. Am J Sports Med 2009;37:1930-7. https://doi.org/10.1177/0363546509335197
9. Evaniew N, Files C, Smith C, Bhandari M, Ghert M, Walsh M, et al. The fragility of statistically significant findings from randomized trials in spine surgery: a systematic survey. Spine J 2015;15:2188-97. https://doi.org/10.1016/j.spinee.2015.06.004
10. Frey K, Rehm M, Chappell D, Eisenlohr J, Crispin A, Saller T, et al. Preemptive volume therapy to prevent hemodynamic changes caused by the beach chair position: hydroxyethyl starch 130/0.4 versus Ringer's acetate—a controlled randomized trial. J Shoulder Elbow Surg 2018;27:2129-38. https://doi.org/10.1016/j.jse.2018.08.003
11. Gartsman GM, Drake G, Edwards TB, Elkousy HA, Hammerman SM, O'Connor DP, et al. Ultrasound evaluation of arthroscopic full-thickness supraspinatus rotator cuff repair: Single-row versus double-row suture bridge (transosseous equivalent) fixation. Results of a prospective, randomized study. J Shoulder Elbow Surg 2013;22:1480-7. https://doi.org/10.1016/j.jse.2013.06.020
12. Hendel MD, Bryan JA, Barsoum WK, Rodriguez EJ, Brems JJ, Evans PJ, et al. Comparison of patient-specific instruments with standard surgical instruments in determining glenoid component

position: a randomized prospective clinical trial. J Bone Joint Surg Am 2012;94:2167-75. https://doi.org/10.2106/JBJS.K.01209
13. Huang TS, Du WY, Wang TG, Tsai YS, Yang JL, Huang CY, et al. Progressive conscious control of scapular orientation with video feedback has improvement in muscle balance ratio in patients with scapular dyskinesis: a randomized controlled trial. J Shoulder Elbow Surg 2018;27:1407-14. https://doi.org/10.1016/j.jse.2018.04.006
14. Jin L, Guo J, Guo J, Yin Y, Hou Z, Zhang Y. Clinical effects of the probing method with depth gauge for determining the screw depth of locking proximal humeral plate. Biomed Res Int 2016;2016:5898161. https://doi.org/10.1155/2016/5898161
15. Jo CH, Shin JS, Lee YG, Shin WH, Kim H, Lee SY, et al. Platelet-rich plasma for arthroscopic repair of large to massive rotator cuff tears: a randomized, single-blind, parallel-group trial. Am J Sports Med 2013; 41:2240-8. https://doi.org/10.1177/0363546513497925
16. Khan M, Evaniew N, Gichuru M, Habib A, Ayeni OR, Bedi A, et al. The fragility of statistically significant findings from randomized trials in sports surgery: a systematic survey. Am J Sports Med 2017;45: 2164-70. https://doi.org/10.1177/0363546516674469
17. Khormaee S, Choe J, Ruzbarsky JJ, Agarwal KN, Blanco JS, Doyle SM, et al. The fragility of statistically significant results in pediatric orthopaedic randomized controlled trials as quantified by the fragility index: a systematic review. J Pediatr Orthop 2018;38:e418-23. https://doi.org/10.1097/BPO.0000000000001201
18. Kim HM, Caldwell JE, Buza JA, Fink LA, Ahmad CS, Bigliani LU, et al. Factors affecting satisfaction and shoulder function in patients with recurrent rotator cuff tear. J Bone Joint Surg 2014;96:106-12. https://doi.org/10.2106/JBJS.L.01649
19. Kolakowski L, Lai JK, Duvall GT, Jauregui JJ, Dubina AG, Jones DL, et al. Neer Award 2018: benzoyl peroxide effectively decreases preoperative *Cutibacterium acnes* shoulder burden: a prospective randomized controlled trial. J Shoulder Elbow Surg 2018;27:1539-44. https://doi.org/10.1016/j.jse.2018.06.012
20. Ko SH, Lee CC, Friedman D, Park KB, Warner JJP. Arthroscopic single-row supraspinatus tendon repair with a modified mattress locking stitch: a prospective, randomized controlled comparison with a simple stitch. Arthroscopy 2008;24:1005-12. https://doi.org/10.1016/j.arthro.2008.04.074
21. Kwak HJ, Lee JS, Lee DC, Kim HS, Kim JY. The effect of a sequential compression device on hemodynamics in arthroscopic shoulder surgery using beach-chair position. Arthroscopy 2010;26:729-33. https://doi.org/10.1016/j.arthro.2009.10.001
22. Li Y, Wang C, Wang M, Huang L, Huang Q. Postoperative malrotation of humeral shaft fracture after plating compared with intramedullary nailing. J Shoulder Elbow Surg 2011;20:947-54. https://doi.org/10.1016/j.jse.2010.12.016
23. Liu Q-H, Fu Z-G, Zhou J-L, Lu T, Liu T, Shan L, et al. Randomized prospective study of olecranon fracture fixation: cable pin system versus tension band wiring. J Int Med Res 2012;40:1055-66. https://doi.org/10.1177/147323001204000324
24. Liu Z, Zhang G, Ge T. Use of a proximal humeral internal locking system enhanced by injectable graft for minimally invasive treatment of osteoporotic proximal humeral fractures in elderly patients. Orthop. Surg 2011;3:253-8. https://doi.org/10.1111/j.1757-7861.2011.00150.x
25. Lopiz Y, Garcia-Coiradas J, Garcia-Fernandez C, Marco F. Proximal humerus nailing: a randomized clinical trial between curvilinear and straight nails. J Shoulder Elbow Surg 2014;23:369-76. https://doi.org/10.1016/j.jse.2013.08.023
26. Lu QF, Tang GL, Zhao XJ, Zhang WJ, Guo SG, Wang HZ. Tension band wiring through double-cannulated screws as a new internal fixation method for treatment of olecranon fractures: a randomized comparative study. Acta Orthop Traumatol Turc 2015;49:654-60. https://doi.org/10.3944/AOTT.2015.14.0330
27. Lyman S. Letters to the editor letters to the editor. Am J Sports Med 2000;28:918.
28. McKee MD, Veillette CJH, Hall JA, Schemitsch EH, Wild LM, McCormack R, et al. A multicenter, prospective, randomized, controlled trial of open reduction-internal fixation versus total elbow arthroplasty for displaced intra-articular distal humeral fractures in elderly patients. J Shoulder Elbow Surg 2009;18:3-12. https://doi.org/10.1016/j.jse.2008.06.005
29. Melean PA, Zuniga A, Marsalli M, Fritis NA, Cook ER, Zilleruelo M, et al. Surgical treatment of displaced middle-third clavicular fractures: a prospective, randomized trial in a working compensation population. J Shoulder Elbow Surg 2015;24:587-92. https://doi.org/10.1016/j.jse.2014.11.041
30. Murray MM, Flutie BM, Kalish LA, Ecklund K, Fleming BC, Proffen BL, et al. The bridge-enhanced anterior cruciate ligament repair (BEAR) procedure an early feasibility cohort study. Orthop J Sports Med 2016;4:1-11. https://doi.org/10.1177/2325967116672176
31. Oh JH, Chung SW, Oh K-S, Yoo JC, Jee W, Choi J-A, et al. Effect of recombinant human growth hormone on rotator cuff healing after arthroscopic repair: preliminary result of a multicenter, prospective, randomized, open-label blinded end point clinical exploratory trial. J Shoulder Elbow Surg 2018;27:777-85. https://doi.org/10.1016/j.jse.2017.11.019
32. Putti AB, Uppin RB, Putti BB. Locked intramedullary nailing versus dynamic compression plating for humeral shaft fractures. J Orthop Surg 2009;17:139-41. https://doi.org/10.1177/230949900901700202
33. Robinson CM, Goudie EB, Murray IR, Jenkins PJ, Ahktar Ma, Read EO, et al. Open reduction and plate fixation versus nonoperative treatment for displaced midshaft clavicular fractures: a multicenter, randomized, controlled trial. J Bone Joint Surg Am 2013;95:1576-84. https://doi.org/10.2106/JBJS.L.00307
34. Robinson CM, Jenkins PJ, White TO, Ker A, Will E. Primary arthroscopic stabilization for a first-time anterior dislocation of the shoulder: a randomized, double-blind trial. J Bone Joint Surg Am 2008;90:708-21. https://doi.org/10.2106/JBJS.G.00679
35. Ridgeon EE, Young PJ, Bellomo R, Mucchetti M, Lembo R, Landoni G. The fragility index in multicenter randomized controlled critical care trials. Crit Care Med 2016;44:1278-84. https://doi.org/10.1097/CCM.0000000000001670
36. Ruzbarsky JJ, Khormaee S, Daluiski A. The fragility index in hand surgery randomized controlled trials. J Hand Surg Am 2019;44:698.e1-698.e7. https://doi.org/10.1016/j.jhsa.2018.10.005
37. Shen Y, Cheng X, Zhang W. The fragility of randomized controlled trials in intracranial hemorrhage. Neurosurg Rev 2017;42:9-14. https://doi.org/10.1007/s10143-017-0870-8
38. Shochet LR, Kerr PG, Polkinghorne KR. The fragility of significant results underscores the need of larger randomized controlled trials in nephrology. Kidney Int 2017;92:1468-75. https://doi.org/10.1016/j.kint.2017.05.011
39. Skinner M, Tritz D, Farahani C, Ross A, Hamilton T, Vassar M. The fragility of statistically significant results in otolaryngology randomized trials. Am J Otolaryngol 2018;40:61-6. https://doi.org/10.1016/j.amjoto.2018.10.011
40. Smekal V, Irenberger A, Struve P, Wambacher M, Krappinger D, Kralinger FS. Elastic stable intramedullary nailing versus nonoperative treatment of displaced midshaft clavicular fractures—a randomized, controlled, clinical trial. J Orthop Trauma 2009;23:106-12. https://doi.org/10.1097/BOT.0b013e318190cf88
41. Soliman OA, Koptan WMT. Proximal humeral fractures treated with hemiarthroplasty: does tenodesis of the long head of the biceps improve results? Injury 2013;44:461-4. https://doi.org/10.1016/j.injury.2012.09.012
42. Syed UAM, Aleem AW, Wowkanech C, Weekes D, Freedman M, Tjoumakaris F, et al. Neer Award 2018: the effect of preoperative education on opioid consumption in patients undergoing

arthroscopic rotator cuff repair: a prospective, randomized clinical trial. J Shoulder Elbow Surg 2018;27:962-7. https://doi.org/10.1016/j.jse.2018.02.039
43. Torrens C, Guirro P, Miquel J, Santana F. Influence of glenosphere size on the development of scapular notching: a prospective randomized study. J Shoulder Elbow Surg 2016;25:1735-41. https://doi.org/10.1016/j.jse.2016.07.006
44. Walsh MR, Nelson BJ, Braman JP, Yonke B, Obermeier M, Raja A, et al. Platelet-rich plasma in fibrin matrix to augment rotator cuff repair: a prospective, single-blinded, randomized study with 2-year follow-up. J Shoulder Elbow Surg 2018;27:1553-63. https://doi.org/10.1016/j.jse.2018.05.003
45. Walsh M, Srinathan SK, McAuley DF, Mrkobrada M, Levine O, Ribic C, et al. The statistical significance of randomized controlled trial results is frequently fragile: a case for a fragility index. J Clin Epidemiol 2014;67:622-8. https://doi.org/10.1016/j.jclinepi.2013.10.019
46. Woltz S, Stegeman SA, Krijnen P, Van Dijkman BA, Van Thiel TPH, Schep NWL, et al. Plate fixation compared with nonoperative treatment for displaced midshaft clavicular fractures a multicenter randomized controlled trial. J Bone Joint Surg Am 2017;99:106-12. https://doi.org/10.2106/JBJS.15.01394.
47. Yan M, Ni J, Song D, Ding M, Liu T, Huang J. Radial head replacement or repair for the terrible triad of the elbow: which procedure is better? ANZ J Surg 2015;85:644-8. https://doi.org/10.1111/ans.13060
48. Zhang L, Zheng J, Wang W, Lin G, Huang Y, Zhang J, et al. The clinical benefit of medial support screws in locking plating of proximal humerus fractures: a prospective randomized study. Int Orthop 2011;35:1655-61. https://doi.org/10.1007/s00264-011-1227-5.