J Shoulder Elbow Surg (2021) 30, 1787–1793

ELSEVIER

Journal of Shoulder and Elbow Surgery

www.elsevier.com/locate/ymse

Statistical fragility of randomized clinical trials in shoulder arthroplasty

Kyle L. McCormick, MD, Liana J. Tedesco, MD, Hasani W. Swindell, MD, Lynn Ann Forrester, MD, Charles M. Jobin, MD, William N. Levine, MD*

Department of Orthopedic Surgery, Columbia University Medical Center, Columbia University, New York, NY, USA

Background: The *P* value is a statistical tool used to assess the statistical significance of clinical trial outcomes in orthopedic surgery. However, the *P* value does not evaluate research quality or clinical significance. The Fragility Index (FI) is an alternative statistical method that can be used to assess the quality and significance of clinical research and is defined as the number of patients in a study intervention group necessary to convert an outcome from statistically significant to statistically insignificant or vice versa. The primary purpose of this study was to evaluate the statistical robustness of clinical trials regarding shoulder arthroplasty using the FI. The secondary goal was to identify trial characteristics associated greater statistical fragility.
Methods: A systematic review of randomized clinical trials in shoulder arthroplasty was performed. The FI was calculated for all dichotomous, categorical study outcomes discussed in the identified studies. Descriptive statistics and the Pearson correlation coefficient were used to evaluate all studies and characterize associations between study variables.
Results: A total of 13 randomized controlled trials were identified and evaluated; these trials had a median sample size of 47 patients (mean, 54 patients; range, 26-102 patients) and a median of 7 patients (mean, 5.8 patients; range, 0-14 patients) lost to follow-up. The median FI was 6 (mean, 5; range, 1-11), a higher FI than what has been observed in other orthopedic subspecialties. However, the majority of outcomes (74.4%) had an FI that was less than the number of patients lost to follow-up, and most outcomes (89.7%) were statistically insignificant.
Conclusion: Randomized controlled trials in shoulder arthroplasty have comparable statistical robustness to the literature in other orthopedic surgical subspecialties. We believe that the inclusion of the FI in future comparative studies in the shoulder arthroplasty literature will allow surgeons to better assess the statistical robustness of future research.
Level of evidence: Basic Science Study; Statistics/Measurement Error

Keywords: Shoulder arthroplasty; Fragility Index; randomized controlled trials shoulder surgery; statistical fragility; clinical trials

The *P* value is a powerful statistical tool used to evaluate the statistical significance of research outcomes in the orthopedic literature. The shortcoming of the *P* value is that it does not provide information on the effect size or the statistical strength of a study outcome.[32] To address this concern, many authors have advocated the broader use of alternative statistical methods to evaluate clinical data.[7,8,22,26,31] One proposed alternative tool is the Fragility Index (FI), defined as the number of patients, if given an alternative result, that would be sufficient to change the statistical significance of a study outcome. The FI is calculated through the stepwise alteration of an outcome in

Institutional review board approval was not required for this basic science study.

*Reprint Requests: William N. Levine, MD, Department of Orthopedic Surgery, Columbia University Medical Center, 622 West 168 St, PH-11, New York, NY 10032.

E-mail address: Wnl1@cumc.columbia.edu (W.N. Levine).

1058-2746/© 2020 The Author(s). This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).
https://doi.org/10.1016/j.jse.2020.10.028

a study arm until the recalculated P value changes from significant to insignificant or vice versa. A small FI suggests that the outcome could have been readily altered with just a few different patient results, whereas a large FI suggests a more statistically robust outcome. The FI provides a useful metric to demonstrate the number of patients required to change the significance of a certain outcome. Because the P value and confidence intervals are viewed dichotomously, significance is determined based on compatibility of data with a null hypothesis. The FI offers information on the effect size, making it a powerful supplement to commonly used statistical tools.

The majority of articles discussing shoulder arthroplasty (82.8%) have been published within the past 5 years, and there has been a substantial increase in the number of shoulder arthroplasty procedures performed in the United States, from 14,000 in 2000 to 70,000 in 2011.[13] Additionally, studies from Finland, Australia, Denmark, and New Zealand have reported a concordant increased incidence of shoulder arthroplasties in the past 2 decades.[9,1,20,25] Given the rising global demand for shoulder arthroplasties, there is a need for critical evaluation of the growing shoulder arthroplasty literature to optimize future patient care. Analyzing randomized controlled trials (RCTs) in shoulder arthroplasty will facilitate the development of higher-quality clinical practice guidelines, as well as help establish a standard for future research in the field.

To our knowledge, no studies to date have provided an in-depth analysis of shoulder arthroplasty research as it relates to FI. The primary purpose of this study was to evaluate the statistical strength of shoulder arthroplasty clinical trials using the FI. A secondary goal was to identify what characteristics of clinical trials are associated with greater statistical fragility.

Methods

Study design and eligibility criteria

A systematic review was performed using methods comparable to prior analyses of statistical fragility,[7,14,15,22,26,30] and the 25 highest-impact journals relevant to orthopedic surgery and shoulder arthroplasty were identified using InCites Journal Citation Reports (Clarivate Analytics, Philadelphia, PA, USA). A query of PubMed was performed to identify all clinical trials in shoulder arthroplasty published in the previously identified journals between January 1, 2009, and June 9, 2019. Specifically, we used the following search strategy: "shoulder" OR "shoulder arthroplasty" OR "arthroplasty" AND "orthopedics" OR "orthopedic procedures" OR "surgery" OR "surgical procedures" AND "randomized controlled trial" AND "2009/01/01" [PDAT]: "2019/06/19" [PDAT] AND "English" [language]. The search was limited to human trials specifically investigating surgical interventions. The inclusion criteria were the use of 1:1 parallel, 2-arm randomization procedures and $\geq$1 dichotomous study outcome. The titles and abstracts of each article were screened independently by 2 authors (K.L.M. and H.W.S.) to ensure that all included studies met the inclusion criteria. Each remaining article was then reviewed in its entirety to identify final eligibility and record data on all dichotomous, categorical outcomes, both primary and secondary. Collected data were placed into a 2 $\times$ 2 contingency table for further analysis.[7,32]

Finally, study sample size, study outcomes, number of patients lost to follow-up, reported P values, publication year, and journal of publication were collected. Data on journal impact factor, number of journal citations, and relative citation ratio (RCR) were also collected from InCites Journal Citation Reports for further analysis of publication-level variables. In addition, the National Institutes of Health iCite database (Bethesda, MD, USA) was used to identify the RCR for each included study.[8,11]

Calculation of FI

The FI for each outcome was defined as the lowest number of outcomes that must be reversed to change the statistical significance of a calculated P value. Individual FI scores were calculated for all categorical, dichotomous outcomes investigated in the collected studies via the method previously described by Walsh et al.[31] The P value for each outcome was recalculated using the 2-sided Fisher exact test. For each outcome group, discrete outcome events were switched from the larger outcome group to the smaller outcome group in a stepwise fashion until the recalculated P value was $>$0.05. Conversely, for statistically insignificant P values, events in the smaller outcome group were changed in the same manner until the P value was $<$.05 and thus statistically significant.

Statistical analysis

Collected data were analyzed using the Student t test to identify differences between the aforementioned study variables. The Pearson correlation coefficient was used to determine associations between publication-level variables, as well as the association between the calculated FI and the P values of included studies. All analyses were performed using SPSS software (version 23; IBM, Armonk, NY, USA) and Microsoft Excel 2016 (Redmond, WA, USA). For studies with multiple dichotomous outcomes, the highest calculated FI was used when comparing publication-level variables to prevent disproportionate weighting of studies with more outcomes.

Results

A total of 76 articles were initially identified. After review, 35 studies were excluded because they did not examine the shoulder or did not evaluate surgical interventions (ie, postoperative pain management). An additional 28 articles were removed from the analysis as they lacked dichotomous, categorical outcomes. Ultimately, 13 studies were included in the final analysis, and a total of 39 dichotomous, categorical outcomes were identified (Fig. 1). The majority of articles (53.8%) were published between 2010

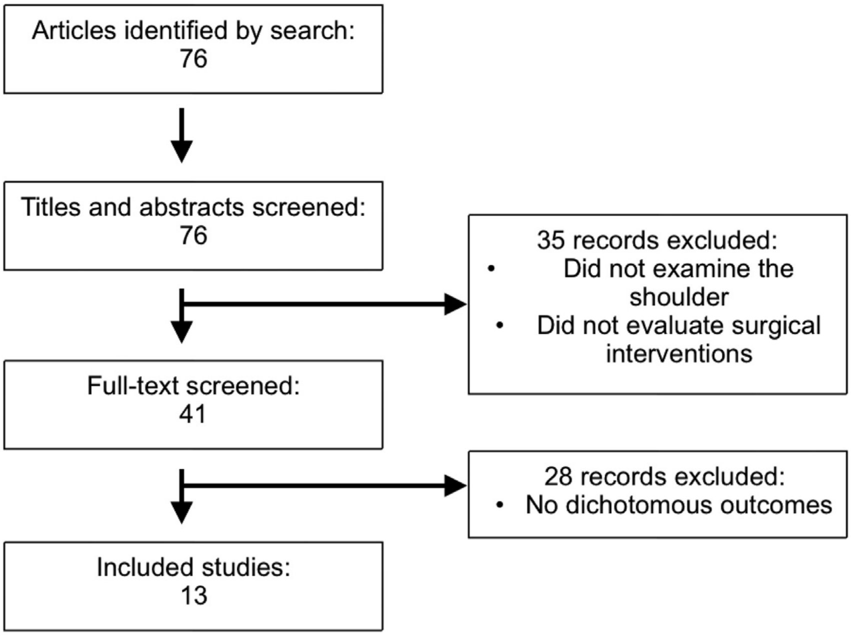

Figure 1 CONSORT (Consolidated Standards of Reporting Trials) diagram for exclusion criteria.

Table I Number of publications by journal

Journal	No. of publications
Journal of Shoulder and Elbow Surgery	10
Journal of Bone and Joint Surgery	2
Clinical Orthopaedics and Related Research	1

and 2014, and most studies (76.9%) were published in the *Journal of Shoulder and Elbow Surgery* (Table I). Two of the studies evaluated the same patient population at different follow-up times, with different numbers of patients lost to follow-up.[16,17]

The outcomes reported in the studies were as follows: radiographic findings (66.7%), failure or revision (10.3%), complications (17.1%), survival (2.6%), and healing on imaging at follow-up (2.6%). We evaluated both primary and secondary outcome variables. The median sample size in the analyzed studies was 47 patients (mean, 54 patients; range, 26-102 patients), with a median of 7 patients (mean, 5.8 patients; range, 0-14 patients) lost to follow-up. All included studies reported the number of patients lost to follow-up (100%).

Fragility Index

The median FI for the 39 outcomes measured was 6 (mean, 5; range, 1-11) (Fig. 2). Of the studied outcomes, 4 were statistically significant whereas 35 were statistically insignificant. The median FI for statistically significant outcomes was 1 (mean, 1.3; range, 1-2), whereas the median FI for statistically insignificant outcomes was 6 (mean, 5.5; range, 1-11). There was a statistically significant difference between the FIs calculated for statistically significant outcomes compared with statistically insignificant outcomes ($P = .0006$).

When exclusively statistically insignificant outcomes were examined, there was a positive correlation between the FI and reported P value ($R = 0.3445$, $P = .042705$). A similar correlation involving statistically significant outcomes could not be calculated because of a smaller sample of statistically significant results.

The FI was ≤ 3 events in 10 of the 39 dichotomous outcomes analyzed (25.6%); 29 outcomes (74.4%) had an FI less than or equal to the total number of patients lost to follow-up. However, no statistically significant association was observed between the number of patients lost to follow-up and the FI ($R = -0.0316$, $P = .918376$).

For publication-level variables, we found no correlation between the FI and sample size, publication year, journal impact factor, or number of journal citations (Table II). Finally, no correlation was observed between patient sample size and number of citations ($R = -0.269$, $P = .374$), between patient sample size and journal impact factor ($R = -0.302$, $P = .316$), or between RCR and publication year ($R = 0.134$, $P = .695$).

Discussion

Shoulder arthritis is a debilitating condition that affects up to one-third of the population aged > 60 years.[4,21] As the quality, longevity, and efficacy of available shoulder

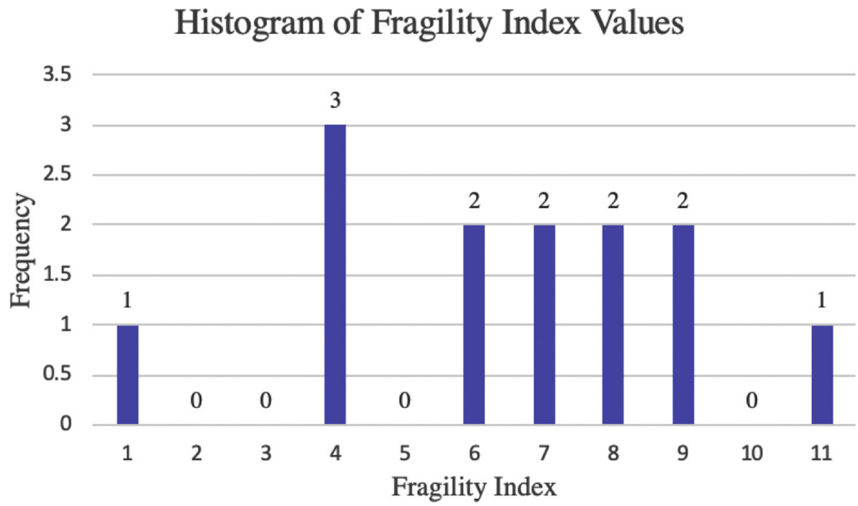

Figure 2 Histogram of Fragility Index values.

Table II Publication-level associations between Fragility Index and study variables

Study variable	Pearson correlation coefficient	*P* value
Patient sample size	0.228	.455
Publication year	0.516	.0713
Journal impact factor	−0.0355	.908
No. of journal citations	−0.0513	.868
Patients lost to follow-up	−0.0316	.918

implants continue to improve,[18] shoulder arthroplasty has become a significantly more prevalent procedure. Issa et al[12] noted an increase in admissions from 8041 to 39,072 for total shoulder arthroplasties in the United States from 1998 and 2010. Additionally, Kim et al[18] found that, following a rise in the number of total shoulder arthroplasties in 2004, there has been a steady increase in demand for this procedure annually. As the body of literature evaluating shoulder arthroplasty techniques and indications grows, closer critique of new research will be necessary to help guide future clinical practice.

The purpose of this study was to evaluate the strength of surgical clinical trials regarding shoulder arthroscopy by using the FI (Table III). Our investigation found the median FI for all included studies to be 6. For all statistically significant outcomes, the FI was 1, and for statistically insignificant outcomes, the FI was 6. The study by Trofa et al,[28] evaluating the association between drain use and postoperative transfusion rates, had an FI of 11, the largest in our study. In their evaluation of the incidence of postoperative complications, they found no statistically significant difference in patients with versus without postoperative drain use. Seven RCTs evaluated different glenoid components, with FIs ranging from 4 to 9,[5,6,16,17,23,24,29] and the lowest FI (1) was calculated in the study by Sebastiá-Forcada et al[27] comparing outcomes in shoulder arthroplasty and hemiarthroplasty.

The American Academy of Orthopaedic Surgeons (AAOS) recently published clinical practice guidelines for evaluating research. For orthopedic research, an article with a median FI of 2 is considered to have ‘‘strong evidence’’ in support of the reported findings.[3] In orthopedics, spine RCTs (40 in total) were found to have an FI of 2,[7] hand surgery (5 total) and pediatric orthopedic (17 total) RCTs had an FI of 3,[14,26] and sports RCTs (40 in total) had an FI of 5.[14,22] These studies exclusively evaluated outcomes that were statistically significant.

The FI has traditionally been used exclusively for statistically significant outcomes. However, there is no literature to suggest it cannot have the same value for statistically insignificant outcomes. Given that statistically insignificant outcomes may influence clinical practice, we included both in our analysis. This contributed to a higher overall FI when looking at shoulder arthroplasty RCTs. Our study found a median FI of 6, which is above the threshold the AAOS would consider strong evidence.

The original FI study by Walsh et al[31] found that the 25th percentile of all FIs in high-impact journals was 3. Thus, by convention, 3 is the threshold used to characterize an outcome as statistically robust. In this analysis, 29 of the included outcomes (74.4%) had FIs > 3, suggesting that RCTs in shoulder arthroplasty have relative statistical strength. However, it is important to note that this is based on both statistically significant and insignificant outcomes. Our study demonstrated a median FI of 1 and a mean of 1.3 for statistically significant outcomes, below the AAOS threshold for strong evidence.

With this in mind, we acknowledge that this study is not without limitations. First, the FI can only assess

Table III Analyzed shoulder arthroplasty articles

Authors	Year	Journal	Type of comparison	Patients enrolled	Patients lost to follow-up	Fragility Index	Dichotomous outcomes
Trofa et al[28]	2019	*JSES*	Drain use vs no drain use	100	0	11	2
Kilian et al[16]	2018	*JSES*	All-polyethylene CL component vs conventional all-polyethylene P component	54	12	7	12
Kilian et al[17]	2017	*JSES*	Pegged vs keeled glenoid component	46	8	6	8
Vara et al[30]	2017	*JSES*	Intravenous tranexamic acid vs none	116	14	8	14
Uschok et al[29]	2017	*JSES*	Stemless vs standard-stem humeral head	40	7	9	7
Sebastiá-Forcada et al[27]	2014	*JSES*	Reverse shoulder arthroplasty vs hemiarthroplasty	62	1	1	1
Poon et al[23]	2014	*JBJS Am*	Concentric vs eccentric glenosphere	50	0	9	1
Ho et al[10]	2013	*JSES*	Radiographs vs CT scans	32	0	7	4
Lapner et al[19]	2013	*JSES*	Lesser tuberosity osteotomy vs subscapularis peel for exposure	87	8	4	8
Edwards et al[6]	2012	*JSES*	Glenoid component with no inferior tilt vs glenoid component inferiorly tilted 10°	52	10	4	10
Boons et al[2]	2012	*CORR*	Operative vs nonoperative	55	8	6	8
Edwards et al[5]	2010	*JSES*	Pegged vs keeled glenoid component	53	6	8	6
Rahme et al[24]	2009	*JBJS Am*	Cemented, all-polyethylene, keeled glenoid component vs cemented, all-polyethylene, in-line 3-pegged glenoid component	28	2	4	2

JSES, Journal of Shoulder and Elbow Surgery; *JBJS Am*, *The Journal of Bone and Joint Surgery*–American edition; *CORR*, *Clinical Orthopaedics and Related Research*; *CT*, computed tomography; *CL*, cementless; *P*, pegged.

dichotomous outcome variables. In research regarding shoulder arthroplasty, there is significant value to assessing continuous outcomes such as range of motion, pain relief, and strength. As a result, the majority of studies identified in our initial literature review were excluded from this analysis (63 of 76). Therefore, the FI cannot be used to draw conclusions regarding the statistical strength of shoulder RCTs that exclusively analyze continuous variables.

Additionally, the literature search used in this study was restricted to the past 10 years, likely excluding older studies. However, this methodology was chosen as it has been established as the convention used in FI research in orthopedics.[7,14,15,22,26] Another potential limitation of this analysis was that articles were selected exclusively from the top 25 highest-impact orthopedic journals, suggesting the study's findings may have been biased toward studies of greater statistical strength. However, previous studies have evaluated < 12 journals; thus, the literature review in this study was likely more comprehensive than that used previously in evaluating the orthopedic literature.[22,26]

Performing an RCT is a tremendous endeavor, with the study's success predicated on the ability to adequately randomize and conceal allocation of patients into separate study groups. Although the convention is that details regarding this process should be revealed in all studies identified as RCTs, it was not possible to verify the randomization process in all included studies. However, as all included studies were published in rigorously peer-reviewed orthopedic journals such as the *Journal of Shoulder and Elbow Surgery*, we believed that the intensive review process of these journals likely limited this potential source of bias.

Conclusion

This analysis uses the FI to assess the strength of the largest cohort of recent, high-impact literature in shoulder arthroplasty to date. Furthermore, this study acts as a comprehensive evaluation of all dichotomous outcomes in the highest-impact journals regarding shoulder arthroplasty, as both statistically significant and statistically insignificant outcomes were assessed.[14] Overall, the FI is an easily quantifiable and interpreted complement to typical statistical evaluation metrics.

Therefore, we argue that inclusion of the FI in future comparative studies in the shoulder arthroplasty literature will allow surgeons to better assess the statistical robustness of future research and will help facilitate the production of higher-quality research in the field.

Disclaimer

Charles M. Jobin reports that he is a paid consultant, presenter, and speaker for Acumed; receives research support from Acumed; receives paid consulting fees from CFO, DePuy (A Johnson & Johnson Company), and Integral Life Sciences; is a paid consultant for Wright Medical Technology and Integrated Shoulder Collaboration; receives paid fees from Wright Medical Technology; is a paid consultant and presenter for Zimmer Biomet; receives speaker fees from Zimmer Biomet; is an American Shoulder and Elbow Surgeons board or committee member; and is on the editorial board of the *Journal of American Academy of Orthopaedic Surgeons*.

William N. Levine is an American Shoulder and Elbow Surgeons board or committee member; is on the editorial board of *Journal of American Academy of Orthopaedic Surgeons*; and reports unpaid consulting agreements with Zimmer.

The other 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.

References

1. Australian Orthopaedic Association. Annual report from the Australian Joint Replacement Registry. Demographics and outcome of shoulder arthroplasty. 2013. https://aoanjrr.sahmri.com/documents/10180/217645/Shoulder Arthroplasty, Accessed 30 January 2020.
2. Boons HW, Goosen JH, Van Grinsven S, Van Susante JL, Van Loon CJ. Hemiarthroplasty for humeral four-part fractures for patients 65 years and older a randomized controlled trial. Clin Orthop Relat Res 2012;470:3483-91. https://doi.org/10.1007/s11999-012-2531-0
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 Am 2018;100:e85. https://doi.org/10.2106/jbjs.17.01039
4. Chillemi C, Franceschini V. Shoulder osteoarthritis. Arthritis 2013; 2013:370231. https://doi.org/10.1155/2013/370231
5. Edwards TB, Labriola JE, Stanley RJ, O'Connor DP, Elkousy HA, Gartsman GM. Radiographic comparison of pegged and keeled glenoid components using modern cementing techniques: a prospective randomized study. J Shoulder Elbow Surg 2010;19:251-7. https://doi.org/10.1016/j.jse.2009.10.013
6. Edwards TB, Trappey GJ, Riley C, O'Connor DP, Elkousy HA, Gartsman GM. Inferior tilt of the glenoid component does not decrease scapular notching in reverse shoulder arthroplasty: results of a prospective randomized study. J Shoulder Elbow Surg 2012;21:641-6. https://doi.org/10.1016/j.jse.2011.08.057
7. 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
8. Grolleau F, Collins GS, Smarandache A, Pirracchio R, Gakuba C, Boutron I, et al. The fragility and reliability of conclusions of anesthesia and critical care randomized trials with statistically significant findings: a systematic review. Crit Care Med 2019;47:456-62. https://doi.org/10.1097/ccm.0000000000003527
9. Harjula JNE, Paloneva J, Haapakoski J, Kukkonen J, Aarimaa V. Increasing incidence of primary shoulder arthroplasty in Finland—a nationwide registry study. BMC Musculoskelet Disord 2018;19:245. https://doi.org/10.1186/s12891-018-2150-3
10. Ho JC, Youderian A, Davidson IU, Bryan J, Iannotti JP. Accuracy and reliability of postoperative radiographic measurements of glenoid anatomy and relationships in patients with total shoulder arthroplasty. J Shoulder Elbow Surg 2013;22:1068-77. https://doi.org/10.1016/j.jse.2012.11.015
11. Hutchins BI, Yuan X, Anderson JM, Santangelo GM. Relative citation ratio (RCR): a new metric that uses citation rates to measure influence at the article level. PLoS Biol 2016;14:e1002541. https://doi.org/10.1371/journal.pbio.1002541
12. Issa K, Pierce CM, Pierce TP, Boylan MR, Zikria BA, Naziri Q, et al. Total shoulder arthroplasty demographics, incidence, and complications—A Nationwide Inpatient Sample database study. Surg Technol Int 2016;29:240-6.
13. Jain NB, Yamaguchi K. The contribution of reverse shoulder arthroplasty to utilization of primary shoulder arthroplasty. J Shoulder Elbow Surg 2014;23:1905-12. https://doi.org/10.1016/j.jse.2014.06.055
14. 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
15. 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
16. Kilian CM, Morris BJ, Sochacki KR, Gombera MM, Haigler RE, O'Connor DP, et al. Radiographic comparison of finned, cementless central pegged glenoid component and conventional cemented pegged glenoid component in total shoulder arthroplasty: a prospective randomized study. J Shoulder Elbow Surg 2018;27:S10-6. https://doi.org/10.1016/j.jse.2017.09.014
17. Kilian CM, Press CM, Smith KM, O'Connor DP, Morris BJ, Elkousy HA, et al. 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 2017;26:2078-85. https://doi.org/10.1016/j.jse.2017.07.016
18. Kim SH, Wise BL, Zhang Y, Szabo RM. Increasing incidence of shoulder arthroplasty in the United States. J Bone Joint Surg Am 2011; 93:2249-54. https://doi.org/10.2106/JBJS.J.01994
19. Lapner PLC, Sabri E, Rakhra K, Bell K, Athwal GS. Healing rates and subscapularis fatty infiltration after lesser tuberosity osteotomy versus subscapularis peel for exposure during shoulder arthroplasty. J Shoulder Elbow Surg 2013;22:396-402. https://doi.org/10.1016/j.jse.2012.05.031
20. New Zealand Orthopaedic Association. The New Zealand Joint Registry: sixteen year report, January 1999 to December 2014. http://www.nzoa.org.nz/system/files/Web_DH7657_NZJR2014Report_v4_12Nov15.pdf; 2015. Accessed 30 January 2020.
21. Pandya J, Johnson T, Low AK. Shoulder replacement for osteoarthritis: a review of surgical management. Maturitas 2018;108:71-6. https://doi.org/10.1016/j.maturitas.2017.11.013

22. Parisien RL, Trofa DP, Dashe J, Cronin PK, Curry EJ, Fu FH, et al. Statistical fragility and the role of P values in the sports medicine literature. J Am Acad Orthop Surg 2019;27:e324-9. https://doi.org/10.5435/jaaos-d-17-00636
23. Poon PC, Chou J, Young SW, Astley T. A comparison of concentric and eccentric glenospheres in reverse shoulder arthroplasty: a randomized controlled trial. J Bone Joint Surg Am 2014;96:e138.1-7. https://doi.org/10.2106/JBJS.M.00941
24. Rahme H, Mattsson P, Wikblad L, Nowak J, Larsson S. Stability of cemented in-line pegged glenoid compared with keeled glenoid components in total shoulder arthroplasty. J Bone Joint Surg Am 2009; 91:1965-72. https://doi.org/10.2106/JBJS.H.00938
25. Rasmussen JV, Jakobsen J, Brorson S, Olsen BS. The Danish Shoulder Arthroplasty Registry: clinical outcome and short-term survival of 2,137 primary shoulder replacements. Acta Orthop 2012;83:171-3. https://doi.org/10.3109/17453674.2012.665327
26. Ruzbarsky JJ, Khormaee S, Daluiski A. The Fragility Index in hand surgery randomized controlled trials. J Hand Surg Am 2019;44:698.e1-e7. https://doi.org/10.1016/j.jhsa.2018.10.005
27. Sebastiá-Forcada E, Cebrián-Gómez R, Lizaur-Utrilla A, Gil-Guillén V. Reverse shoulder arthroplasty versus hemiarthroplasty for acute proximal humeral fractures. A blinded, randomized, controlled, prospective study. J Shoulder Elbow Surg 2014;23:1419-26. https://doi.org/10.1016/j.jse.2014.06.035
28. Trofa DP, Paulino FE, Munoz J, Villacis DC, Irvine JN, Jobin CM, et al. Short-term outcomes associated with drain use in shoulder arthroplasties: a prospective, randomized controlled trial. J Shoulder Elbow Surg 2019;28:205-11. https://doi.org/10.1016/j.jse.2018.10.014
29. Uschok S, Magosch P, Moe M, Lichtenberg S, Habermeyer P. Is the stemless humeral head replacement clinically and radiographically a secure equivalent to standard stem humeral head replacement in the long-term follow-up? A prospective randomized trial. J Shoulder Elbow Surg 2017;26:225-32. https://doi.org/10.1016/j.jse.2016.09.001
30. Vara AD, Koueiter DM, Pinkas DE, Gowda A, Wiater BP, Wiater JM. Intravenous tranexamic acid reduces total blood loss in reverse total shoulder arthroplasty: a prospective, double-blinded, randomized, controlled trial. J Shoulder Elbow Surg 2017;26:1383-9. https://doi.org/10.1016/j.jse.2017.01.005
31. 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
32. Wasserstein RL, Lazar NA. The ASA statement on p-values: context, process, and purpose. Am Stat 2016;70:129-33. https://doi.org/10.1080/00031305.2016.1154108