



# The Fragility of Statistical Findings in Cervical Disc Arthroplasty: a Systematic Review of Randomized Controlled Trials

Emmanuel C. Megafu<sup>1</sup> · Michael N. Megafu<sup>2</sup> · Janet T. Nguyen<sup>1</sup> · Elisabeth Point Du Jour<sup>1</sup> · Wesley H. Bronson<sup>3</sup> · James D. Lin<sup>3</sup> · Andrew C. Hecht<sup>3</sup> · Robert L. Parisien<sup>3</sup>

Received: 6 March 2024 / Accepted: 26 April 2024  
© The Author(s) 2024

## Abstract

**Purpose** This study employs both the fragility index (FI) and fragility quotient (FQ) to assess the level of robustness in the cervical disc arthroplasty (CDA) literature. We hypothesize that dichotomous outcomes involving CDA would exhibit statistical vulnerability.

**Methods** A PubMed search was conducted to evaluate dichotomous data for randomized controlled trials (RCTs) in CDA literature from 2000 to 2023. The FI of each outcome was calculated through the reversal of a single outcome event until significance was reversed. The FQ was calculated by dividing each fragility index by the study sample size. The interquartile range (IQR) was also calculated for the FI and FQ.

**Results** Of the 1561 articles screened, 111 met the search criteria, with 35 RCTs evaluating CDA included for analysis. Six hundred and ninety-three outcome events with 130 significant ( $P < 0.05$ ) outcomes and 563 nonsignificant ( $P \geq 0.05$ ) outcomes were identified. The overall FI and FQ for all 693 outcomes were 5 (IQR 3–7) and 0.019 (IQR 0.011–0.043). Fragility analysis of statistically significant outcomes and nonsignificant outcomes both revealed an FI of 5. All of the studies reported loss to follow-up (LTF) data where 65.7% (23) did not report or reported an LTF greater or equal to 5.

**Conclusions** The literature regarding CDA RCTs lacks statistical robustness and may misrepresent the conclusions with the sole use of the  $P$  value. By implementing the FI and FQ along with the  $P$  value, we believe the interpretation and contextualization of the clinical data surrounding CDA will be better understood.

**Keywords** Statistical fragility · Cervical disc arthroplasty · Randomized controlled trials · Fragility index · Fragility quotient

## Introduction

Cervical disc arthroplasty (CDA) is an evolving surgical approach for addressing degenerative disc diseases in the cervical spine and has gained attention as an alternative to anterior cervical discectomy and fusion (ACDF). This procedure involves replacing degenerative discs with artificial

implants, aiming to preserve motion and potentially mitigate issues commonly associated with fusion, such as adjacent segment degeneration [1, 2]. Several clinical trials have shown promising outcomes in the short term, emphasizing benefits like preserved motion and reduced rates of adjacent segment pathology [3]. However, a comprehensive understanding of the long-term efficacy, safety, and comparative outcomes of cervical disc arthroplasty remains a focus of ongoing research [4, 5]. The need for further investigation into the long-term clinical outcomes, complication rates, and the effects of adjacent-level post-cervical disc arthroplasty is evident [6, 7]. Studies have explored the long-term results of cervical arthroplasty and provided valuable insights into the outcomes and complications; these research efforts aim to provide evidence-based guidelines for clinical decision-making and to enhance patient outcomes by filling

✉ Michael N. Megafu  
Michael.Megafu@atsu.edu

<sup>1</sup> Geisinger Commonwealth School of Medicine, Scranton, PA, USA

<sup>2</sup> A.T. Still University, Kirksville College of Osteopathic Medicine, Kirksville, MO, USA

<sup>3</sup> Department of Orthopedic Surgery, Mount Sinai Hospital, New York, NY, USA

the knowledge gaps in the field of cervical disc arthroplasty [8–10].

Considering all possible complications of cervical disc diseases, applications of evidence-based medical research become essential when choosing the best treatment method. Statistical approaches that utilize the  $P$  value reveal the significance of potential divergences between clinical interventions [11]. In the case of CDA, evidence-based medical practice can apply the foundational implications of said statistical results to determine treatment outcomes for damaged cervical discs. Various clinical literature has used the  $P$  value as a conventional evaluation method for interpreting the marginal significance of statistical results. In recent research findings, novel methodologies have been adopted with the utilization of the fragility index (FI) and fragility quotient (FQ) to assess the robustness of clinical trial results. When evaluating the fragility of randomized controlled trials (RCTs), the FI is to reveal the necessary number required to reverse a trial from high degrees of significance to low or none [12]. In addition, as a measurement of fragility, the FQ further assesses the FI concerning the sample size by dividing the FI by sample size [13]. Combined, the FI and FQ complement the  $P$  value and promote confidence that low degrees of fragility (higher fragility index) indicate more robust clinical trial data.

Evidence-based medical research will enable physicians to practice evidence-based patient care when treating cervical disc diseases. Despite the debate between conservative and nonconservative management of cervical disc diseases, available literary evidence continues to explore the overall patient outcomes of CDA [14–16]. With these considerations, it becomes even more essential to analyze the robustness of the results in the literature. This study evaluates the degree of statistical fragility in the cervical disc arthroplasty (CDA) literature. We hypothesize that the dichotomous outcomes within the CDA literature are statistically fragile and will be vulnerable to a small number of outcome event reversals.

## Methods and materials

A systematic review was conducted to compare statistical robustness for cervical disc injuries. Since this was a systematic review, no ethical consent or IRB consultation was needed to continue this study. The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines were used for article identification and consequent selection. The PubMed database was queried from 2000 to 2023, for all RCTs relating to cervical disc arthroplasty. The language was restricted to English. Search criteria involved articles containing: “cervical disc arthroplasty” OR “cervical disc replacement”. Inclusion criteria consisted of RCTs describing dichotomous outcomes with associated  $P$  values. Studies evaluating non-RCT, non-dichotomous outcomes, systematic reviews, meta-analyses, animal studies, cadaveric studies, biomechanical studies, case reports, greater than two intervention group studies, and cervical disc articles before 2000 were excluded from this review. For the included studies, the following were extracted: journal name, publication year, authors, PubMed Identifier, loss to follow-up (LTF), study design, and  $P$  values. The journals that met the search criteria included: *Clinical Orthopaedics and Related Research*, *Clinical Spine Surgery*, *International Orthopaedics*, *The Journal of Bone and Joint Surgery*, *Journal of Neurosurgery: Spine (Phila Pa 1976)*, *Journal of the Pakistan Medical Association*, *Journal of Spinal Disorders and Techniques*, *Neurosurgical Focus*, *Orthopedics*, and *The Spine Journal*.

Study outcomes with levels of statistical significance displayed a  $P$  value of less than 0.05. Conversely, statistically nonsignificant data had a set  $P$  value greater than or equal to 0.05. A cross-tabulation was created to compare and categorize the FI of each dichotomous outcome event. The number of outcome events was modified to reverse significance. The FI is defined as the designated number that was required to reverse the significance of patient outcomes (Table 1). The FI was calculated and recorded for each outcome, event, and nonevent. The median FI that incorporated all outcomes was marked as the overall FI for this present study. The FQ was calculated by dividing the FI by the sample size. Characteristics of RCTs were split into subgroups for comparison. FI and FQ subgroups consisted of primary versus secondary outcomes, initial significance ( $P < 0.05$  vs.  $P \geq 0.05$ ), complications, comparing outcomes FI to LTF (FI < LTF vs. FI > LTF), and year (Table 2). Overall FI and FQ were measured by adding all outcome events. FI outcomes were calculated using the two-tailed Fisher exact test. Characteristics organized within each subgroup were calculated via interquartile ranges (IQRs) for FI and FQ. The IQR measures the difference between the upper quartile (Q3, 75th

**Table 1** Demonstration of reversal significance with a fragility of 1

	Outcome A	Outcome B	$P$ Value
Scenario 1			
Treatment A	13	27	<b>0.059</b>
Treatment B	5	35	
Scenario 2			
Treatment A	<b>14</b>	<b>26</b>	<b>0.034</b>
Treatment B	5	35	

**Table 2** Overall fragility data and analysis of subgroups<sup>a</sup>

Characteristic	Events	Fragility Index (IQR)	Fragility Quotient (IQR)
<b>All trials</b>	693	5 (3-7)	0.019 (0.011-0.043)
<b>Outcome significance<sup>b</sup></b>			
<i>P</i> < 0.05	130	5 (2-11)	0.017 (0.006-0.044)
<i>P</i> > 0.05	563	5 (4-7)	0.019 (0.011-0.043)
<b>Comparing outcome FI to LTF<sup>c</sup></b>			
FI < LTF	565	5 (3-7)	0.016 (0.009-0.033)
FI > LTF	128	5 (4-7)	0.055 (0.025-0.092)
<b>Year of publication</b>			
2000 – 2007	43	5 (4-5)	0.052 (0.043-0.073)
2008 – 2015	434	5 (3-7)	0.016 (0.009-0.036)
2016 – 2023	216	5 (3-10)	0.021 (0.011-0.044)
<b>Journals</b>			
Clinical Orthopaedics and Related Research	19	5 (4-7)	0.060 (0.048-0.084)
Clinical Spine Surgery	26	4.5 (4-8)	0.059 (0.040-0.133)
International Orthopaedics	20	4 (4-6)	0.054 (0.045-0.062)
Journal of Bone and Joint Surgery	62	5 (3-6)	0.017 (0.011-0.030)
Journal of Neurosurgery: Spine	156	6 (3-10)	0.025 (0.011-0.045)
Spine (Phila Pa 1976)	246	5 (3-7)	0.014 (0.009-0.024)
Journal of the Pakistan Medical Association	6	4.5 (4-7)	0.107 (0.095-0.167)
Journal of Spinal Disorders and Techniques	42	5 (5-9)	0.043 (0.011-0.067)
Neurosurgical Focus	18	4 (4-5)	0.073 (0.073-0.091)
Orthopedics	5	4 (4-5)	0.167 (0.167-0.208)
The Spine Journal	93	5 (3-8.5)	0.014 (0.008-0.028)

<sup>a</sup>FI, fragility index; IQR, interquartile range; LTF, lost to follow-up

<sup>b</sup>*P* < 0.05 represents the significant outcome subgroup and *P* > 0.05 represents the insignificant outcome subgroup

<sup>c</sup>FI < LTF represents the outcome subgroup where the FI was less than the number of patients LTF. FI > LTF represents the outcome subgroup where the FI was greater than the number of patients LTF

percentile) and lower quartile (Q1, 75th percentile). A risk-of-bias assessment was also performed (Table 3).

## Results

Of the 1561 studies screened, 111 met the search criteria with 35 RCTs included in the analysis (Fig. 1). A total of 693 events with 130 significant (*P* < 0.05) outcomes and 563 with nonsignificant (*P* > 0.05) outcomes were identified. The overall FI, incorporating all 693 outcome events from the 35 RCTs was 5 (IQR 3–7). The overall FQ was 0.019 (IQR 0.011–0.043), indicating that the reversal of 13 of 100 outcomes may change the study significance of the included RCTs. Of the 9 included RCTs, 3 studies reported loss to follow-up (LTF) data greater than the overall FI of 5. Therefore, 33.3% of studies reported an LTF value that was greater than the overall FI. For the 130 outcomes that were reported as significant, the median FI of events required to change significance was 5 (IQR 2–11) (Table 2). The FQ for significant outcomes was 0.017 (IQR 0.006–0.044). For the 563 outcomes that were reported as nonsignificant, the number of events required to change significance was 5 (IQR, 4–7). The FQ for nonsignificant outcomes was 0.019 (IQR

0.011–0.043). For the outcomes where FI ≤ LTF (*n* = 565), the median FI was found to be 5 (IQR 3–7). For the outcomes where FI > LTF (*n* = 128), the median FI was found to be 5 (IQR 4–7). The associated median FQs were 0.016 (IQR 0.009–0.033) and 0.055 (IQR 0.025–0.092), respectively. Fragility subanalysis per year of publication identified an FI of 5 (IQR 4–5) from 2000 to 2007, an FI of 5 (IQR 3–7) from 2008 to 2015, and an FI of 5 (IQR 3–10) from 2016 to 2023, thus demonstrating consistent statistical fragility over the 23 years.

## Discussion

Our fragility analysis exploring the cervical disc arthroplasty revealed an overall median FI of 5 (IQR 3–7) and a median FQ of 0.019 (IQR 0.011–0.043). A fragility index of 5 indicates that 5 events were needed to alter the statistical significance from significant to nonsignificant or vice versa. In correlation with sample size, the FQ of 0.019 demonstrates that 2 out of 100 patients is necessary for the median number to reverse significance across the total outcomes. In the RCTs included in this paper, 65.7% (23) of studies did not report or reported an LTF greater or equal to an overall

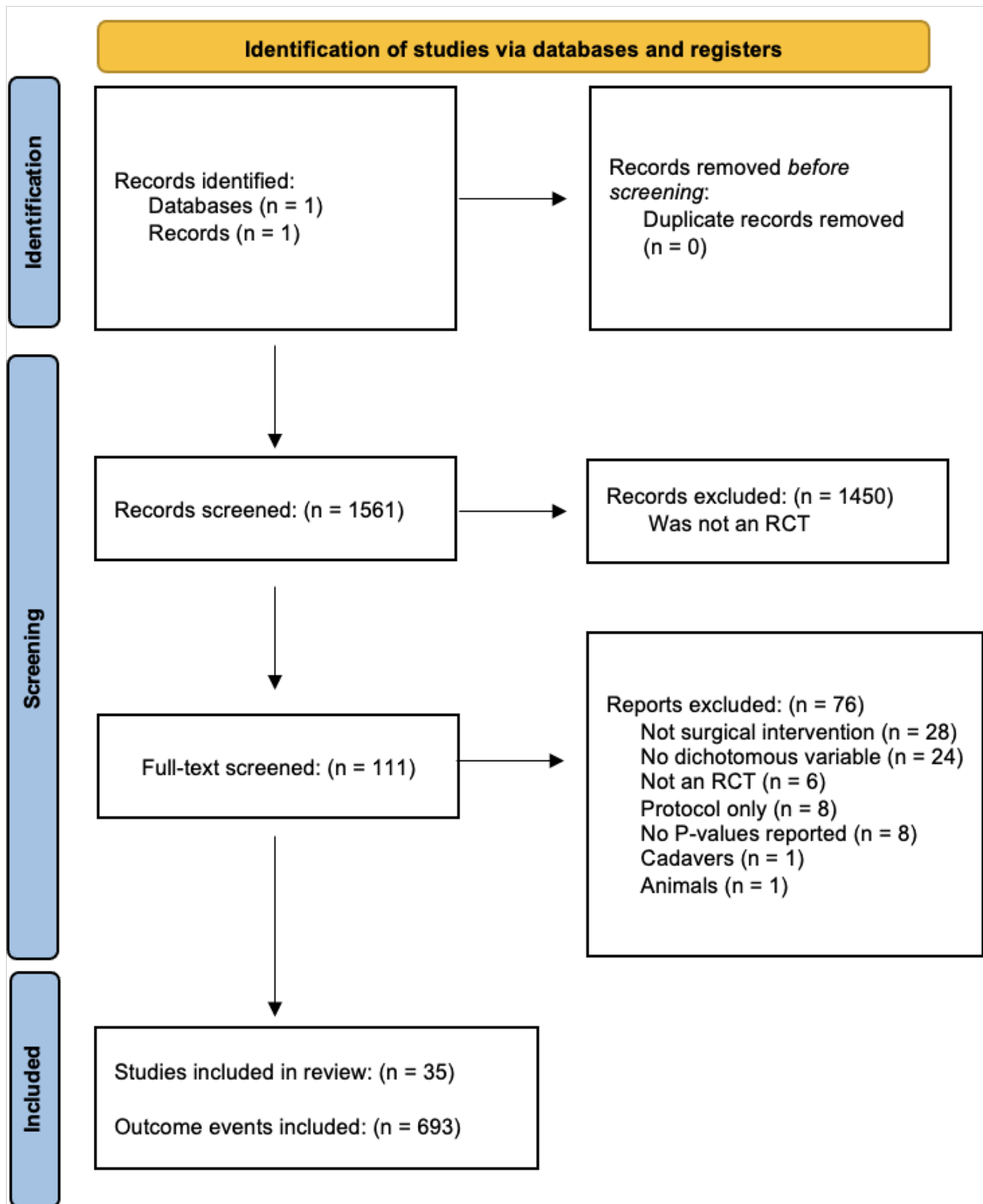


Fig. 1 PRISMA flow diagram of included studies

**Table 3** Risk of bias assessment

	Domain 1: Risk of Bias Arising from Randomization Process	Domain 2: Risk of bias due to deviations from the intended interventions	Domain 3: Risk of Bias Due to Missing Outcome Data	Domain 4: Risk of bias in measurement of the outcome	Domain 5: Risk of bias in selection of the reported result	Overall Risk of Bias
Loidolt et al. [29]	Low Risk	Low Risk	Low Risk	Low Risk	Low Risk	Low Risk
Lombardi et al. [30]	Low Risk	Low Risk	Unclear	Low Risk	Low Risk	Low Risk
MacDowall et al. [31]	Low Risk	Low Risk	Low Risk	Low Risk	Low Risk	Low Risk
Lavelle et al. [32]	Low Risk	Low Risk	Low Risk	Low Risk	Low Risk	Low Risk
Coric et al. [33]	Low Risk	Low Risk	Low Risk	Low Risk	Unclear	Low Risk
Rožanković et al. [34]	Low Risk	Low Risk	Low Risk	Low Risk	Low Risk	Low Risk
Gornet et al. [34]	Low Risk	Low Risk	Low Risk	Low Risk	Low Risk	Low Risk
Lanman et al. [35]	Low Risk	Low Risk	Low Risk	Low Risk	Low Risk	Low Risk
Pandey et al. [36]	Low Risk	Low Risk	Low Risk	Low Risk	Low Risk	Low Risk
Qizhi et al. [37]	Low Risk	Low Risk	Low Risk	Low Risk	Unclear	Low Risk
Radcliff et al. [38]	Low Risk	Low Risk	Unclear	Low Risk	Unclear	Low Risk
Janssen et al. [39]	Low Risk	Low Risk	Unclear	Low Risk	Unclear	Low Risk
Davis RJ et al. [40]	Low Risk	Low Risk	Unclear	Low Risk	Low Risk	Low Risk
Graham et al. [41]	Low Risk	Low Risk	Unclear	Low Risk	Low Risk	Low Risk
Phillips et al. [42]	Low Risk	Low Risk	Low Risk	Low Risk	Low Risk	Low Risk
Skeppholm et al. [43]	Low Risk	Low Risk	Low Risk	Low Risk	High Risk	Low Risk
Zhang et al. [44]	Low Risk	Low Risk	Low Risk	Low Risk	Low Risk	Low Risk
Karabag et al. [45]	Low Risk	Low Risk	Low Risk	Low Risk	Low Risk	Low Risk
Davis et al. [46]	Low Risk	Low Risk	Low Risk	Low Risk	Low Risk	Low Risk
Kang et al. [47]	Low Risk	Low Risk	Low Risk	Low Risk	Low Risk	Low Risk
Phillips et al. [2]	Low Risk	Low Risk	Low Risk	Low Risk	Low Risk	Low Risk
Vaccaro et al. [48]	Low Risk	Low Risk	Low Risk	Low Risk	Low Risk	Low Risk
Zigler et al. [49]	Low Risk	Low Risk	Low Risk	Low Risk	Unclear	Low Risk
Zhang et al. [50]	Low Risk	Low Risk	Low Risk	Low Risk	Low Risk	Low Risk
Cheng et al. [51]	Low Risk	Low Risk	Low Risk	Low Risk	Low Risk	Low Risk
Sasso et al. [7]	Low Risk	Low Risk	Low Risk	Low Risk	Low Risk	Low Risk
Garrido et al. [52]	Low Risk	Low Risk	Low Risk	Low Risk	Low Risk	Low Risk
Cheng et al. [53]	Low Risk	Low Risk	Low Risk	Low Risk	Low Risk	Low Risk
Heller et al. [3]	Low Risk	Low Risk	Low Risk	Low Risk	Low Risk	Low Risk
Phillips et al. [54]	Low Risk	Low Risk	Low Risk	Low Risk	Low Risk	Low Risk
Sasso et al. [55]	Low Risk	Low Risk	Low Risk	Low Risk	Low Risk	Low Risk
Anderson et al. [56]	Low Risk	Low Risk	Low Risk	Low Risk	Low Risk	Low Risk
Sasso et al. [57]	Low Risk	Low Risk	Low Risk	Low Risk	High Risk	Low Risk
Sasso et al. [58]	Low Risk	Low Risk	Low Risk	Low Risk	Low Risk	Low Risk
Porchet et al. [59]	Low Risk	Low Risk	Low Risk	Low Risk	Low Risk	Low Risk

FI of 5. Through this finding, LTF maintenance is an important tool that can help increase the overall robustness of a given study. The low median FI and FQ depict the statistical fragility within the CDA literature as only a small number of events is necessary to alter statistical significance. The application of a fragility analysis gives us an objective measure to compare clinical robustness where a higher FI would increase the confidence in the effectiveness of the data depicted in that study. This metric simply allows us to assess the quality of the studies that guide management algorithms in evidence-based medicine and accurately interpret clinical data that guide clinical care.

In this systematic review, the overall FI of 5 reported was similar to many of the previous spine and orthopedic

literature related to statistical fragility. Evaniew et al. were the first to apply a fragility index to the spine literature and reported an overall median FI of 2 [17]. In a recent study, Muthu et al. reexamined the spine literature and reported an overall median FI of 2 [18]. However, these studies focus on the overall spine literature and demonstrate significant fragility while our study highlights CDA literature to be more robust compared to the previous analyses. Gupta et al. examined the use of ketamine infusion during scoliosis surgery and reported a FI of 2. Cordero et al. calculated the statistical fragility of tourniquet use in total knee arthroplasty (TKA) and obtained an FI of 4 [19]. The anterior cruciate ligament reconstruction (ACLR) literature comparing bone-patellar tendon-bone grafts and hamstring tendon

autografts revealed a FI of 5 [20]. When looking at autograft vs. allograft comparison in ACLR, the authors found a FI of 6 [21]. A recent fragility analysis looking at calcaneus fractures demonstrated an FI of 6 as well [22]. In an examination of the distal biceps repair literature, the authors demonstrated an FI of 6.5 [23]. Megafu and Megafu in examining the distal radius fracture literature reported an FI of 9, the highest FI within all orthopedic and spine literature to date [24]. Other trauma literature looking at fibula fractures, distal femur fractures, and orbital fractures reported a similar FI of 5 across all studies [25–27]. By the spine and orthopedic literature, the FI consistently demonstrates statistical fragility, thus confirming our hypothesis and continual evidence of fragile results within the spine and orthopedic literature.

This review paper analyzing the spine literature regarding cervical disc arthroplasty has many strengths and limitations. An inherent strength in this review is the unbiased, comprehensive nature of having a systematic review of RCTs with the utilization of the PRISMA methodology. Also, the sole use of RCTs is a notable strength as management protocols and evidence-based medicine guidelines derive their decision-making abilities based on these studies. However, it is important to acknowledge this study's limitations. The first potential weakness is that fragility analysis is solely applicable to binary data with dichotomous endpoints. This limits FI utility for non-dichotomous studies. Studies with ordinal and continuous outcomes were filtered out to prevent overlap. Adding more study designs may have influenced the ability to further investigate CDA results. However, our team was able to comprehensively assess all possible binary outcomes by utilizing a dichotomous approach. Another limitation was that studies needed to be RCTs to be eligible for inclusion in this present review. As a result cohort studies, case-control studies, cross-sectional studies, prospective studies, and retrospective studies were excluded from analysis. Despite reduced incorporation of subject outcome events, our systematic review with RCTs provided a suitable, reliable way to assess effective CDA treatment options.

With the ever-expanding advancements in medicine, our systematic review aimed to quantify robustness in CDA literature by exploring the FI and FQ values to confirm the clinical significance of CDA. Our findings show that the majority of RCTs in cervical disc arthroscopic surgery have consistently been less robust for the last 23 years as speculated by our team. Possible implications for this loss of statistical significance may be due to solitary reliance on *P* value analysis, which inadvertently overlooks LTFs, impacts of varying sample size, and event outcomes [28]. Integrating fragility analysis in conjunction with *P* value analysis can be helpful when trying to carefully measure the

robustness of medical procedures. Due to the recent development of fragility analysis, it would be in the clinician's best interest to further support clinical data with additional statistical tools, such as the minimal clinically important difference (MCID), substantial clinical benefit (SCB), and patient-acceptable symptomatic state (PASS). These research tools can be used to leverage FI and FQ data to strengthen understanding of relationships and patterns for significant outcomes in respective trials.

## Conclusion

The literature regarding CDA RCTs lacks statistical robustness and may misrepresent the conclusions with the sole use of the *P* value. By implementing of the FI and FQ along with the *P* value, we believe the interpretation and contextualization of the clinical data surrounding CDA will be better understood.

**Author contributions** All authors contributed to the study's conception and design. Material preparation was performed by Michael Megafu MPH, Emmanuel C. Megafu, MS, and Janet T. Nguyen, MS. Data collection and analysis were performed by Michael Megafu MPH, Emmanuel C. Megafu, MS and Janet T. Nguyen, MS, and Elisabeth Point Du Jour, MS. The first draft was written by Emmanuel C. Megafu, MS and Michael Megafu, MPH. Multiple comments and edits were made by Wesley H. Bronson MD, James D. Lin MD, MS, Andrew C. Hecht MD, and Robert L. Parisien, MD. However, all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

**Funding** The authors declare that no funds, grants, or other support were received during the preparation of this manuscript. Open access funding provided by SCEL, Statewide California Electronic Library Consortium

## Declarations

**Ethical approval** Informed consent was not needed as this is a systematic review.

**Consent to participate** Not applicable.

**Consent to publish** Not applicable.

**Competing Interests** All authors declare they have no financial interests, except Robert L. Parisien, MD: The author, their immediate family, and any research foundation with which they are affiliated report the following potential conflicts: American Board of Orthopaedic Surgery: Diplomate; American Orthopaedic Society for Sports Medicine: Council of Delegates; American Orthopaedic Society for Sports Medicine: Research Committee; Arthroscopy Association of North America: Research Committee; Arthroscopy Association of North America: Surgical Skills Committee; Arthroscopy Association of North America: International Committee; Arthroscopy Association of North America: Social Media Committee; Eastern Orthopaedic Association: Committee Member; New England Orthopaedic Society: Committee Member; Society of Military Orthopaedic Surgeons: Committee Member; Jour-



nal of Cartilage & Joint Preservation: Associate Editor; Arthroscopy: Editorial or Governing Board; Arthroscopy, Sports Medicine and Rehabilitation: Editorial or Governing Board; Journal of Sport Rehabilitation: Editorial or Governing Board; Gotham Surgical Solutions and Devices: Grant/Education and Hospitality Financial or Material Support; Arthrex Inc.: Grant/Education and Hospitality Financial or Material Support.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Murray D, Janssen M, Delamarter R et al (2009) Results of the prospective, randomized, controlled multicenter Food and Drug Administration investigational device exemption study of the ProDisc-C total disc replacement versus anterior discectomy and fusion for the treatment of 1-level symptomatic cervical disc disease. *Spine J* 9:275–286. <https://doi.org/10.1016/j.spinee.2008.05.006>
- Phillips FM, Lee JYB, Geisler FH et al (2013) A prospective, randomized, controlled clinical investigation comparing PCM cervical disc arthroplasty with anterior cervical discectomy and fusion. 2-year results from the US FDA IDE clinical trial. *Spine* 38:E907–E918. <https://doi.org/10.1097/BRS.0b013e318296232f>
- Heller JG, Sasso RC, Papadopoulos SM et al (2009) Comparison of BRYAN cervical disc arthroplasty with anterior cervical decompression and fusion: clinical and radiographic results of a randomized, controlled, clinical trial. *Spine* 34:101–107. <https://doi.org/10.1097/BRS.0b013e31818ee263>
- Gornet MF, Burkus JK, Shaffrey ME et al (2015) Cervical disc arthroplasty with PRESTIGE LP disc versus anterior cervical discectomy and fusion: a prospective, multicenter investigational device exemption study. *J Neurosurg Spine* 23:558–573. <https://doi.org/10.3171/2015.1.SPINE14589>
- Radcliff K, Davis RJ, Hisey MS et al (2017) Long-term evaluation of cervical disc arthroplasty with the Mobi-C® cervical disc: a randomized, prospective, Multicenter Clinical Trial with seven-year follow-up. *Int J Spine Surg* 11:31. <https://doi.org/10.14444/4031>
- Delamarter RB, Murray D, Janssen ME et al (2010) Results at 24 months from the prospective, randomized, multicenter investigational device exemption trial of ProDisc-C versus anterior cervical discectomy and fusion with 4-year follow-up and continued access patients. *SAS J* 4:122–128. <https://doi.org/10.1016/j.esas.2010.09.001>
- Sasso RC, Anderson PA, Riew KD, Heller JG (2011) Results of cervical arthroplasty compared with anterior discectomy and fusion: four-year clinical outcomes in a prospective, randomized controlled trial. *J Bone Joint Surg Am* 93:1684–1692. <https://doi.org/10.2106/JBJS.J.00476>
- Coric D, Nunley PD, Guyer RD et al (2011) Prospective, randomized, multicenter study of cervical arthroplasty: 269 patients from the Kineflex®C artificial disc investigational device exemption study with a minimum 2-year follow-up: clinical article. *J Neurosurg Spine* 15:348–358. <https://doi.org/10.3171/2011.5.SPINE10769>
- Ahn P-G, Kim KN, Moon SW, Kim KS (2009) Changes in cervical range of motion and sagittal alignment in early and late phases after total disc replacement: radiographic follow-up exceeding 2 years. *J Neurosurg Spine* 11:688–695. <https://doi.org/10.3171/2009.7.SPINE0946>
- McAfee PC, Reah C, Gilder K et al (2012) A meta-analysis of comparative outcomes following cervical arthroplasty or anterior cervical fusion: results from 4 prospective multicenter randomized clinical trials and up to 1226 patients. *Spine* 37:943–952. <https://doi.org/10.1097/BRS.0b013e31823da169>
- Tsushima E (2022) Interpreting results from statistical hypothesis testing: understanding the appropriate P-value. *Phys Ther Res* 25:49–55. <https://doi.org/10.1298/ptr.R0019>
- Walsh M, Srinathan SK, McAuley DF et al (2014) The statistical significance of randomized controlled trial results is frequently fragile: a case for a Fragility Index. *J Clin Epidemiol* 67:622–628. <https://doi.org/10.1016/j.jclinepi.2013.10.019>
- Dettori JR, Norvell DC (2020) How fragile are the results of a trial? The Fragility Index. *Global Spine J* 10:940–942. <https://doi.org/10.1177/2192568220941684>
- Fay L-Y, Huang W-C, Tsai T-Y et al (2014) Differences between arthroplasty and anterior cervical fusion in two-level cervical degenerative disc disease. *Eur Spine J* 23:627–634. <https://doi.org/10.1007/s00586-013-3123-8>
- Parish JM, Coric D (2020) Cervical arthroplasty: long-term outcomes of FDA IDE trials. *Global Spine J* 10:61S–64S. <https://doi.org/10.1177/2192568219898154>
- Dufour T, Beaurain J, Huppert J et al (2019) Clinical and radiological evaluation of cervical disc arthroplasty with 5-year follow-up: a prospective study of 384 patients. *Eur Spine J* 28:2371–2379. <https://doi.org/10.1007/s00586-019-06069-z>
- Evaniew N, Files C, Smith C et al (2015) The fragility of statistically significant findings from randomized trials in spine surgery: a systematic survey. *Spine J* 15:2188–2197. <https://doi.org/10.1016/j.spinee.2015.06.004>
- Bryan SK, Moses RE (1988) Sufficiency of the Klenow fragment for survival of polC(Ts) pcbA1 *Escherichia coli* at 43 degrees C. *J Bacteriol* 170:456–458. <https://doi.org/10.1128/jb.170.1.456-458.1988>
- Cordero JK, Lawrence KW, Brown AN et al (2023) The fragility of Tourniquet Use in total knee arthroplasty: a systematic review of Randomized controlled trials. *J Arthroplasty* 38:1177–1183. <https://doi.org/10.1016/j.arth.2022.12.035>
- Lawrence KW, Okewunmi JO, Chakrani Z et al (2023) Randomized controlled trials comparing bone-patellar tendon-bone Versus Hamstring Tendon autografts in Anterior Cruciate Ligament Reconstruction surgery are statistically fragile: a systematic review. <https://doi.org/10.1016/j.arthro.2023.07.039>. *Arthroscopy*
- Megafu M, Megafu E, Mian H et al (2023) Fragile statistical findings in Randomized controlled trials evaluating Autograft Versus Allograft Use in Anterior Cruciate Ligament Reconstruction: a systematic review. <https://doi.org/10.1016/j.arthro.2023.07.055>. *Arthroscopy*
- Megafu M, Megafu E, Mian H et al (2023) The statistical fragility of outcomes in calcaneus fractures: a systematic review of randomized controlled trials. *Foot* 57:102047. <https://doi.org/10.1016/j.foot.2023.102047>
- Megafu MN, Mian HS, Hassan MM et al (2023) The fragility of statistical findings in distal biceps tendon repairs: a systematic review of randomized controlled trials. *J Shoulder Elb Surg* 32:e379–e386. <https://doi.org/10.1016/j.jse.2023.03.022>

24. Megafu M, Megafu E (2022) The fragility of statistical findings in distal Radius fractures: a systematic review of Randomized controlled trials. *Injury* 53:3352–3356. <https://doi.org/10.1016/j.injury.2022.07.017>
25. Mian H, Megafu M, Megafu E et al (2023) The statistical fragility of the distal fibula fracture literature: a systematic review of randomized controlled trials. <https://doi.org/10.1016/j.injury.2023.03.022>. *Injury*
26. Megafu M, Mian H, Megafu E et al (2023) The fragility of statistical significance in distal femur fractures: systematic review of randomized controlled trials. *Eur J Orthop Surg Traumatol* 33:2411–2418. <https://doi.org/10.1007/s00590-022-03452-3>
27. Megafu MN, Megafu EC, Nguyen JT et al (2023) The statistical fragility of Orbital fractures: a systematic review of Randomized controlled trials. *J Oral Maxillofac Surg* 81:752–758. <https://doi.org/10.1016/j.joms.2023.02.012>
28. Baduashvili A, Evans AT, Cutler T (2020) How to understand and teach P values: a diagnostic test framework. *J Clin Epidemiol* 122:49–55. <https://doi.org/10.1016/j.jclinepi.2020.03.003>
29. Loidolt T, Kurra S, Riew KD et al (2021) Comparison of adverse events between cervical disc arthroplasty and anterior cervical discectomy and fusion: a 10-year follow-up. *Spine J* 21:253–264. <https://doi.org/10.1016/j.spinee.2020.10.013>
30. Lombardi JM, Vivas AC, Gornet MF et al (2020) The Effect of ACDF or Arthroplasty on Cervicogenic headaches: a Post Hoc Analysis of a prospective, Multicenter Study with 10-Year follow-up. *Clin Spine Surg* 33:339–344. <https://doi.org/10.1097/BSD.0000000000001087>
31. MacDowall A, Canto Moreira N, Marques C et al (2019) Artificial disc replacement versus fusion in patients with cervical degenerative disc disease and radiculopathy: a randomized controlled trial with 5-year outcomes. *J Neurosurg Spine* 30:323–331. <https://doi.org/10.3171/2018.9.SPINE18659>
32. Lavelle WF, Riew KD, Levi AD, Florman JE (2019) Ten-year outcomes of cervical disc replacement with the BRYAN Cervical Disc: results from a prospective, randomized, controlled clinical trial. *Spine* 44:601–608. <https://doi.org/10.1097/BRS.0000000000002907>
33. Coric D, Guyer RD, Nunley PD et al (2018) Prospective, randomized multicenter study of cervical arthroplasty versus anterior cervical discectomy and fusion: 5-year results with a metal-on-metal artificial disc. *J Neurosurg Spine* 28:252–261. <https://doi.org/10.3171/2017.5.SPINE16824>
34. Rožanković M, Marasanov SM, Vukić M (2017) Cervical disk replacement with Discover Versus Fusion in a single-level cervical disk disease: a prospective single-Center Randomized Trial with a Minimum 2-Year follow-up. *Clin Spine Surg* 30:E515–E522. <https://doi.org/10.1097/BSD.0000000000000170>
35. Lanman TH, Burkus JK, Dryer RG et al (2017) Long-term clinical and radiographic outcomes of the Prestige LP artificial cervical disc replacement at 2 levels: results from a prospective randomized controlled clinical trial. *J Neurosurg Spine* 27:7–19. <https://doi.org/10.3171/2016.11.SPINE16746>
36. Pandey PK, Pawar I, Gupta J, Verma RR (2017) Comparison of outcomes of single-level Anterior Cervical Discectomy with Fusion and single-level Artificial cervical disc replacement for single-level cervical degenerative disc disease. *Spine* 42:E41–E49. <https://doi.org/10.1097/BRS.0000000000001696>
37. Qizhi S, Lei S, Peijia L et al (2016) A comparison of Zero-Profile devices and Artificial Cervical disks in patients with 2 noncontiguous levels of cervical spondylosis. *Clin Spine Surg* 29:E61–E66. <https://doi.org/10.1097/BSD.0000000000000096>
38. Radcliff K, Coric D, Albert T (2016) Five-year clinical results of cervical total disc replacement compared with anterior discectomy and fusion for treatment of 2-level symptomatic degenerative disc disease: a prospective, randomized, controlled, multicenter investigational device exemption clinical trial. *J Neurosurg Spine* 25:213–224. <https://doi.org/10.3171/2015.12.SPINE15824>
39. Janssen ME, Zigler JE, Spivak JM et al (2015) ProDisc-C total disc replacement Versus Anterior Cervical Discectomy and Fusion for single-level symptomatic cervical disc disease: seven-year follow-up of the Prospective Randomized U.S. Food and Drug Administration Investigational device exemption study. *J Bone Joint Surg Am* 97:1738–1747. <https://doi.org/10.2106/JBJS.N.01186>
40. Davis RJ, Nunley PD, Kim KD et al (2015) Two-level total disc replacement with Mobi-C cervical artificial disc versus anterior discectomy and fusion: a prospective, randomized, controlled multicenter clinical trial with 4-year follow-up results. *J Neurosurg Spine* 22:15–25. <https://doi.org/10.3171/2014.7.SPINE13953>
41. Graham RS, Samsell BJ, Proffer A et al (2015) Evaluation of glycerol-preserved bone allografts in cervical spine fusion: a prospective, randomized controlled trial. *J Neurosurg Spine* 22:1–10. <https://doi.org/10.3171/2014.9.SPINE131005>
42. Phillips FM, Geisler FH, Gilder KM et al (2015) Long-term outcomes of the US FDA IDE prospective, randomized controlled clinical trial comparing PCM cervical disc Arthroplasty with Anterior Cervical Discectomy and Fusion. *Spine* 40:674–683. <https://doi.org/10.1097/BRS.0000000000000869>
43. Skeppholm M, Lindgren L, Henriques T et al (2015) The Discover artificial disc replacement versus fusion in cervical radiculopathy—a randomized controlled outcome trial with 2-year follow-up. *Spine J* 15:1284–1294. <https://doi.org/10.1016/j.spinee.2015.02.039>
44. Zhang H-X, Shao Y-D, Chen Y et al (2014) A prospective, randomised, controlled multicentre study comparing cervical disc replacement with anterior cervical decompression and fusion. *Int Orthop* 38:2533–2541. <https://doi.org/10.1007/s00264-014-2497-5>
45. Karabag H, Cakmak E, Celik B et al (2014) Arthroplasty versus fusion for single-level cervical disc disease. *J Pak Med Assoc* 64:1348–1351
46. Davis RJ, Kim KD, Hisey MS et al (2013) Cervical total disc replacement with the Mobi-C cervical artificial disc compared with anterior discectomy and fusion for treatment of 2-level symptomatic degenerative disc disease: a prospective, randomized, controlled multicenter clinical trial: clinical article. *J Neurosurg Spine* 19:532–545. <https://doi.org/10.3171/2013.6.SPINE12527>
47. Kang L, Lin D, Ding Z et al (2013) Artificial disk replacement combined with midlevel ACDF versus multilevel fusion for cervical disk disease involving 3 levels. *Orthopedics* 36:e88–94. <https://doi.org/10.3928/01477447-20121217-24>
48. Vaccaro A, Beutler W, Poppelman W et al (2013) Clinical outcomes with selectively constrained SECURE-C cervical disc arthroplasty: two-year results from a prospective, randomized, controlled, multicenter investigational device exemption study. *Spine* 38:2227–2239. <https://doi.org/10.1097/BRS.0000000000000031>
49. Zigler JE, Delamarter R, Murrey D et al (2013) ProDisc-C and anterior cervical discectomy and fusion as surgical treatment for single-level cervical symptomatic degenerative disc disease: five-year results of a Food and Drug Administration study. *Spine* 38:203–209. <https://doi.org/10.1097/BRS.0b013e318278eb38>
50. Zhang X, Zhang X, Chen C et al (2012) Randomized, controlled, multicenter, clinical trial comparing BRYAN cervical disc arthroplasty with anterior cervical decompression and fusion in China. *Spine* 37:433–438. <https://doi.org/10.1097/BRS.0b013e31822699fa>
51. Cheng L, Nie L, Li M et al (2011) Superiority of the Bryan® disc prosthesis for cervical myelopathy: a randomized study with



- 3-year followup. *Clin Orthop Relat Res* 469:3408–3414. <https://doi.org/10.1007/s11999-011-2039-z>
52. Garrido BJ, Taha TA, Sasso RC (2010) Clinical outcomes of Bryan cervical disc arthroplasty a prospective, randomized, controlled, single site trial with 48-month follow-up. *J Spinal Disord Tech* 23:367–371. <https://doi.org/10.1097/BSD.0b013e3181bb8568>
  53. Cheng L, Nie L, Zhang L, Hou Y (2009) Fusion versus Bryan Cervical Disc in two-level cervical disc disease: a prospective, randomised study. *Int Orthop* 33:1347–1351. <https://doi.org/10.1007/s00264-008-0655-3>
  54. Phillips FM, Allen TR, Regan JJ et al (2009) Cervical disc replacement in patients with and without previous adjacent level fusion surgery: a prospective study. *Spine* 34:556–565. <https://doi.org/10.1097/BRS.0b013e31819b061c>
  55. Sasso RC, Best NM, Metcalf NH, Anderson PA (2008) Motion analysis of bryan cervical disc arthroplasty versus anterior discectomy and fusion: results from a prospective, randomized, multicenter, clinical trial. *J Spinal Disord Tech* 21:393–399. <https://doi.org/10.1097/BSD.0b013e318150d121>
  56. Anderson PA, Sasso RC, Riew KD (2008) Comparison of adverse events between the Bryan artificial cervical disc and anterior cervical arthrodesis. *Spine* 33:1305–1312. <https://doi.org/10.1097/BRS.0b013e31817329a1>
  57. Sasso RC, Smucker JD, Hacker RJ, Heller JG (2007) Clinical outcomes of BRYAN cervical disc arthroplasty: a prospective, randomized, controlled, multicenter trial with 24-month follow-up. *J Spinal Disord Tech* 20:481–491. <https://doi.org/10.1097/BSD.0b013e3180310534>
  58. Sasso RC, Smucker JD, Hacker RJ, Heller JG (2007) Artificial disc versus fusion: a prospective, randomized study with 2-year follow-up on 99 patients. *Spine* 32:2933–2940 discussion 2941–2. <https://doi.org/10.1097/BRS.0b013e31815d0034>
  59. Porchet F, Metcalf NH (2004) Clinical outcomes with the Prestige II cervical disc: preliminary results from a prospective randomized clinical trial. *Neurosurg Focus* 17:E6. <https://doi.org/10.3171/foc.2004.17.3.6>

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.