

# Statistical Fragility of Randomized Controlled Trials Evaluating Rehabilitation After Arthroscopic Rotator Cuff Repair

Sean B. Sequeira,<sup>\*†</sup> MD, Melissa A. Wright,<sup>†</sup> MD, and Anand M. Murthi,<sup>†</sup> MD

*Investigation performed at MedStar Union Memorial Hospital, Baltimore, Maryland, USA*

**Background:** Clinical decision-making often relies on evidence-based medicine, derived from objective data with conventional and rigorous statistical tests to evaluate significance. The literature surrounding rehabilitation after rotator cuff repair (RCR) is conflicting, with no defined standard of practice.

**Purpose:** To determine the fragility index (FI) and the fragility quotient (FQ) of randomized controlled trials (RCTs) evaluating rehabilitation protocols after RCR.

**Study Design:** Systematic review.

**Methods:** A systematic review was performed according to PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines by searching the PubMed, Cochrane Library, and Embase databases for RCTs evaluating rehabilitation protocols after arthroscopic RCRs from 2000 to June 1, 2022. The FI was determined by manipulating the dichotomous outcome events from each article until a reversal of significance with  $2 \times 2$  contingency tables was achieved. The FQ was determined by dividing the FI by the sample size.

**Results:** Fourteen RCTs with 48 dichotomous outcomes were ultimately included for analysis. The mean FI for the included dichotomous outcomes was 4 (interquartile range, 3-6), suggesting that the reversal of 4 events is required to change study significance. The mean FQ was 0.048. Of the RCTs that reported data regarding loss to follow-up, most studies (58.5%) indicated that  $>4$  patients had been lost to follow-up.

**Conclusion:** The results of RCT studies of RCR rehabilitation protocols are moderately fragile, something clinicians should be aware of when implementing study results into practice. We recommend the inclusion of FI and FQ in addition to standard P values when reporting statistical results in future RCTs with dichotomous outcome variables on this topic.

**Keywords:** P value; fragility index; fragility quotient; immobilization; rehabilitation; rotator cuff; statistical significance

\*Address correspondence to Sean B. Sequeira, MD, Department of Orthopaedic Surgery, MedStar Union Memorial Hospital, 3333 North Calvert Street, Suite 400, Baltimore, MD 21218, USA (email: sean.sequeira@medstar.net).

<sup>†</sup>Department of Orthopaedic Surgery, MedStar Union Memorial Hospital, Baltimore, Maryland, USA.

Final revision submitted February 20, 2023; accepted March 2, 2023.

One or more of the authors has declared the following potential conflict of interest or source of funding: M.A.W. has received education payments from Arthrex, Elite Orthopedics, Smith & Nephew, and Supreme Orthopedic Systems. A.M.M. has received education payments from Supreme Orthopedic Systems; consulting fees from Catalyst OrthoScience, DePuy/Medical Device Business Services, Globus Medical, Ignite Orthopedics, Stryker, and Zimmer Biomet; nonconsulting fees from Globus Medical; royalties from DePuy, Ignite Orthopedics, and Globus Medical; honoraria from Wright Medical; and has an investment interest in Ignite Orthopedics. AOSSM checks author disclosures against the Open Payments Database (OPD). AOSSM has not conducted an independent investigation on the OPD and disclaims any liability or responsibility relating thereto.

The Orthopaedic Journal of Sports Medicine, 11(7), 23259671231184946

DOI: 10.1177/23259671231184946

© The Author(s) 2023

Clinical decision-making in orthopaedic surgery is usually based on current literature, with randomized controlled trials (RCTs) that compare  $\geq 2$  interventions and evaluate a series of continuous and categorical outcomes, providing the highest level of evidence. The ideal rehabilitation protocol after rotator cuff repair (RCR) remains controversial among orthopaedic surgeons.<sup>11,20</sup> Large RCTs have sought to define a standard of care, with most studies comparing prolonged immobilization with early range of motion (ROM) protocols.<sup>25</sup> Systematic reviews and meta-analyses have also been published on rehabilitation after RCR with conflicting conclusions, suggesting that the optimal postoperative protocol remains unknown.<sup>23-25</sup> These systematic reviews are limited by the poor quality of included studies, mostly because of a small sample size of included studies, and heterogeneity among studies that can affect the strength of final conclusions. Despite their limitations, both RCTs and systematic reviews that evaluate the existing literature on rehabilitation after RCR often affect surgeons' practices. Therefore, the robustness of these conclusions, or

This open-access article is published and distributed under the Creative Commons Attribution - NonCommercial - No Derivatives License (<https://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits the noncommercial use, distribution, and reproduction of the article in any medium, provided the original author and source are credited. You may not alter, transform, or build upon this article without the permission of the Author(s). For article reuse guidelines, please visit SAGE's website at <http://www.sagepub.com/journals-permissions>.

lack thereof, should be better scrutinized and reported in a transparent fashion to help surgeons use the best evidence-based medicine.

The  $P$  value is an important metric, along with other metrics such as effect size, that RCTs use to test significance and justify the conclusions they draw. Most often, the  $\alpha$  value, or the chance that an alternative hypothesis found true is actually due to chance, is used, with statistical significance set at  $P < .05$ . Although statistical tests are imperative to help the surgeon draw conclusions from a study, the use of  $P$  values alone to ascribe significance may not optimize statistical rigor.<sup>32</sup> Because significance is usually assigned an otherwise arbitrary value of less than an  $\alpha$  of .05, outcomes sometimes require a reversal of only 1 to 2 events to change the significance of an outcome itself.<sup>2,17,34</sup> The fragility index (FI) is a relatively new concept, developed by Feinstein<sup>8</sup> in 1990 and used to characterize the stability (or fragility) of a given dichotomous outcome. The AAOS guidelines suggested that an FI of  $>2$  was considered statistically robust.<sup>9</sup> The FI is calculated by manipulating outcome events until a reversal of significance is achieved. A low FI signifies that the outcome is statistically fragile since it would require minimal manipulation of the outcome event to reverse significance.

To mitigate the shortcomings of FI and its independence of sample size, the fragility quotient (FQ) was developed not long after.<sup>36</sup> The FQ is calculated by dividing the FI by the sample size. Together, the FI and FQ can help augment RCTs' statistical reporting and better characterize each outcome's statistical stability. Several studies have commented on the fragility of the literature surrounding shoulder surgery and RCR.<sup>27-29</sup> However, none of the current published RCTs evaluating rehabilitation after RCR include fragility analysis (either FI or FQ). The lack of fragility analysis in the current literature on this controversial topic limits the confidence a surgeon can have in the robustness of the conclusions of these studies and whether to implement study recommendations into practice.

The purpose of this study was to analyze dichotomous outcomes in RCTs evaluating rehabilitation after RCR to determine the FI and the FQ of these trials. Our hypothesis was that the conclusions drawn regarding rehabilitation after RCR would be statistically fragile and support inclusion of FI and FQ in future RCTs on this topic.

## METHODS

### Search Strategy

This systematic review was performed according to PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines. Three databases, including Embase, Cochrane Library, and PubMed, were searched by 2 reviewers (S.B.S. and M.A.W.) up to June 1, 2022, using the search string "rehabilitation" OR "immobilization" AND "rotator cuff repair."

### Eligibility Criteria

To meet inclusion criteria, selected articles had to be a comparative RCT that evaluated outcomes for a rehabilitation intervention after RCR. They also had to evaluate at least 1 dichotomous outcome variable and report  $P$  values for that variable. Studies were excluded if they were in a language other than English or if the full text was not available. Data extraction from each study was performed independently by 2 reviewers (S.B.S. and M.A.W.) and reconciled by a third reviewer (A.M.M.).

The primary outcome of our study was the mean FI and FQ across all dichotomous outcome variables reported in the original study as statistically significant. The secondary outcomes were the FI and FQ examined separately for both significant and nonsignificant outcome variables.

### Statistical Analysis

The FI and FQ were calculated for all dichotomous outcome variables in the included RCTs. To calculate the FI and FQ, we recorded outcome events in a  $2 \times 2$  contingency table. Both significant and nonsignificant dichotomous outcomes were evaluated. The original  $P$  value was recorded for each outcome, and the Fisher exact test was used to verify the accuracy of the original, reported  $P$  value. Iterative manipulation of each outcome event was subsequently performed until a reversal of significance ( $P < .05$ ) was achieved. At this point, the number of events required for a reversal of significance was recorded as the FI. The FI of all dichotomous outcomes within included RCTs was calculated in an identical manner. The FQ was determined by taking each FI as a proportion of the total sample size. Means and interquartile ranges (IQRs) were computed for the FI and FQ of each outcome to better comment on the variability in the statistical fragility between the 25th and 75th percentiles.

Data regarding loss to follow-up were also evaluated for all studies by determining the sample size of patients who were initially included and then subsequently analyzed for each outcome. For example, if 450 patients were included in the study but only 437 returned to undergo imaging for a postoperative cuff tear, the number lost to follow-up was documented as 13. However, if 445 returned for their first postoperative visit and were evaluated for stiffness, the number lost to follow-up for stiffness was documented as 5.

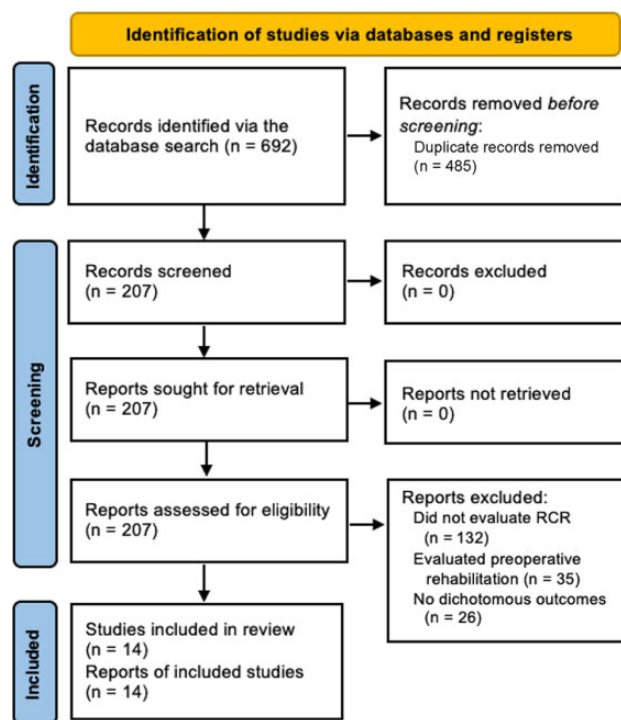
## RESULTS

### Descriptive Summary of Included RCTs

Of 692 RCTs that were identified from the PRISMA search, 14 RCTs\* met all inclusion criteria and were included in the statistical fragility analysis (Figure 1). The characteristics of the included studies are shown in Appendix Table A1. Of those RCTs that met inclusion criteria, 9<sup>†</sup> (64.3%) were classified as having level 1 evidence and 5<sup>3,6,12,21,26</sup>

\* References 3, 5, 10, 12, 15, 16, 18, 19, 21, 22, 26, 30, 31, 35.

†References 5, 10, 15, 16, 18, 22, 30, 31, 35.



**Figure 1.** PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines for systematic review of rehabilitation after rotator cuff repair (RCR).

(35.7%) as having level 2 evidence. Eleven studies<sup>‡</sup> (78.6%) performed an a priori power analysis, 2 studies<sup>18,31</sup> (14.3%) had no information regarding power analysis, and 1 study<sup>5</sup> conducted a post hoc power analysis. All included studies with a power analysis were found to be adequately powered. The mean sample size for the included RCTs was  $100.6 \pm 35.5$  patients. A total of 48 dichotomous outcomes from the included articles were evaluated for this study, with 7 initially reported as statistically significant and 41 as insignificant. Eighteen (37.5%) of the dichotomous outcomes evaluated were primary outcomes. The rate of retear at various time points was the most common dichotomous outcome across the 14 RCTs included within this study, with some studies including multiple retear outcomes (n = 16; 33.3%). Other commonly evaluated dichotomous outcomes included complication rates and progression to different stages of rotator cuff degeneration, as measured by classification systems such as the one by Sugaya et al.<sup>33</sup>

### Basic FI and FQ Analysis

The mean FI of the 48 dichotomous outcome events was 4 (IQR, 3-6). The mean FQ of the 48 dichotomous outcomes was 0.048 (IQR, 0.032-0.063). The mean FI of the 7 significant dichotomous outcome events was 2 (IQR, 1-3), and the FQ of significant outcome events was 0.036 (IQR, 0.025-0.050). The mean FI of the 41 insignificant dichotomous

outcome events was 5 (IQR, 4-6), and the FQ of insignificant outcome events was 0.05 (IQR, 0.034-0.067). The mean FI was 5 when evaluating outcomes involving retear (Table 1), and the mean FQ was 0.02.

### Analysis of Loss to Follow-up

Of the 48 outcome events across 14 RCTs, 7 outcome events in 3 studies<sup>3,5,35</sup> did not have data regarding loss to follow-up. These outcomes included bursitis,<sup>35</sup> echogenicity at 6 weeks and 3 months,<sup>35</sup> and 4 retears.<sup>5,30,35</sup> Of the outcome events with data regarding loss to follow-up, 24 of 41 (58.5%) had lost >4 patients.

### DISCUSSION

This study found that the conclusions in RCTs evaluating rehabilitation after RCR were moderately fragile, with a mean FI of 4 (IQR, 3-6) and FQ of 0.048. This finding suggests that on average, only 4 outcomes are required to reverse the significance for the dichotomous outcomes included within this study. Furthermore, for 58.5% of outcomes, the lost to follow-up numbers were >4, suggesting that had compliance with follow-up been maintained, conclusions drawn in these RCTs could have been different. Knowing the fragility of a study may influence clinicians' willingness to adopt study recommendations into practice. Future statistical reporting on this subject should include FI and FQ in addition to *P* values to provide clinicians with a more complete picture of the robustness of the data and aid in clinical decision-making.

Rehabilitation protocols after surgical RCR have been extensively studied in the literature, though the data are conflicting.<sup>23,25</sup> A recent systematic review on early versus delayed rehabilitation after surgical RCR concluded that although there was no difference between the 2 rehabilitation groups for most clinical outcomes and retear rates, the early rehabilitation group exhibited superior ROM compared with the late-rehabilitation group.<sup>25</sup> Another systematic review that evaluated 16 level I to II studies also found that there was no difference in functional or retear rates between early and delayed ROM, although external rotation was better in the early ROM cohort.<sup>23</sup> A third recent systematic review evaluating postoperative rehabilitation protocols found a possible benefit of better functional outcomes at the risk of increased retear rates, conflicting the aforementioned reviews.<sup>4</sup> These reviews and meta-analyses are limited by the quality of the data in the primary studies and also by the heterogeneity of the included studies. None of the previous studies, reviews, or meta-analyses assessed the statistical fragility of results, although the conflicting nature of the data and lack of consistent conclusions may suggest a lack of robustness to the data.

While the fragility of studies on rehabilitation after RCR has not been previously assessed, our findings that studies on this topic are statistically fragile are consistent with reviews that have evaluated statistical fragility of other orthopaedic literature. Khan et al<sup>13</sup> evaluated statistical fragility in the orthopaedic sports medicine literature and

<sup>‡</sup>References 3, 10, 12, 15, 16, 19, 21, 22, 26, 30, 35.

TABLE 1  
The FI and FQ for All Analyzed Outcomes (n = 48) From the 14 Included Randomized Controlled Trials<sup>a</sup>

Outcome	FI	Sample Size	FQ	Outcome	FI	Sample Size	FQ
Retear rate <sup>21</sup>	2	74	0.027	Goutallier 1 at 3 mo <sup>10</sup>	5	118	0.042
Retear rate <sup>12</sup>	5	116	0.043	Goutallier 2 at 3 mo <sup>10</sup>	3	118	0.025
Retear rate <sup>16</sup>	4	83	0.048	Goutallier 0 at 3 mo <sup>10</sup>	6	113	0.053
Bursitis <sup>35</sup>	7	80	0.089	Thomazeau 1 at 3 mo <sup>10</sup>	2	117	0.017
Echogenicity at 6 wk <sup>35</sup>	4	80	0.050	Thomazeau 2 at 3 mo <sup>10</sup>	4	118	0.033
Echogenicity at 12 wk <sup>35</sup>	5	80	0.063	Thomazeau 1 at 12 mo <sup>10</sup>	6	113	0.053
Retear rate <sup>35</sup>	4	80	0.050	Thomazeau 2 at 12 mo <sup>10</sup>	8	120	0.066
Full-thickness tear <sup>31</sup>	12	165	0.072	Sugaya 2 at 3 mo <sup>10</sup>	1	117	0.008
Infraspinatus tear <sup>31</sup>	7	165	0.042	Sugaya 3 at 3 mo <sup>10</sup>	2	114	0.0175
Atrophy at 6 mo <sup>31</sup>	6	165	0.036	Sugaya 2 at 12 mo <sup>10</sup>	3	112	0.026
Atrophy at 1 y <sup>31</sup>	7	165	0.042	Sugaya 3 at 12 mo <sup>10</sup>	2	110	0.0181
Reoperation rate <sup>31</sup>	3	165	0.018	Nonhealed rotator cuff <sup>10</sup>	6	117	0.051
Complication rate <sup>31</sup>	5	165	0.030	Small to medium retear <sup>5</sup>	2	30	0.066
Persistent pain <sup>31</sup>	4	165	0.024	Large retear <sup>5</sup>	1	18	0.055
Retear rate <sup>15</sup>	5	105	0.047	External rotation <20° <sup>3</sup>	3	92	0.032
Retear rate <sup>30</sup>	7	105	0.067	External rotation >30° <sup>3</sup>	4	92	0.043
Complication rate <sup>30</sup>	4	105	0.038	Adhesive capsulitis <sup>3</sup>	1	92	0.011
Retear rate <sup>6</sup>	4	68	0.058	Nonintact cuff <sup>3</sup>	7	82	0.085
Supraspinatus atrophy <sup>18</sup>	3	88	0.034	Recurrent tear <sup>3</sup>	4	82	0.048
Stiffness <sup>18</sup>	1	88	0.011	Completely healed rotator cuff <sup>3</sup>	3	82	0.036
Retear rate <sup>18</sup>	5	88	0.056	Retear <sup>22</sup>	4	50	0.080
Sugaya 1 <sup>18</sup>	7	88	0.079	Nonserious adverse events <sup>22</sup>	3	50	0.060
Sugaya 2 <sup>18</sup>	7	88	0.079	Serious adverse events <sup>22</sup>	5	50	0.100
Sugaya 3 <sup>18</sup>	6	88	0.068	Failure <sup>26</sup>	6	61	0.098

<sup>a</sup>FI, fragility index; FQ, fragility quotient.

found that, over a 10-year period, the mean FI of study outcomes was 2. A more recent study of the statistical fragility of the orthopaedic sports literature determined that the FI was 5.<sup>28</sup> Parisien et al<sup>27</sup> found that conclusions regarding the efficacy of platelet-rich plasma were statistically fragile, with a mean FI of 4 and FQ of 0.092. In this analysis, they also found that for about one-third of outcomes, the study had a lost to follow-up number greater than the FI, suggesting that had better follow-up been maintained, statistical significance and conclusions may have been reversed, assuming the outcomes of the patients lost to follow-up trended in the opposite direction from those who were evaluated. Even studies outside of orthopaedic surgery, including those in gynecologic surgery and cardiovascular research, have found similarly low FI and FQ, suggesting that the poor rigor of statistical reporting is not unique to orthopaedic surgery.<sup>7,29</sup> A study of journals with the highest impact factors, including the *New England Journal of Medicine* and *Lancet*, found that study conclusions were comparatively less fragile than what is found in other journals but ultimately still statistically fragile.<sup>14</sup> Despite focusing on higher-impact journals and more recent literature, our study still found that the literature comparing rotator cuff rehabilitation protocols is quite fragile.

This study is unique in demonstrating the fragility of a specific and important practice among shoulder and elbow surgeons that has not been previously studied, to our knowledge. By examining the literature on rehabilitation after RCR in this manner, the included previously published RCTs can all be better interpreted. FI and FQ add

information beyond the published *P* values that can help clinicians better interpret the robustness of study results and determine whether conclusions should be incorporated into clinical practice.

Based on the moderate statistical fragility found in this analysis, we recommend that future RCTs examining rehabilitation after RCR tailor their study design and statistical analysis to incorporate FI and FQ. As suggested by a previous fragility study in the literature, RCTs with a larger sample size and greater power will inherently produce higher FI and FQ, optimizing their statistical rigor and the strength of the subsequent conclusions of the study.<sup>1</sup> We posit that the consistent and regular reporting of FI and FQ in tandem with *P* values, as well as larger sample sizes and greater power in future RCTs, will help to specifically address previous deficiencies in the literature and determine a gold standard for rehabilitation after arthroscopic RCR. FI and FQ provide physicians who review literature another way to critically examine the significance of findings and gauge the clinical relevance of each trial with respect to patient care. Although the results of this analysis are related to rehabilitation after RCR, the concepts of FI and FQ can and should be broadly applied to other facets of the orthopaedic surgery literature to enhance the critical examination of RCT findings to best inform future clinical practice.

### Limitations

Although, to our knowledge, this study is the first of its kind to evaluate the statistical fragility of conclusions drawn

regarding rehabilitation after RCR, it does have its limitations. First, the inclusion of only high-impact orthopaedic and physical medicine and rehabilitation, while intentional, may have excluded RCTs that would have otherwise fit inclusion criteria. In addition, the concept of FI has intrinsic limitations itself. FI is a stand-alone value with no prescribed threshold to indicate fragility or stability of the study in question and, moreover, does not incorporate the study's sample size into consideration. FQ was introduced to mitigate some of these limitations, but even FQ still is limited by no true threshold to confer fragility or lack thereof. Also, this study did not include an evaluation of basic patient demographic factors that could influence outcomes in this analysis of rehabilitation outcomes after RCR. Finally, only dichotomous outcomes were included in the analysis of fragility. The inability to assess the fragility of continuous outcome variables limits the generalizability of the study findings.

## CONCLUSION

The results of RCT studies of RCR rehabilitation protocols are moderately fragile, something clinicians should be aware of when implementing study results into practice. We recommend the inclusion of FI and FQ in addition to standard *P* values when reporting statistical results in future RCTs on this topic.

## REFERENCES

- Ahmed W, Fowler RA, McCredie VA. Does sample size matter when interpreting the fragility index? *Crit Care Med*. 2016;44(11):e1142-e1143.
- Andrade C. The *P* value and statistical significance: misunderstandings, explanations, challenges, and alternatives. *Indian J Psychol Med*. 2019;41(3):210-215.
- Arndt J, Clavert P, Mielcarek P, et al. Immediate passive motion versus immobilization after endoscopic supraspinatus tendon repair: a prospective randomized study. *Orthop Traumatol Surg Res*. 2012;98(6)(suppl):S131-S138.
- Bandara U, An VVG, Imani S, Nandapalan H, Sivakumar BS. Rehabilitation protocols following rotator cuff repair: a meta-analysis of current evidence. *ANZ J Surg*. 2021;91(12):2773-2779.
- Chou CT, Hu W, Wen CS, Wang SF, Lieu FK, Teng JT. Efficacy of informed versus uninformed physiotherapy on postoperative retear rates of medium-sized and large rotator cuff tears. *J Shoulder Elbow Surg*. 2015;24(9):1413-1420.
- Cuff DJ, Pupello DR. Prospective randomized study of arthroscopic rotator cuff repair using an early versus delayed postoperative physical therapy protocol. *J Shoulder Elbow Surg*. 2012;21(11):1450-1455.
- Ennezat PV, Alavi Z, Le Jemtel TH, Hansen MR. Consideration regarding the analysis of randomized controlled trials in the era of evidence-based medicine. *J Cardiovasc Pharmacol*. 2022;79(5):605-619.
- Feinstein AR. The unit fragility index: an additional appraisal of "statistical significance" for a contrast of two proportions. *J Clin Epidemiol*. 1990;43(2):201-209.
- Forrester LA, McCormick KL, Bonsignore-Opp L, et al. Statistical fragility of surgical clinical trials in orthopaedic trauma. *JAAOS Global Res Rev*. 2021;5(11):e20.00197.
- Jenssen KK, Lundgreen K, Madsen J, Kvakestad R, Pripp A, Dimmen S. No Functional difference between three and six weeks of immobilization after arthroscopic rotator cuff repair: a prospective randomized controlled non-inferiority trial. *Arthroscopy*. 2018;34(10):2765-2774.
- Jung C, Tepohl L, Tholen R, et al. Rehabilitation following rotator cuff repair. *Oberer Extrem*. 2018;13(1):45-61.
- Keener JD, Galatz LM, Stobbs-Cucchi G, Patton R, Yamaguchi K. Rehabilitation following arthroscopic rotator cuff repair: a prospective randomized trial of immobilization compared with early motion. *J Bone Joint Surg Am*. 2014;96(1):11-19.
- Khan M, Evaniew N, Gichuru M, et al. The fragility of statistically significant findings from randomized trials in sports surgery: a systematic survey. *Am J Sports Med*. 2017;45(9):2164-2170.
- Khan MS, Fonarow GC, Friede T, et al. Application of the reverse fragility index to statistically nonsignificant randomized clinical trial results. *JAMA Netw Open*. 2020;3(8):e2012469.
- Kim YS, Chung SW, Kim JY, Ok JH, Park I, Oh JH. Is early passive motion exercise necessary after arthroscopic rotator cuff repair? *Am J Sports Med*. 2012;40(4):815-821.
- Kjær BH, Magnusson SP, Henriksen M, et al. Effects of 12 weeks of progressive early active exercise therapy after surgical rotator cuff repair: 12 weeks and 1-year results from the CUT-N-MOVE randomized controlled trial. *Am J Sports Med*. 2021;49(2):321-331.
- Kjaer P, Leboeuf-Yde C, Korsholm L, Sorensen JS, Bendix T. Magnetic resonance imaging and low back pain in adults: a diagnostic imaging study of 40-year-old men and women. *Spine*. 2005;30(10):1173-1180.
- Koh KH, Lim TK, Shon MS, Park YE, Lee SW, Yoo JC. Effect of immobilization without passive exercise after rotator cuff repair: randomized clinical trial comparing four and eight weeks of immobilization. *J Bone Joint Surg Am*. 2014;96(6):e44.
- Koh YG, Kim SJ, Chun YM, Kim YC, Park YS. Arthroscopic treatment of patellofemoral soft tissue impingement after posterior stabilized total knee arthroplasty. *Knee*. 2008;15(1):36-39.
- Kokmeyer D, Dube E, Millett PJ. Prognosis driven rehabilitation after rotator cuff repair surgery. *Open Orthop J*. 2016;10:339-345.
- Lee BG, Cho NS, Rhee YG. Effect of two rehabilitation protocols on range of motion and healing rates after arthroscopic rotator cuff repair: aggressive versus limited early passive exercises. *Arthroscopy*. 2012;28(1):34-42.
- Littlewood C, Bateman M, Butler-Walley S, et al. Rehabilitation following rotator cuff repair: a multi-centre pilot & feasibility randomised controlled trial (RaCeR). *Clin Rehabil*. 2021;35(6):829-839.
- Longo UG, Risi Ambrogioni L, Berton A, et al. Conservative versus accelerated rehabilitation after rotator cuff repair: a systematic review and meta-analysis. *BMC Musculoskelet Disord*. 2021;22:637.
- Matlak S, Andrews A, Looney A, Tepper KB. Postoperative rehabilitation of rotator cuff repair: a systematic review. *Sports Med Arthrosc Rev*. 2021;29(2):119-129.
- Mazuquin B, Moffatt M, Gill P, et al. Effectiveness of early versus delayed rehabilitation following rotator cuff repair: systematic review and meta-analyses. *PLoS One*. 2021;16(5):e0252137.
- Mazzocca AD, Arciero RA, Shea KP, et al. The effect of early range of motion on quality of life, clinical outcome, and repair integrity after arthroscopic rotator cuff repair. *Arthroscopy*. 2017;33(6):1138-1148.
- Parisien RL, Ehlers C, Cusano A, Tornetta P, Li X, Wang D. The statistical fragility of platelet-rich plasma in rotator cuff surgery: a systematic review and meta-analysis. *Am J Sports Med*. 2021;49(12):3437-3442.
- Parisien RL, Trofa DP, Dashe J, et al. Statistical fragility and the role of *P* values in the sports medicine literature. *J Am Acad Orthop Surg*. 2019;27(7):e324-e329.
- Pascoal E, Liu M, Lin L, Luketic L. The fragility of statistically significant results in gynecologic surgery: a systematic review. *J Obstet Gynaecol Can*. 2022;44(5):508-514.
- Sheps DM, Bouliane M, Styles-Tripp F, et al. Early mobilisation following mini-open rotator cuff repair: a randomised control trial. *Bone Joint J*. 2015;97(9):1257-1263.
- Sheps DM, Silveira A, Beaupre L, et al. Early active motion versus sling immobilization after arthroscopic rotator cuff repair: a randomized controlled trial. *Arthroscopy*. 2019;35(3):749-760.e2.

32. Solla F, Tran A, Bertonecchi D, Musoff C, Bertonecchi CM. Why a P-value is not enough. *Clin Spine Surg.* 2018;31(9):385-388.
33. Sugaya H, Maeda K, Matsuki K, Moriishi J. Functional and structural outcome after arthroscopic full-thickness rotator cuff repair: single-row versus dual-row fixation. *Arthroscopy.* 2005;21(11):1307-1316.
34. Sullivan GM, Feinn R. Using effect size—or why the P value is not enough. *J Grad Med Educ.* 2012;4(3):279-282.
35. Tirefort J, Schwitzguebel AJ, Collin P, Nowak A, Plomb-Holmes C, Lädermann A. Postoperative mobilization after superior rotator cuff repair: sling versus no sling: a randomized prospective study. *J Bone Joint Surg Am.* 2019;101(6):494-503.
36. Vargas M, Marra A, Buonanno P, Coviello A, Iacovazzo C, Servillo G. Fragility index and fragility quotient in randomized controlled trials on corticosteroids in ARDS due to COVID-19 and non-COVID-19 etiology. *J Clin Med.* 2021;10(22):5287.

## APPENDIX

TABLE A1  
Details of the Included Studies<sup>a</sup>

Lead Author (Year)	LOE	Sample Size	Dichotomous Outcomes Included in Analysis
Lee (2012) <sup>21</sup>	2	74	Retear rates
Chou (2015) <sup>5</sup>	1	24	Small to medium retears, large retears
Jenssen (2018) <sup>10</sup>	1	120	Goutallier 0, Goutallier 1, Goutallier 2, Thomazeau 1 (early), Thomazeau 2 (early), Thomazeau 1 (late), Thomazeau 2 (late), Sugaya 2, Sugaya 3, Nonhealed rotator cuff
Keener (2014) <sup>12</sup>	2	114	Retear rates
Kim (2012) <sup>15</sup>	1	105	Retear rates
Kjær (2021) <sup>16</sup>	1	82	Retear rates
Koh (2014) <sup>18</sup>	1	100	Supraspinatus atrophy, stiffness, full-thickness re-tear, Sugaya 1, Sugaya 2, Sugaya 3
Littlewood (2021) <sup>22</sup>	1	73	Retear rate, nonserious adverse event, serious adverse event
Mazzocca (2017) <sup>26</sup>	2	73	Failure
Sheps (2015) <sup>30</sup>	1	189	Retear rate, complication rate
Sheps (2019) <sup>31</sup>	1	206	Full-thickness tear, infraspinatus tears, infraspinatus atrophy (early), infraspinatus atrophy (late), reoperation rate, complication rate, persistent pain
Tirefort (2019) <sup>35</sup>	1	80	Bursitis, echogenicity (early), echogenicity (late), re-tear
Arndt (2012) <sup>3</sup>	2	100	External rotation <20°, external rotation >30°, adhesive capsulitis, nonintact cuff, recurrent tear, complete healing
Cuff (2012) <sup>6</sup>	2	68	Retear rates

<sup>a</sup>LOE, level of evidence.