

# UC Irvine

## UC Irvine Previously Published Works

### Title

Statistical Fragility of Single-Row Versus Double-Row Anchoring for Rotator Cuff Repair: A Systematic Review of Comparative Studies

### Permalink

<https://escholarship.org/uc/item/7qv64700>

### Journal

Orthopaedic Journal of Sports Medicine, 10(5)

### ISSN

2325-9671

### Authors

Fackler, Nathan P  
Ehlers, Cooper B  
Callan, Kylie T  
et al.

### Publication Date

2022-05-01

### DOI

10.1177/23259671221093391

### Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at <https://creativecommons.org/licenses/by/4.0/>

Peer reviewed

# Statistical Fragility of Single-Row Versus Double-Row Anchoring for Rotator Cuff Repair

## A Systematic Review of Comparative Studies

Nathan P. Fackler,<sup>\*†</sup> MS, Cooper B. Ehlers,<sup>‡</sup> MD, Kylie T. Callan,<sup>\*</sup> BS, Arya Amirhekmat,<sup>\*</sup> BS, Eric J. Smith,<sup>\*</sup> MD, Robert L. Parisien,<sup>§</sup> MD, and Dean Wang,<sup>\*||</sup> MD

*Investigation performed at University of California, Irvine, Irvine, California, USA*

**Background:** Comparative studies and randomized controlled trials (RCTs) often use the *P* (probability) value to convey the statistical significance of their findings. *P* values are an imperfect measure, however, and are vulnerable to a small number of outcome reversals to alter statistical significance. The inclusion of a fragility index (FI) and fragility quotient (FQ) may aid in the interpretation of a study's statistical strength.

**Purpose/Hypothesis:** The purpose of this study was to examine the statistical stability of studies comparing single-row to double-row rotator cuff repair. It was hypothesized that the findings of these studies would be vulnerable to a small number of outcome event reversals, often fewer than the number of patients lost to follow-up.

**Study Design:** Systematic review; Level of evidence, 3.

**Methods:** We analyzed comparative studies and RCTs on primary single-row versus double-row rotator cuff repair that were published between 2000 and 2021 in 10 leading orthopaedic journals. Statistical significance was defined as a  $P < .05$ . The FI for each outcome was determined by the number of event reversals necessary to alter significance. The FQ was calculated by dividing the FI by the respective sample size.

**Results:** Of 4896 studies screened, 22 comparative studies, 10 of which were RCTs, were ultimately included for analysis. A total of 74 outcomes were examined. Overall, the median FI was 2 (interquartile range [IQR], 1-3), and the median FQ was 0.035 (IQR, 0.020-0.057). The mean FI was  $2.55 \pm 1.29$ , and the mean FQ was  $0.043 \pm 0.027$ . In 64% of outcomes, the FI was less than the number of patients lost to follow-up.) Additionally, 81% of significant outcomes needed just a single outcome reversal to lose their significance.

**Conclusion:** Over half of the studies currently used to guide clinical practice have a number of patients lost to follow-up greater than their FI. The results of these studies should be interpreted within the context of these limitations. Future analyses may benefit from the inclusion of the FI and the FQ in their statistical analyses.

**Keywords:** fragility index; fragility quotient; statistical significance; *P* value; rotator cuff; single row; double row

As the number of patients who undergo arthroscopic rotator cuff repair continues to rise, clinical research has focused on delineating the best possible technique for repair. Many anatomic and biomechanical factors are considered when comparing different techniques, including restoration of the anatomic footprint,<sup>2</sup> biomechanical strength of the construct,<sup>30</sup> and minimized motion at the bone-tendon interface.<sup>18</sup> Currently, the most frequently

used techniques are the single-row (SR) and double-row (DR) repairs.<sup>9</sup> However, the designations "single row" and "double row" are inherently broad and encompass techniques such as the modified Mason-Allen (MMA) stitch, triple-loaded anchor repair (TLAR), and classic SR techniques for SR repairs, and the transosseous equivalent (TOE) and classic DR techniques for DR repairs.<sup>30</sup>

Despite the wide range of techniques described in the literature, no consensus currently exists as to what should be the standard repair technique, and comparative studies between SR and DR constructs are ongoing.<sup>9</sup> Recently, the literature has suggested that DR repair may offer a more

The Orthopaedic Journal of Sports Medicine, 10(5), 23259671221093391  
DOI: 10.1177/23259671221093391  
© The Author(s) 2022

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>.

biomechanically sound construct, but it is more costly and time demanding than is SR repair.<sup>2,9</sup> The surgeon's choice between SR and DR repair is guided primarily by evidence-based medicine, with an emphasis placed on these comparative studies and randomized controlled trials (RCTs). However, the statistical stability of these studies is subject to question and may be more fragile than previously thought.

The significance of the conclusions reached in comparative studies and RCTs is often portrayed using statistical thresholds such as the *P* (probability) value. The *P* value is a useful tool for distilling the results of a trial down to either significant or nonsignificant; however, it is limited and fails to consider important aspects of a trial, such as sample size and number of events taking place.<sup>46</sup> Statistical significance, as defined by the *P* value, can be altered by a small number of event reversals. Furthermore, if the number of event reversals needed to change the significance of the result is smaller than the number of patients lost to follow-up, this biases the study.<sup>46</sup> To account for the shortcomings of the *P* value and aid in the statistical interpretation of a study, Feinstein<sup>11</sup> introduced the concept of the fragility index (FI). The FI is defined as the number of event reversals needed to change the findings of a study from significant to nonsignificant and vice versa. This value adds to the strength of the study; however, it does not consider the study population size. To control for this, Ahmed et al<sup>1</sup> proposed the fragility quotient (FQ), which is calculated by dividing the FI by the total sample size to arrive at a value of relative fragility for the study. A larger FQ indicates a more stable *P* value, as an FI of 5 in a study population of 50 (FQ, 0.100) is substantially more stable than an FI of 5 in a study population of 5000 (FQ, 0.001). The inclusion of both the FI and the FQ allows for the reporting of absolute (FI) and relative (FQ) fragility of a study and would therefore contribute greatly to the interpretation of comparative studies; however, they are not often reported.

The purpose of this study was to determine the statistical stability of comparative studies and RCTs evaluating SR versus DR repair for rotator cuff tears by calculating the FI and FQ of the included studies. The secondary objective was to perform a subgroup analysis to determine the proportion of studies for which the fragility was smaller than the number of patients lost to follow-up. We hypothesized that the number of event outcome reversals needed to alter the significance of the study would be fewer than the number of patients lost to follow-up for a majority of the studies included in the analysis.

## METHODS

Comparative studies and RCTs focusing on the utilization of the SR and DR techniques for rotator cuff tears published in select journals between 2000 and 2021 were identified and collected. Ten journals were selected for their prominence within the field of orthopaedic surgery and shoulder and elbow surgery. According to the 2020 InCites Journal Citation Reports index ([https://www.researchgate.net/publication/347504391\\_InCites\\_Journal\\_Citation\\_Reports\\_2020\\_with\\_Impact\\_Factors\\_and\\_Quartiles\\_included](https://www.researchgate.net/publication/347504391_InCites_Journal_Citation_Reports_2020_with_Impact_Factors_and_Quartiles_included)), these journals are recognized as the most influential in the field of orthopaedic and shoulder and elbow surgery. The 10 orthopaedic journals included and their impact factors were the *British Journal of Sports Medicine* (12.022), *American Journal of Sports Medicine* (5.810), *Journal of Bone and Joint Surgery–American* (4.578), *Clinical Orthopaedics and Related Research* (4.329), *Arthroscopy* (4.325), *Bone & Joint Journal* (4.306), *Acta Orthopaedica* (2.965), *International Orthopaedics* (2.854), *Journal of Shoulder and Elbow Surgery* (2.817), and *Journal of Orthopaedic Research* (2.728).

Studies from these journals were queried for relevance, screened, and ultimately included or excluded based on criteria symmetric with the current surgical trends for rotator cuff repair. The initial PubMed search criteria included all studies on the rotator cuff, SR repair, and DR repair. This search was then limited to include only comparative studies and RCTs. The titles and abstracts of these comparative studies and RCTs were then screened for the utilization of the SR versus DR technique. Each article was then examined, and studies were included if (1) SR or DR repair for rotator cuff tear was implemented; (2) the patients underwent primary rotator cuff repair for small, medium, or large/massive tears; and (3) the study reported a 12-month minimum follow-up period. The studies were excluded if (1) the surgical technique was not explicitly described or referenced; (2) the study focused on something other than clinical outcomes (cost, operative time, blood flow, etc); (3) the patients underwent revision rotator cuff repair; (4) the study included patients undergoing a procedure other than SR or DR repair; (5) the studies were cadaveric, in vitro, or animal studies; (6) the study used population databases, national registries, or cross-sectional data; or (7) no dichotomous outcomes were reported anywhere in the study. From the studies meeting these criteria, all categorical outcomes were included (Figure 1). Nondichotomous data points were not included, as

<sup>||</sup>Address correspondence to Dean Wang, MD, 101 The City Drive South, Pavilion III, Building 29A, Orange, CA 92686, USA (email: deanwangmd@gmail.com) (Twitter: @deanwangmd).

<sup>\*</sup>Department of Orthopaedic Surgery, University of California, Irvine, Irvine, California, USA.

<sup>†</sup>Georgetown University School of Medicine, Washington, DC, USA.

<sup>‡</sup>Department of Orthopaedic Surgery, University of California, San Diego, San Diego, California, USA.

<sup>§</sup>Department of Orthopaedic Surgery, Mount Sinai, New York, New York, USA.

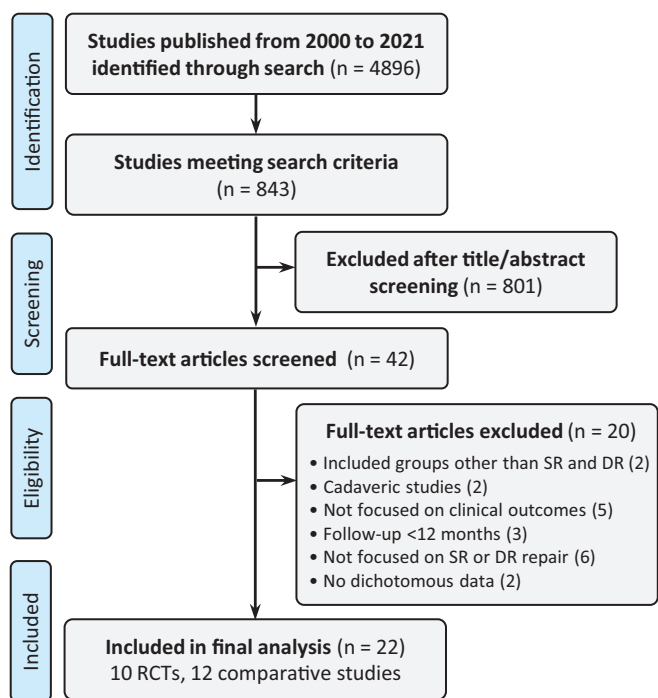
Final revision submitted January 6, 2022; accepted February 17, 2022.

One or more of the authors has declared the following potential conflict of interest or source of funding: R.L.P. has received grant support and education payments from Arthrex. D.W. has received education payments from Arthrex and Smith & Nephew; consulting fees from Newclip, Ipsen Biosciences, DePuy/Medical Device Business Services, and Vericel; and hospitality payments from Linvatec and Stryker. 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.

these are unable to be analyzed using current fragility methodology.

We extracted data involved in the dichotomous outcomes of each study, which included the number of patients in each outcome group, the outcome being measured, the total population size, and the number of patients lost to follow-up. The reported *P* value associated with each dichotomous outcome measure was recorded and verified for accuracy using a Fisher exact test. Statistical significance was set at *P* < .05.

Using a contingency table, we manipulated the results of the outcomes until the significance was reversed. For example, if the *P* value of a certain outcome was reported as < .05 by a study, the number of outcome reversals needed to increase the *P* value to > .05 was determined and vice versa (Figure 2). FI was recorded as the number of outcome reversals needed to change the significance of the study. FQ was determined by dividing the FI by the respective sample size,



**Figure 1.** Study inclusion flowchart. DR, double row; RCT, randomized controlled trial; SR, single row.

and interquartile ranges (IQRs) were calculated. Studies in which FI was less than the number of patients lost to follow-up were identified.

Four subgroups were analyzed for significant differences via independent *t* tests at 95% confidence: (1) significant (*P* < .05) versus nonsignificant (*P* > .05) outcomes, (2) outcomes for which the FI was fewer than the number of patients lost to follow-up versus outcomes for which the FI was greater than the number of patients lost to follow-up, (3) outcomes between traditional DR repair and its variants (eg, TOE), and (4) outcomes between traditional SR repair and its variants (eg, MMA, TLAR). All data analysis was performed using Microsoft Excel (Version 16.37; Microsoft Corp).

**RESULTS**

Of the 4896 studies identified, 843 comparative studies were screened. Ultimately, 22 studies, including 10 RCTs, were included for the analysis. The characteristics of the included studies can be found in Table 1. Overall, the median FI was 2 (IQR 1-3) and the median FQ was 0.035 (IQR 0.020-0.057). The overall mean FI was 2.55 ± 1.29 and the mean FQ was 0.043 ± 0.027.

We performed a subgroup analysis in which we examined groups reporting statistically significant outcomes (*P* < .05), outcomes in which number of patients lost to follow-up was less than the FI, outcomes from variants of classic SR repair, and outcomes from variants of classic DR repair (Table 2).

Significant (n = 16) and nonsignificant (n = 58) outcomes were analyzed across multiple parameters. The median FIs were 1 (IQR, 1-1) and 3 (IQR, 2-4) for significant and nonsignificant outcomes, respectively. The median FQs were 0.02 (IQR, 0.015-0.029) and 0.05 (IQR, 0.030-0.061), respectively. The mean FIs were 1.25 ± 0.58 and 2.96 ± 1.20, and the mean FQs were 0.023 ± 0.015 and 0.049 ± 0.027, respectively. A significant difference was found between these 2 groups for both FI (*P* < .001; 95% CI, 1.09-2.32) and FQ (*P* < .001; 95% CI, 0.019-0.033). A majority of the significant outcomes, 13/16 (81%), needed a single outcome reversal to become statistically nonsignificant.

When comparing the fragility index to the number of patients lost to follow up for a given outcome, the median FIs were 3 (IQR, 2-4) and 1.5 (IQR, 1-4) for FI < LTF and FI > LTF, respectively. The median FQs were 0.054 (IQR,

Empty Can Test	Positive	Negative	Lost to Follow Up
Single Row	9	15	13
Double Row	16	7	11
<i>P</i> value	0.041		

Empty Can Test	Positive	Negative	Lost to Follow Up
Single Row	10	14	13
Double Row	16	7	11
<i>P</i> value	0.08		

**Figure 2.** Hypothetical demonstration using an outcome with a fragility index of 1 of a single outcome event reversal resulting in altered statistical significance. The left panel shows the original data (in number of patients), and the right panel shows the modified data with a single outcome reversal.

TABLE 1  
Characteristics of the Included Studies<sup>a</sup>

Study	Year	Study Design	Journal	Mean FI	Mean FQ
Lapner et al <sup>28</sup>	2012	RCT	<i>JBJS</i>	2.0	0.024
Imam et al <sup>21</sup>	2020	RCT	<i>JSES</i>	4.0	0.050
Jeong et al <sup>22</sup>	2018	Comparative study	<i>JSES</i>	2.7	0.029
Ide et al <sup>20</sup>	2015	Comparative study	<i>JSES</i>	1.9	0.031
Plachel et al <sup>41</sup>	2020	Comparative study	<i>AJSM</i>	3.0	0.111
Yoon et al <sup>48</sup>	2019	Comparative study	<i>AJSM</i>	3.0	0.054
Hantes et al <sup>17</sup>	2018	Comparative study	<i>AJSM</i>	2.5	0.038
Franceschi et al <sup>12</sup>	2016	RCT	<i>AJSM</i>	3.1	0.063
Shin et al <sup>42</sup>	2015	Comparative study	<i>AJSM</i>	3.0	0.036
Kim et al <sup>25</sup>	2013	Comparative study	<i>AJSM</i>	1.0	0.027
Park et al <sup>39</sup>	2008	Comparative study	<i>AJSM</i>	2.0	0.026
Franceschi et al <sup>13</sup>	2007	RCT	<i>AJSM</i>	2.0	0.039
Charousset et al <sup>6</sup>	2007	Comparative study	<i>AJSM</i>	3.0	0.046
Yamakado <sup>47</sup>	2019	RCT	<i>Arthroscopy</i>	3.8	0.041
Barber <sup>3</sup>	2016	RCT	<i>Arthroscopy</i>	3.0	0.075
Denard et al <sup>8</sup>	2012	Comparative study	<i>Arthroscopy</i>	1.0	0.009
Ma et al <sup>29</sup>	2012	RCT	<i>Arthroscopy</i>	2.8	0.052
Koh et al <sup>26</sup>	2011	RCT	<i>Arthroscopy</i>	2.5	0.038
Pennington et al <sup>40</sup>	2010	Comparative study	<i>Arthroscopy</i>	2.5	0.047
Grasso et al <sup>14</sup>	2009	RCT	<i>Arthroscopy</i>	3.0	0.042
Sugaya et al <sup>43</sup>	2005	Comparative study	<i>Arthroscopy</i>	1.3	0.033
Carbonel et al <sup>5</sup>	2012	RCT	<i>Int Orthop</i>	2.0	0.013

<sup>a</sup>*AJSM*, American Journal of Sports Medicine; FI, fragility index; FQ, fragility quotient; *JBJS*, Journal of Bone and Joint Surgery—American; *JSES*, Journal of Shoulder and Elbow Surgery; *Int Orthop*, International Orthopaedics; RCT, randomized controlled trial.

TABLE 2  
Fragility of Analyzed Subgroups<sup>a</sup>

Characteristic	No. of Events	FI, mean ± SD	FQ, mean ± SD
All trials	74	2.55 ± 1.29	0.043 ± 0.027
Reported <i>P</i> value			
Significant (<.05)	16	1.25 ± 0.58	0.023 ± 0.015
Nonsignificant (>.05)	58	2.96 ± 1.20	0.049 ± 0.027
<i>P</i> value		<.001	<.001
Lost to follow-up			
FI < LTF	47	2.74 ± 1.15	0.049 ± 0.027
FI > LTF	27	2.22 ± 1.48	0.033 ± 0.025
<i>P</i> value		.047	.008
Double-row variants			
Classic double row	38	2.42 ± 1.18	0.040 ± 0.023
Variant (TOE)	36	2.69 ± 1.41	0.046 ± 0.031
<i>P</i> value		.184	.207
Single-row variants			
Classic single row	59	2.39 ± 1.31	0.037 ± 0.022
Variants (MMA, TLAR)	15	3.20 ± 1.01	0.062 ± 0.035
<i>P</i> value		.015	<.001
Outcome reported as retear			
Retear outcomes	21	1.95 ± 1.11	0.035 ± 0.024
Nonretear outcomes	53	2.79 ± 1.29	0.046 ± 0.028
<i>P</i> value		.005	.067

<sup>a</sup>FI, fragility index; FQ, fragility quotient; LTF, lost to follow-up; MMA, modified Mason-Allen; TLAR, triple-loaded anchor repair; TOE, transosseus equivalent.

0.027-0.061) and 0.024 (IQR, 0.009-0.050), respectively. The mean FIs were  $2.74 \pm 1.15$  and  $2.22 \pm 1.48$ , and the mean FQs were  $0.049 \pm 0.027$  and  $0.033 \pm 0.025$ ,

respectively. A significant difference was found between these 2 groups for both FI ( $P = .047$ ; 95% CI, 0.090-1.131) and FQ ( $P = .008$ ; 95% CI, 0.003-0.028).

For outcomes in studies examining classic DR repair ( $n = 38$ ) versus a variant version of DR repair (TOE) ( $n = 36$ ), the median FIs were 2 (IQR, 1-3) and 3 (IQR, 1-4), respectively. The median FQs were 0.036 (IQR, 0.020-0.058) and 0.036 (IQR, 0.021-0.054), respectively. The mean FIs were  $2.42 \pm 1.18$  and  $2.69 \pm 1.41$ , and the mean FQs were  $0.040 \pm 0.023$  and  $0.046 \pm 0.031$ , respectively. No significant difference was found between these 2 groups for both FI ( $P = .184$ ; 95% CI, -0.331 to 0.871) and FQ ( $P = .207$ , 95% CI, -0.007 to 0.017).

For outcomes in studies examining classic SR repair ( $n = 59$ ) versus a variant version of SR repair (MMA, TLAR) ( $n = 15$ ), the median FIs were 2 (IQR, 1-3) and 3 (IQR, 2-4), respectively. The median FQs were 0.033 (IQR, 0.016-0.056) and 0.052 (IQR, 0.034-0.075), respectively. The mean FIs were  $2.39 \pm 1.31$  and  $3.20 \pm 1.01$ , and the mean FQs were  $0.037 \pm 0.022$  and  $0.062 \pm 0.035$ , respectively. A significant difference was found between these 2 groups for both FI ( $P = .015$ ; 95% CI, 0.085-1.535) and FQ ( $P < .001$ ; 95% CI, 0.009-0.039).

Finally, for outcomes reported as retear ( $n = 21$ ) versus outcomes not reported as retear ( $n = 53$ ), the median FIs were 2 (IQR, 1-2.5) and 3 (IQR, 2-4), respectively. The median FQs were 0.026 (IQR, 0.018-0.054) and 0.044 (IQR, 0.027-0.061), respectively. The mean FIs were  $1.95 \pm 1.11$  and  $2.79 \pm 1.29$ , and the mean FQs were  $0.035 \pm 0.024$  and  $0.046 \pm 0.028$ , respectively. A significant difference was found between these 2 groups for FI ( $P = .005$ ; 95% CI, 0.200 to 1.480) but not for FQ ( $P = .067$ ; 95% CI, -0.004 to 0.024).

## DISCUSSION

For this review, the median FI was found to be 2, and the median FQ was 0.035, which is consistent with previous orthopaedic literature reporting an average median FI of 3.67<sup>†</sup> and FQ of 0.036.<sup>9,19,27,35-37,45</sup> This study's findings demonstrated that statistical significance can be changed with <3 outcome reversals, or reversal of 3.4% of outcome events. Furthermore, 13 of the 16 significant outcomes (81%) reported needed just a single outcome reversal to become statistically nonsignificant. Our hypothesis was confirmed in that more than half (63%) of the outcomes analyzed had an FI that was smaller than the number of patients lost to follow-up. This study adds to the growing body of evidence in support of the inclusion of FI and FQ into the RCTs and comparative studies that guide clinical decision making.

While  $P$  values are helpful in distilling complex statistical information down to a single value, there is growing concern in the medical community that these values may be oversimplifications and that continuing to use  $P$  values on their own may be detrimental to our understanding of the clinical significance of a study.<sup>7,46</sup>  $P$  values help to explain the probability of obtaining an outcome similar to or more extreme than the one observed if the null hypothesis is true,<sup>4</sup> which may be a good starting point for

assessing the strength of a given study. However, they do not account for factors such as size of the study and number of events observed, which are critical to the overall assessment of the strength of a study.<sup>46</sup> The results of a study may be considered significant but only need 1 or 2 outcome reversals to change their significance, which is frequently smaller than the number of patients lost to follow-up.<sup>37,46</sup> Despite these concerns, the reporting of  $P$  values continues to increase in the literature by approximately 8.2% annually, with 79% of clinical trials reporting  $P$  values in 2015.<sup>7</sup>

Inclusion of an FI can add clarity to these studies by ascribing a numerical value to the number of outcome reversals needed to alter the significance of a given study, although it is currently underused in the literature.<sup>9,36</sup> Recent studies have shown that research in orthopaedic sports medicine could benefit from the inclusion of the FI, with studies published in top orthopaedic journals needing an average of 5 outcome reversals, or 5.4% of the study population, to reverse significance.<sup>38</sup> The current study demonstrated similar fragility of orthopaedic literature, with 2.55 outcome reversals, or 3.4% of the study population, needed to reverse significance. With most power analyses accepting <20% loss to follow-up, this level of fragility highlights a risk of bias present within the current orthopaedic literature.<sup>46</sup>

While the FI is a useful tool in assessing the overall significance of a study, it fails to consider the size of the population being studied. Ahmed et al<sup>1</sup> saw this weakness in FI and sought to address it by dividing the FI of a study by the total number of patients included in the study, creating the FQ. Svantesson et al demonstrated the importance of the FQ by examining a population database study<sup>44</sup> that reported a median FI of 116, which is far above the average median FI in orthopaedic literature (3.67). However, this study had a very large patient population compared with the average orthopaedic study, with 5540 patients. By dividing the FI by the 5540 patients included in this population database study, Ehlers et al<sup>9</sup> found an FQ of 0.021, which is more consistent with the average FQ reported in orthopaedic literature (0.036) and in this study (0.043). This example highlights how FI can be misinterpreted when used on its own. The inclusion of FQ in concert with FI is warranted for a more holistic interpretation of the results of a study.

The focus of this study was on the clinical outcomes of SR versus DR repair for rotator cuff tears. A systematic review of meta analyses comparing SR repair with DR repair found that a majority of RCTs in the literature demonstrated very few differences in clinical outcomes between the 2 groups; however, DR repair had a lower rate of retear, particularly in patients with a tear >3 cm in size.<sup>32</sup> However, a more recent systematic review found that the labels "single row" and "double row" may be an oversimplification and that new, more biomechanically stable techniques which have yet to be adequately compared exist for both groups.<sup>30</sup> Despite the large number of studies that have been conducted comparing the 2, a consensus has yet to be reached, and further study is needed. Furthermore, no comparative studies to date have

<sup>†</sup>References 9, 10, 15, 19, 23, 24, 27, 31, 33, 34, 35, 37, 45.

included fragility in their statistical analysis. As demonstrated in this analysis, retear rates had a significantly smaller FI than all other dichotomous outcomes observed, needing a single outcome reversal to change the significance of this important outcome. Future studies could benefit from the inclusion of FI and FQ particularly, for primary outcomes such as retear rates, to demonstrate the stability of their findings relative to the current orthopaedic literature.

### Strengths and Limitations

There are several strengths to this study. The mean impact factor of the journals examined for this study was 4.67, which is higher than those of recent fragility analyses examining sports medicine (4.28),<sup>38</sup> spine (2.4),<sup>10</sup> and trauma (3.31).<sup>36</sup> The focus on the highest-impact orthopaedic journals limits risk of bias of the studies included in this analysis compared with studies published in journals with lower impact factors. Furthermore, this study included both primary and secondary outcome measures, making the application of FI and FQ in this study more generalizable. While an analysis of a commonly reported dichotomous primary measure, such as rate of retear, may prove valuable, the inclusion of secondary outcome measures allows this technique to be applied more broadly.

This study is not without limitations. Primarily, FI and FQ are only able to be calculated using dichotomous data and cannot be applied to continuous variables such as pain scores and outcome measurements, which make up a majority of the primary outcomes of the studies included in this analysis. Because only dichotomous data could be analyzed, 2 studies that otherwise met inclusion criteria for this review had to be excluded. Additionally, while we consider the inclusion of only the top 10 highest-impact journals in orthopaedic shoulder and elbow surgery to be a strength, there is indeed potential for other influential studies on SR versus DR rotator cuff repair to exist outside of these journals, and this limitation could be considered a relative weakness.

### CONCLUSION

The significance of studies examining SR versus DR repair for rotator cuff tears is fragile and subject to change from even a small number of outcome reversals. The average number of outcome reversals needed to reverse the significance was 2.55. The number of outcome reversals needed was smaller than the number of patients lost to follow-up for 63% of studies. Furthermore, 13 of the 16 outcomes reported as “significant” required just a single outcome reversal needed to alter their significance. Future studies may benefit from the inclusion of an FI and/or FQ in their statistical analyses.

### REFERENCES

1. Ahmed W, Fowler RA, McCredie VA. Does sample size matter when interpreting the fragility index? *Crit Care Med*. 2016;44(11):e1142-e1143.

2. Aydin N, Karaismailoglu B, Gurcan M, Ozsahin MK. Arthroscopic double-row rotator cuff repair: a comprehensive review of the literature. *SICOT J*. 2018;4:57.
3. Barber FA. Triple-loaded single-row versus suture-bridge double-row rotator cuff tendon repair with platelet-rich plasma fibrin membrane: a randomized controlled trial. *Arthroscopy*. 2016;32(5):753-761.
4. Biau DJ, Jolles BM, Porcher R. P value and the theory of hypothesis testing: an explanation for new researchers. *Clin Orthop Relat Res*. 2010;468(3):885-892.
5. Carbonel I, Martinez AA, Calvo A, Ripalda J, Herrera A. Single-row versus double-row arthroscopic repair in the treatment of rotator cuff tears: a prospective randomized clinical study. *Int Orthop*. 2012;36(9):1877-1883.
6. Charousset C, Grimberg J, Duranthon LD, Bellaiche L, Petrover D. Can a double-row anchorage technique improve tendon healing in arthroscopic rotator cuff repair? A prospective, nonrandomized, comparative study of double-row and single-row anchorage techniques with computed tomographic arthrography tendon healing assess. *Am J Sports Med*. 2007;35(8):1247-1253.
7. Chavalarias D, Wallach JD, Li AHT, Ioannidis JPA. Evolution of reporting P values in the biomedical literature, 1990-2015. *JAMA*. 2016;315(11):1141-1148.
8. Denard PJ, Jiwani AZ, Lädermann A, Burkhart SS. Long-term outcome of arthroscopic massive rotator cuff repair: the importance of double-row fixation. *Arthroscopy*. 2012;28(7):909-915.
9. Ehlers CB, Curley AJ, Fackler NP, Minhas A, Chang ES. The statistical fragility of hamstring versus patellar tendon autografts for anterior cruciate ligament reconstruction: a systematic review of comparative studies. *Am J Sports Med*. 2021;49(10):2827-2833.
10. Evaniew N, Files C, Smith C, et al. The fragility of statistically significant findings from randomized trials in spine surgery: a systematic survey. *Spine J*. 2015;15(10):2188-2197.
11. 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.
12. Franceschi F, Papalia R, Franceschetti E, et al. Double-row repair lowers the retear risk after accelerated rehabilitation. *Am J Sports Med*. 2016;44(4):948-956.
13. Franceschi F, Ruzzini L, Longo UG, et al. Equivalent clinical results of arthroscopic single-row and double-row suture anchor repair for rotator cuff tears: a randomized controlled trial. *Am J Sports Med*. 2007;35(8):1254-1260.
14. Grasso A, Milano G, Salvatore M, Falcone G, Deriu L, Fabbriciani C. Single-row versus double-row arthroscopic rotator cuff repair: a prospective randomized clinical study. *Arthroscopy*. 2009;25(1):4-12.
15. Grolleau F, Collins GS, Smarandache A, 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(3):456-462.
16. Hantes ME, Chalatsis GI, Mpakagiannis G. Single-row vs double-row repair in rotator cuff tears. In: Sozen S, ed. *Surgical Recovery*. Intech-Open; 2020.
17. Hantes ME, Ono Y, Raoulis VA, et al. Arthroscopic single-row versus double-row suture bridge technique for rotator cuff tears in patients younger than 55 years: a prospective comparative study. *Am J Sports Med*. 2018;46(1):116-121.
18. Hohmann E, König A, Kat CJ, Glatt V, Tetsworth K, Keough N. Single-versus double-row repair for full-thickness rotator cuff tears using suture anchors: a systematic review and meta-analysis of basic biomechanical studies. *Eur J Orthop Surg Traumatol*. 2018;28(5):859-868.
19. Huang X, Chen B, Thabane L, Adachi JD, Li G. Fragility of results from randomized controlled trials supporting the guidelines for the treatment of osteoporosis: a retrospective analysis. *Osteoporos Int*. 2021;32(9):1713-1723.
20. Ide J, Karasugi T, Okamoto N, Taniwaki T, Oka K, Mizuta H. Functional and structural comparisons of the arthroscopic knotless double-row suture bridge and single-row repair for anterosuperior rotator cuff tears. *J Shoulder Elbow Surg*. 2015;24(10):1544-1554.

21. Imam M, Sallam A, Ernstbrunner L, et al. Three-year functional outcome of transosseous-equivalent double-row vs. single-row repair of small and large rotator cuff tears: a double-blinded randomized controlled trial. *J Shoulder Elbow Surg.* 2020;29(10):2015-2026.
22. Jeong JY, Park KM, Sundar S, Yoo JC. Clinical and radiologic outcome of arthroscopic rotator cuff repair: single-row versus transosseous equivalent repair. *J Shoulder Elbow Surg.* 2018;27(6):1021-1029.
23. 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.
24. Khormaei S, Choe J, Ruzbarsky JJ, 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(8):418-423.
25. Kim YK, Moon SH, Cho SH. Treatment outcomes of single- versus double-row repair for larger than medium-sized rotator cuff tears: the effect of preoperative remnant tendon length. *Am J Sports Med.* 2013;41(10):2270-2277.
26. Koh KH, Kang KC, Lim TK, Shon MS, Yoo JC. Prospective randomized clinical trial of single- versus double-row suture anchor repair in 2- to 4-cm rotator cuff tears: clinical and magnetic resonance imaging results. *Arthroscopy.* 2011;27(4):453-462.
27. Kyriakides PW, Schultz BJ, Egol K, Leucht P. The fragility and reverse fragility indices of proximal humerus fracture randomized controlled trials: a systematic review. *Eur J Trauma Emerg Surg.* Published online May 31, 2021. doi:10.1007/s00068-021-01684-2
28. Lapner PLC, Sabri E, Rakhra K, et al. A multicenter randomized controlled trial comparing single-row with double-row fixation in arthroscopic rotator cuff repair. *J Bone Joint Surg Am.* 2012;94(14):1249-1257.
29. Ma HL, Chiang ER, Wu HTH, et al. Clinical outcome and imaging of arthroscopic single-row and double-row rotator cuff repair: a prospective randomized trial. *Arthroscopy.* 2012;28(1):16-24.
30. Maassen NH, Somerson JS. A majority of single versus double-row rotator cuff repair comparisons fail to consider modern single-row techniques: a systematic review. *JBJS Rev.* 2020;8(5):e203.
31. Maldonado DR, Go CC, Huang BH, Domb BG. The fragility index of hip arthroscopy randomized controlled trials: a systematic survey. *Arthroscopy.* 2021;37(6):1983-1989.
32. Mascarenhas R, Chalmers PN, Sayegh ET, et al. Is double-row rotator cuff repair clinically superior to single-row rotator cuff repair: a systematic review of overlapping meta-analyses. *Arthroscopy.* 2014;30(9):1156-1165.
33. Matics TJ, Khan N, Jani P, Kane JM. The fragility of statistically significant findings in pediatric critical care randomized controlled trials. *Pediatr Crit Care Med.* 2019;20(6):258-262.
34. Muthu S, Ramakrishnan E. Fragility analysis of statistically significant outcomes of randomized control trials in spine surgery: a systematic review. *Spine (Phila Pa 1976).* 2021;46(3):198-208.
35. Parisien RL, Constant M, Saltzman BM, et al. The fragility of statistical significance in cartilage restoration of the knee: a systematic review of randomized controlled trials. *Cartilage.* 2021;13(1)(suppl):147S-155S.
36. Parisien RL, Dashe J, Cronin PK, Bhandari M, Tornetta P. Statistical significance in trauma research: too unstable to trust? *J Orthop Trauma.* 2019;33(12):466-470.
37. 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.
38. 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):324-329.
39. Park JY, Lhee SH, Choi JH, Park HK, Yu JW, Seo JB. Comparison of the clinical outcomes of single- and double-row repairs in rotator cuff tears. *Am J Sports Med.* 2008;36(7):1310-1316.
40. Pennington WT, Gibbons DJ, Bartz BA, et al. Comparative analysis of single-row versus double-row repair of rotator cuff tears. *Arthroscopy.* 2010;26(11):1419-1426.
41. Plachel F, Siegert P, Rüttershoff K, et al. Long-term results of arthroscopic rotator cuff repair: a follow-up study comparing single-row versus double-row fixation techniques. *Am J Sports Med.* 2020;48(7):1568-1574.
42. Shin SJ, Kook SH, Rao N, Seo MJ. Clinical outcomes of modified Mason-Allen single-row repair for bursal-sided partial-thickness rotator cuff tears: comparison with the double-row suture-bridge technique. *Am J Sports Med.* 2015;43(8):1976-1982.
43. 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.
44. Svantesson E, Hamrin Senorski E, Danielsson A, et al. Strength in numbers? The fragility index of studies from the Scandinavian knee ligament registries. *Knee Surg Sport Traumatol Arthrosc.* 2020;28(2):339-352.
45. Tignanelli CJ, Napolitano LM. The fragility index in randomized clinical trials as a means of optimizing patient care. *JAMA Surg.* 2019;154(1):74-79.
46. Walsh M, Srinathan SK, McAuley DF, et al. The statistical significance of randomized controlled trial results is frequently fragile: a case for a fragility index. *J Clin Epidemiol.* 2014;67(6):622-628.
47. Yamakado K. A prospective randomized trial comparing suture bridge and medially based single-row rotator cuff repair in medium-sized supraspinatus tears. *Arthroscopy.* 2019;35(10):2803-2813.
48. Yoon JS, Kim SJ, Choi YR, Kim SH, Chun YM. Arthroscopic repair of the isolated subscapularis full-thickness tear: single- versus double-row suture-bridge technique. *Am J Sports Med.* 2019;47(6):1427-1433.