Article

AMERICAN ORTHOPAEDIC FOOT & ANKLE SOCIETY

Foot & Ankle International®
2024, Vol. 45(4) 298–308
© The Author(s) 2024
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/10711007231224797
journals.sagepub.com/home/fai

Statistical Fragility Analysis of Open Reduction Internal Fixation vs Primary Arthrodesis to Treat Lisfranc Injuries: A Systematic Review

Benjamin J. Ahn, BS[1], Matthew Quinn, MD[2], Leon Zhao, BS[2], Elaine W. He, BA[1], Myles Dworkin, MD[2], Om Naphade[1], Rory A. Byrne, BA[2], Janine Molino, PhD[1], and Brad Blankenhorn, MD[2]

Abstract

Background: There is a lack of consensus in the use of open reduction internal fixation (ORIF) vs primary arthrodesis (PA) in the management of Lisfranc injuries. Statistical fragility represents the number of events needed to flip statistical significance and provides context to interpret *P* values of outcomes from conflicting studies. The current study evaluates the statistical fragility of existing research with an outcome-specific approach to provide statistical clarity to the ORIF vs PA discussion. We hypothesized that statistical fragility analysis would offer clinically relevant insight when interpreting conflicting outcomes regarding ORIF vs PA management of Lisfranc injuries.
Methods: All comparative studies, RCTs, and case-series investigating ORIF vs PA management of Lisfranc injuries published through October 5, 2023, were identified. Descriptive characteristics, dichotomous outcomes, and continuous outcomes were extracted. Fragility index and continuous fragility index were calculated by the number of event reversals needed to alter significance. Outcomes were categorized by clinical relevance, and median FI and CFI were reported.
Results: A total of 244 studies were screened. Ten studies and 67 outcomes (44 dichotomous, 23 continuous) were included in the fragility analysis. Of the 10 studies, 4 studies claimed PA to correlate with superior outcomes compared to ORIF with regard to functional scores and return to function outcomes. Of these 4 studies, 3 were statistically robust. Six studies claimed PA and ORIF to have no differences in outcomes, in which only 2 studies were statistically robust.
Conclusion: The overall research regarding ORIF vs PA is relatively robust compared with other orthopaedic areas of controversy. Although the full statistical context of each article must be considered, studies supporting PA superiority with regard to functional scores and return to function metrics were found to be statistically robust. Outcome-specific analysis revealed moderate fragility in several clinically relevant outcomes such as functional score, return to function, and wound complications.

Keywords: statistical fragility, fragility index, Lisfranc, open reduction internal fixation, primary arthrodesis

Introduction

Lisfranc injuries describe a spectrum of ligamentous or bony injuries involving the intertarsal or tarsometatarsal joints, ranging from sprain to fracture-dislocation of the midfoot.[8] These injuries can result from either low-energy or high-energy trauma and are often the result of axial load on a plantarflexed foot with rotational stress.[18] Estimated incidence has ranged from 1/60 000 up to 14/100 000

[1]The Warren Alpert Medical School of Brown University, Providence, RI, USA
[2]Department of Orthopaedic Surgery, The Warren Alpert Medical School of Brown University, Providence, RI, USA

Corresponding Author:
Benjamin J. Ahn, BS, The Warren Alpert Medical School of Brown University, Box #G-9006, 222 Richmond Street, Providence, RI 02903, USA.
Email: benjamin_ahn@brown.edu

person-years, with the higher estimates coming from more recent publications as increased awareness and imaging advancements have improved detection.[19,36,41]

For Lisfranc injuries requiring operative management, there is currently no definitive consensus on open reduction and internal fixation (ORIF) vs primary arthrodesis (PA). Traditionally, the gold standard was ORIF with either transarticular screw fixation or dorsal bridge plating.[2,34] Over the last decade, PA has become increasingly popular particularly for ligamentous injuries because of their poor healing potential as well as comminuted fractures with disruption of the articular surface.[18,45] Proponents of PA also argue that there is limited inherent motion through the medial and middle column joints so any losses in range of motion or function will be minimal. The long-term effects of these losses, however, are still unknown.[9] However, ORIF has been associated with more surgeries predominantly for hardware removal or secondary fusion.[3,20,21,23]

Given the current lack of consensus, there is a need to evaluate the relative merits of previous studies and determine how these results should impact clinical decision making. Although *P* values are commonly used to demonstrate statistical significance, these values can be misleading because of the impact of study design and power. Statistical fragility analyses offer a way to quantify the robustness of statistically significant results. The statistical fragility of dichotomous outcomes can be measured using the fragility index (FI) and fragility quotient (FQ). The FI is the minimum number of patients for whom reversal of their outcome would flip significance.[44] The FQ is the FI number divided by total sample size, which is a relative measure that accounts for trial size and patients lost to follow-up (LTF).[1] The continuous fragility index (CFI) and continuous fragility quotient (CFQ)[44] were developed to measure the statistical fragility of continuous outcomes, expanding the usefulness of fragility analysis. In general, higher fragility indices indicate greater confidence in a study's results. Meanwhile, lower fragility indices or those lower than the LTF weaken the confidence in a study's results as outcomes may have changed with better participant retention.

The aim of this study was to evaluate the statistical fragility of outcomes found in existing comparative literature on Lisfranc ORIF vs PA to aid in the interpretation of existing literature to facilitate clinical decision making. Previous statistical fragility studies within orthopaedic surgery have found low median FIs ranging from 2 to 6, including Achilles tendon management (median FI = 2.5), spine (median FI = 2), adult reconstruction (median FI = 6), hip arthroscopy (median FI = 4), hand surgery (median FI = 3), sports (median FI = 2), trauma (median FI = 3), and pediatrics (median FI = 3).[14-17,22,25,26,31,39] We hypothesized that the available literature for management of Lisfranc injuries would be consistent with other orthopaedic specialties with low FI and CFIs.

Methods

Search Strategy

We systematically collected all published studies that evaluated the outcomes of Lisfranc procedures through October 5, 2023 (Figure 1). For the initial search, we used the PubMed database with the following terms: ((lisfranc) OR (tarsometatarsal fracture) OR (tarsometatarsal dislocation) OR (tarsometatarsal fusion) OR (midfoot fusion)) AND ((open reduction internal fixation) OR (ORIF) OR (fixation)) AND ((primary arthrodesis) OR (PA) OR (fusion)). We did not rely on PubMed filters to exclusively search for comparative studies and RCTs, and instead, manually screened studies and excluded them based on their design when applicable.

We conducted title and abstract screening to identify relevant studies for inclusion in the analysis. Studies were included if they compared ORIF to PA. The full text of each article was examined, and studies were excluded using the following criteria: (1) unrelated to Lisfranc, ORIF, or PA; (2) systematic reviews, meta-analyses, or review articles; (3) cadaveric, nonhuman, in vitro, laboratory, or surgical technique studies; (4) case reports or series without a comparison group; (5) commentary, editorial, letter to the editor, conference reports, protocols, or future study design; (6) studies for which only an abstract was available; (7) studies not published in English; or (8) studies lacking the necessary statistics for fragility analysis. Studies were independently reviewed by 5 authors, and discrepancies were resolved through paired discussions.

Data Extraction and Fragility Analysis

The data collected for each article included the published journal, publication year, level of evidence, length of follow-up, and trial type. All dichotomous and continuous outcomes relevant to the postoperative results of ORIF and PA were extracted. For each dichotomous outcome, the specific outcome assessed, sample size, number of participants lost to follow-up, reported *P* value, and the number of events were recorded. For each continuous outcome, the sample size, number lost to follow-up, reported *P* value, SE and/or SD, and sample means were recorded.

To evaluate the fragility of dichotomous outcomes, we calculated a Fragility Index (FI). This was accomplished using a 2-by-2 contingency table and Fisher exact test.[44] In an iterative process, we moved 1 patient from a negative to a positive event until the statistical significance was reversed. The FI represents the number of patients needed to be moved to reverse significance (Figure 2). This analysis was performed for both significant and nonsignificant outcomes. The Fragility Quotient (FQ) was calculated by dividing the FI by the sample size.[43]

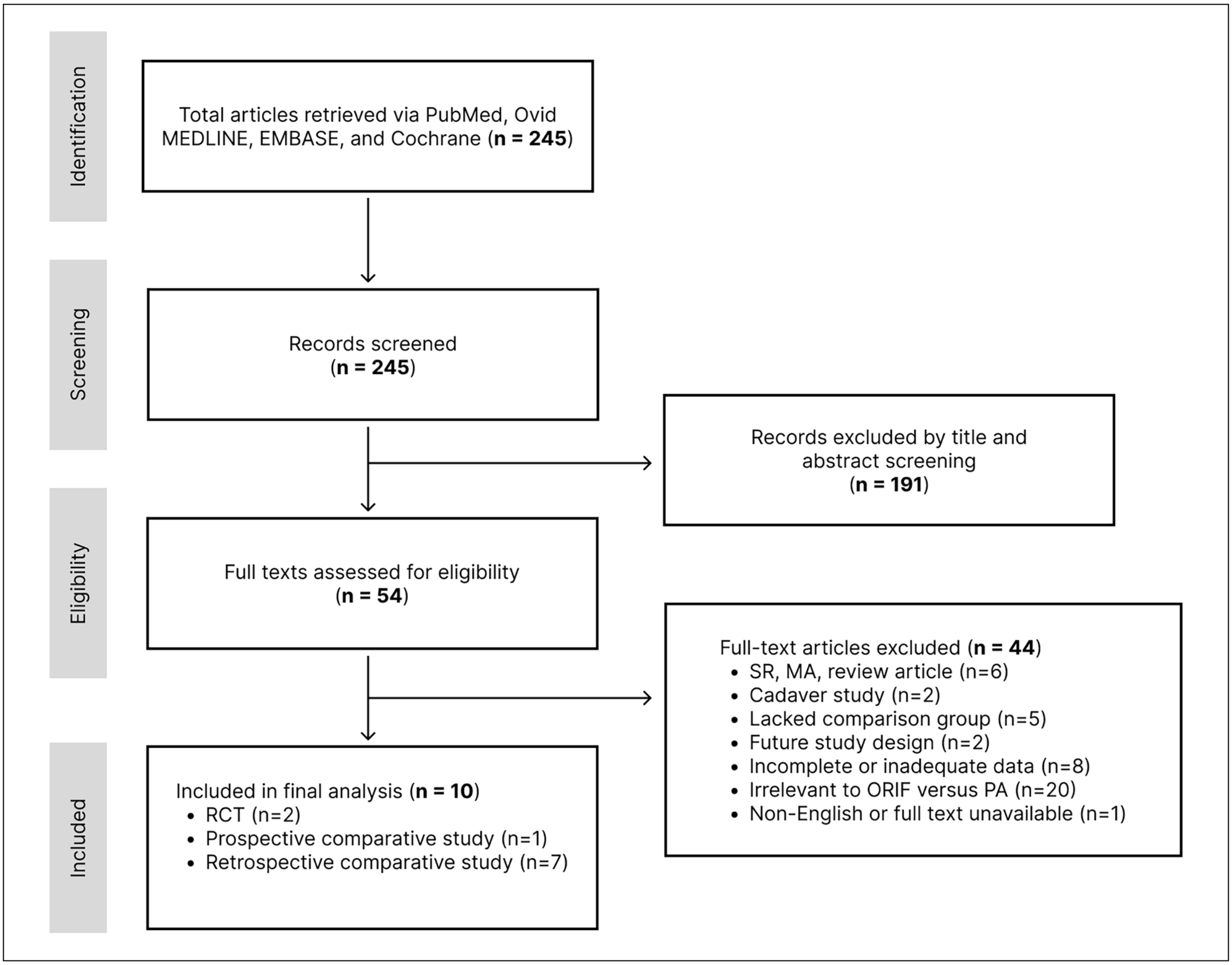


Figure 1. Study inclusion flowchart.
Abbreviations: MA, meta-analysis; RCT, randomized controlled trial; SR, systematic review.

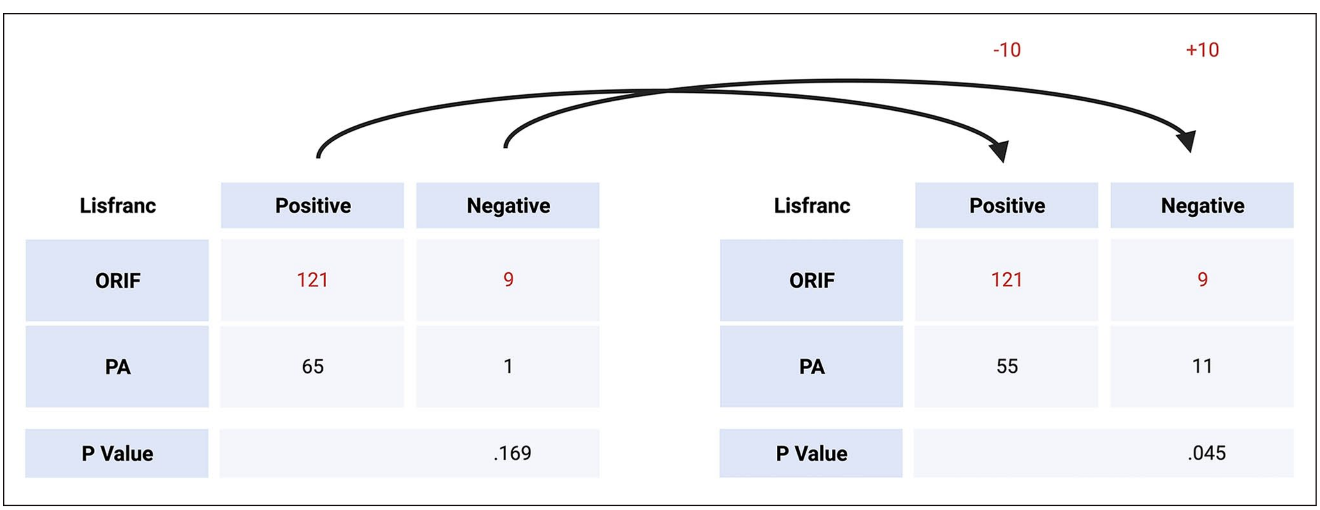
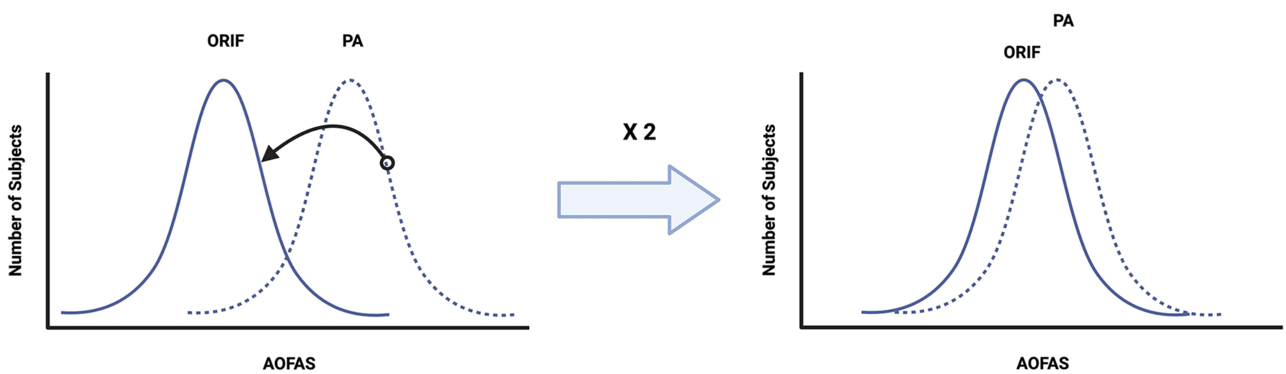


Figure 2. A demonstration of how the fragility index (FI) is calculated for dichotomous variables. In this example, a 10-subject event reversal (FI = 10) resulted in altered statistical significance.



	μ	σ	t-test P Value
ORIF	88.58	6.36	0.012
PA	94.00	3.16	

	μ	σ	t-test P Value
ORIF	89.02	6.30	0.063
PA	93.62	3.63	

Figure 3. A demonstration of how the continuous fragility index (CFI) is calculated for continuous variables. In this example, a 2-subject event reversal (CFI = 2) resulted in altered statistical significance.

Continuous outcomes must be analyzed differently from dichotomous outcomes, as they are not represented by discrete data points, such as number of patients, but instead a range of values, such as a score. To evaluate the fragility of continuous outcomes, we calculated a Continuous Fragility Index (CFI) using Welch *t* test.[7] This statistical method has been refined by other studies, enabling its application to outcomes that do not report raw data, increasing its utility.[29,46] For each outcome, synthetic representative data were generated based on the reported sample mean, SD, and sample size for the experimental and control groups. Through an iterative process, a subject was transferred from one data set to another until the statistical significance was reversed (Figure 3). The CFI represents the number of patients that need to be moved in this process. This process was simulated 5 times and the CFI was averaged. The Continuous Fragility Quotient (CFQ) was calculated by dividing the CFI by the sample size.

We further analyzed the statistical fragility of outcomes categorized by outcome type. Dichotomous and continuous outcomes were divided into categories decided by BA and MQ, and the median FI/CFI and interquartile range were reported.

Lastly, we used both the FI and CFI to analyze the statistical fragility of individual studies. For each study, the authors' main conclusions were extracted, and relevant outcomes and their associated FI or CFI were organized. Studies with an FI or CFI less than the LTF were categorized as fragile, and studies with an FI or CFI greater than the LTF were categorized as robust. Additionally, studies with an absolute FI or CFI below 2, the smallest FI reported in other orthopaedic fragility studies,[14] were classified as weak. Data analysis and plotting was performed using Python 3.7.

Results

Out of the initially identified 244 studies through a primary search, a total of 53 studies underwent full-text screening. After inclusion and exclusion criteria, 10 studies were deemed eligible and included in the final analysis. Sixty-seven total outcomes were identified for analysis, including 44 dichotomous outcomes and 23 continuous outcomes.

Table 1 presents a summary of the characteristics of the included studies and the measured outcomes. The majority of included studies were retrospective (70%). Figure 4 reports the distribution of FI values for all dichotomous and continuous outcomes.

Table 2 describes the fragility analysis separated by dichotomous outcome categories. Out of 44 total dichotomous outcomes, the median FI was 7.0 (25-75th percentile, 5.0-8.0), and the median FQ was 0.092 (25-75th percentile, 0.047-0.209). Dichotomous outcome categories commonly reported across different studies included return to function metrics, reoperation/revision, and wound complications/infections. Table 3 describes the fragility analysis separated by continuous outcome categories. There were 23 continuous outcomes in total, and the median CFI was 8.8 (25-75th percentile, 3.5-12.5), with a median CFQ of 0.100 (25-75th percentile, 0.072-0.142). Continuous outcome categories commonly reported across different studies included return to function, AOFAS, and SF-36 scores.

Table 4 compares each of the study's original claims with major findings from the fragility analysis as well as

Table 1. General Characteristics of Included Outcomes From 10 Studies.

Characteristics of Included Studies	Number of Studies (n = 10)	Number of Outcomes (n = 67)
Journal, n (%)		
Foot & Ankle International	3 (30)	18 (26.7)
Other	7 (70)	49 (73.1)
Outcome type		
Dichotomous	NA	44 (65.7)
Continuous	NA	23 (34.3)
Study type, n (%)		
RCT	2 (20)	14 (20.9)
Prospective case series	1 (10)	10 (14.9)
Retrospective comparative/cohort	7 (70)	43 (64.2)
Reported *P* value, n (%)		
≥.05	NA	47 (70.1)
<.05	NA	20 (29.9)
Publication year, n (%)		
2021-2023	4 (40)	37 (55.2)
2016-2020	6 (60)	21 (31.3)
<2016	1 (10)	9 (13.4)
Level of evidence, n (%)		
I	3 (30)	24 (35.8)
II	0 (0)	0 (0)
III	6 (60)	39 (58.2)
IV	1 (10)	4 (6.0)

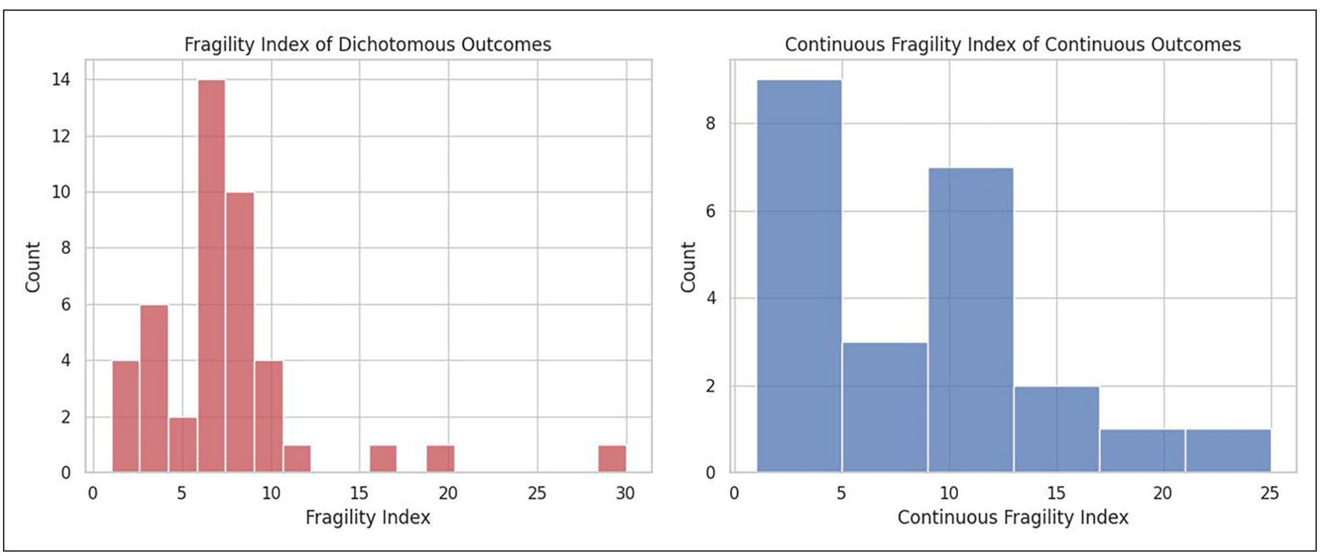


Figure 4. Fragility indices and continuous fragility indices of dichotomous and continuous outcomes across all studies.

study size and LTF. Of the 10 studies included in the analysis, 5 were determined to be statistically fragile and 5 were found to be statistically robust. Of the 10 studies, 4 studies claimed PA to correlate with superior outcomes compared to ORIF, in which 3 of these studies were statistically robust and 1 was statistically fragile. Six studies claimed PA and ORIF to have no differences in outcomes, in which 2 studies were statistically robust and 4 studies were statistically fragile.

Discussion

Historically, the literature comparing the PA and ORIF for Lisfranc injuries has been diverse with varying degrees of

Table 2. Fragility Analysis by Outcome, Dichotomous.[a]

Outcome	No. of Outcomes	Fragility Index, Median (IQR)	Fragility Quotient, Median (IQR)	No. (%) of Outcomes Where Loss to Follow-up Was Greater Than FI
All studies	44	7.0 (5.0, 8.0)	0.092 (0.047, 0.209)	7 (15.9%)
Arthritis	4	6.5 (2.75, 12.25)	0.059 (0.051, 0.074)	1 (25%)
CRPS	2	5.0 (4.0, 6.0)	0.124 (0.070, 0.179)	0 (0%)
Hardware complications	5	7.0 (7.0, 8.0)	0.156 (0.078, 0.219)	2 (40%)
Quality of anatomic reduction	5	6.0 (4.0, 8.0)	0.188 (0.033, 0.205)	1 (20%)
Reoperation/revision	9	8 (5.0, 9.0)	0.072 (0.037, 0.111)	0 (0%)
Return to function	10	7.0 (6.0, 7.0)	0.106 (0.078, 0.174)	3 (30%)
Wound complications/infections	6	7.0 (4.5, 8.0)	0.077 (0.023, 0.218)	0 (0%)
Miscellaneous	3	8 (7.5, 9.0)	0.25 (0.215, 0.281)	0 (0%)

Abbreviations: CRPS, Complex Regional Pain Syndrome; FI, Fragility Index.
[a]Outcomes from the included studies were clustered into broader categories for further analysis.

Table 3. Fragility Analysis by Outcome, Continuous.[a]

Outcome	No. Outcomes	Continuous fragility index, median (IQR)	Continuous fragility quotient, median (IQR)	No. (%) outcomes where lost to follow up was greater than CFI
All studies	23	8.8 (3.5, 12.5)	0.100 (0.072, 0.142)	7 (30.43)
AOFAS	5	5.8 (3.4, 8.2)	0.08 (0.076, 0.136)	2 (40)
FAAM	1	8.8 (NA, NA)	0.100 (NA, NA)	1 (100)
FFI	2	11.40 (9.5, 13.30)	0.186 (0.152, 0.220)	1 (50)
MOXFQ	1	2.4 (NA, NA)	0.053 (NA, NA)	1 (100)
Operation logistics	2	20.5 (18.15, 22.85)	0.123 (0.109, 0.138)	0 (0)
Return to function	6	9.2 (4.95, 11.20)	0.092 (0.071, 0.115)	0 (0)
SF-36	3	4.4 (3.5, 8.7)	0.104 (0.077, 0.126)	1 (33.33)
VAS	3	12.4 (6.7, 15.8)	0.097 (0.068, 0.158)	1 (33.33)

Abbreviations: AOFAS, American Orthopaedic Foot & Ankle Society Ankle-Hindfoot Score; CFI, Continuous Fragility Index; FAAM, Foot and Ankle Ability Measure; FFI, Foot Function Index; MOXFQ, Manchester–Oxford Foot Questionnaire; SF-36, Short Form-36; VAS, Visual Analog Scale.
[a]Outcomes from included studies were clustered into broader categories for further analysis.

consensus. More recently, however, authors have proposed that PA may offer superior outcomes compared with ORIF.[21,23,27,42] This manuscript uses fragility analysis as a novel framework to appraise the robustness of these comparative outcomes and help clinicians make more informed decisions based on the resilience of available evidence. For all included studies, we found a median FI of 7.0 (IQR: 5.0-8.0), FQ of 0.09 (0.047-0.209), CFI of 8.8 (3.5-12.5), and CFQ of 0.100 (0.072-0.142). These values are higher than previously reported scores in orthopaedic literature,[16-23] suggesting more statistically robust outcomes. Interestingly, studies that claimed to support no difference between ORIF and PA were generally statistically fragile whereas those that supported superior outcomes with PA tended to be statistically robust. Taken together, the results of the current study suggest that the available literature for management of Lisfranc injuries is more robust than what has been previously published in other subspecialties, and that the strongest evidence available supports superior outcomes with PA over ORIF.

The statistical fragility analysis provides an intuitive way to interpret the significance of an outcome, affecting one's confidence in a study's findings and improving clinical translation of results. *P* values are frequently misinterpreted, and statistical significance may be conflated with result importance by medical providers.[4,24,33] *P* values are also influenced by sample size, effect size, and variance, and these factors change when study results are applied in clinical practice. For example, if an underpowered study produces a nominally significant finding, the reported effect size will be an overestimation compared with the population value.[10] Fragility analysis provides a means for clinicians to better contextualize the results from a study, especially given the high fragility found in many orthopaedic studies.[14-17,22,25,26,31,39]

Fragility Analysis of "Functional Score" Outcomes

Our review of the literature found that a number of variables were used to evaluate functional outcomes in ORIF vs PA managed injuries with no clear consensus based on *P* value alone. Studies reporting superior outcomes with PA

Table 4. Summary Table Comparing the Claims of Papers With the Fragility Analysis and Interpretation.[a]

Study	Author	Year	Level of Evidence	Study's Main Conclusion	Fragility Statistics	Main Interpretation
Open Reduction Internal Fixation Versus Primary Arthrodesis for Lisfranc Injuries: A Prospective Randomized Study	Henning et al[21]	2009	I	PA and ORIF had no differences in functional scores or satisfaction	Quality of anatomic reduction: FI = 6 Return to function: FI = 10, 8, 7, 6 n = 40 LTF = 8	Fragile: severely underpowered study with LTF greater than many FIs
Clinical Outcomes and Development of Symptomatic Osteoarthritis 2 to 24 Years After Surgical Treatment of Tarsometatarsal Joint Complex Injuries	Dubois-Ferriere et al[11]	2016	IV	PA and ORIF had no differences in functional outcomes	Functional scores: VAS FI = 14.2 FFI FI = 12.8 AOFAS FI = 4 n = 128 LTF = 67	Fragile: large LTF greater than FI
Outcomes of Lisfranc Injuries in an Active Duty Military Population	Hawkinson et al[20]	2017	III	PA and ORIF had no differences in return to function	Return to function: FI = 8, 8, 4, 6 Reoperation rates: CFI = 4 n = 111 LTF = 0	**Robust:** mostly large FIs greater than LTF
Comparison of Arthrodesis and Non-fusion to Treat Lisfranc Injuries	Qiao et al[37]	2017	III	PA has improved functional scores over ORIF	Functional scores: AOFAS CFI = 2 VAS CFI = 1 SF-36 CFI = 1.6 n = 25 LTF = 0	Fragile: small CFIs
Reoperation Rate Differences Between Open Reduction Internal Fixation and Primary Arthrodesis of Lisfranc Injuries	Buda et al[6]	2018	III	PA and ORIF had no differences in reoperation rates	Reoperation rates: FI = 8 n = 217 LTF = 0	**Robust:** large FIs
Primary Arthrodesis Versus Open Reduction Internal Fixation for Complete Lisfranc Fracture Dislocations: A Retrospective Study Comparing Functional and Radiological Outcomes	Kirzner et al[27]	2020	III	PA results in a better functional outcome via improved anatomical reduction than ORIF	Quality of anatomic reduction: FI = 11, 8 n = 45 LTF = 6	**Robust**: large FIs greater than LTF
A Comparison of Complications and Reoperations Between Open Reduction and Internal Fixation Versus Primary Arthrodesis Following Lisfranc Injury	So et al[40]	2021	III	No difference between ORIF and PA in reoperation and wound complication rates	Wound complications: FI = 4, 6 Reoperation: CFI = 5 n = 196 LTF = 0	Fragile: small FIs especially considering large study size
Primary Fusion Versus Open Reduction Internal Fixation for Purely Ligamentous Lisfranc Injuries: A Prospective Comparative Study and Analysis of Factors Affecting the Outcomes	Kandil et al[23]	2022	I	PA lead to better functional outcomes than ORIF	Functional scores: AOFAS CFI = 7.8 FFI CFI = 7.4 n = 30 LTF = 0	**Robust:** large CFIs especially compared to smaller study size
Lisfranc Injuries With Dislocation the First Tarsometatarsal Joint: Primary Arthrodesis or Internal Fixation (a Randomized Controlled Trial)	Sun et al[42]	2022	I	PA has improved functional scores over ORIF	Functional scores: AOFAS CFI = 12.4 FAAM CFI = 7.4 VAS CFI = 19.0 SF-36 physical function CFI = 3.6 SF-36 bodily pain CFI = 12.8 n = 88 LTF = 10	**Robust:** large CFIs greater than LTF
Functional Outcomes of Primary Arthrodesis (PA) Versus Open Reduction and Internal Fixation (ORIF) in the Treatment of Lisfranc Injuries	Aneja et al[3]	2023	III	No difference in return to function	Return to function: FI = 10, 7, 7, 6 n = 272 LTF = 191	Fragile: large LTF compared to FIs

Abbreviations: AOFAS, American Orthopaedic Foot & Ankle Society Ankle-Hindfoot Score; CFI, continuous fragility index; FAAM, Foot and Ankle Ability Measure; FFI, Foot Function Index; FI, fragility index; LTF, loss to follow-up; n, study size; PA, Primary Arthrodesis; ORIF, Open Reduction Internal Fixation; SF-36, Short Form-36; VAS, Visual Analogue Scale.

[a]Studies with LTF greater than FI or CFI highlight statistical fragility. Main interpretation was determined by evaluating the FI or CFI vs the LTF. Studies with a small FI or CFI compared with LTF were classified as fragile.

were generally found to be statistically robust whereas those reporting no difference between PA and ORIF were generally found to be fragile. The most common individual functional outcome score was the AOFAS, a nonvalidated score based on patient- and physician-reported metrics such as pain, function, and alignment for foot and ankle conditions[28] reported by 5 studies. The studies supporting PA superiority over ORIF were found to be statistically robust. Sun et al and Kandil et al both conducted RCTs with 88 and 30 patients, respectively, and reported strong CFIs (Sun CFI = 12.4; Kandil CFI = 7.8) with little LTF (Sun LTF = 10; Kandil LTF = 0). These CFIs are greater than their study's LTF, and come from level I evidence studies as well, supporting their claims of PA leading to improved functional scores over ORIF. Kandil et al does acknowledge a small study size and lack of a power analysis, potentially limiting the overall statistical strength of this study despite the robust CFI. Ly and Coetzee,[30] a highly cited prospective randomized trial, also found that PA led to superior AOFAS scores over ORIF ($AOFAS_{PA} = 88$, $AOFAS_{ORIF} = 68.6$). This study could not be included in our fragility analysis because of unreported SDs, and they reported a wide range for $AOFAS_{ORIF}$ scores (16-100), potentially limiting the statistical rigor. However, this study conducted its own power analysis in which the authors determined the study to be sufficiently powered despite a total study size of 41 patients. Nonetheless, the robust outcomes from Sun et al and Kandil et al in addition to the Ly and Coetzee study provide a body of evidence supporting superior AOFAS scores with PA over ORIF. In contrast, the studies claiming no difference between PA and ORIF were found to be statistically weak. For example, Dubois-Ferriere et al found no significant difference in AOFAS between ORIF- and PA-managed injuries in a retrospective study of 127 patients. Fragility analysis demonstrated their results to be statically weak. Although the number of patients lost to follow-up was 67 (52%), the CFI was only 5.8.

Another commonly used functional outcome measure was SF-36, a patient-reported survey that assesses health-related quality of life.[5] Similar to results found with AOFAS scores, studies reporting PA superiority over ORIF were found to be statistically robust while studies reporting no difference between PA and ORIF were statistically fragile. Sun et al[42] found significant improvement with PA over ORIF with a difference in SF-36 greater than 5 points. In this study, the SF-36 physical function score had a CFI of 13, suggesting greater confidence in their results. Although no specific MCID has been established for Lisfranc injuries, previously orthopaedic publications have found that a change of 5 points represents a clinically important change, further supporting the findings of Sun et al.[42] In contrast, Qiao et al[37] found no significant difference in their retrospective cohort study. The CFI for Qiao et al[37] was only 2.6, indicating that if only 3 participants switched groups the results would have been significantly different. Similarly, Henning et al[21] found no significant difference between PA and primary ORIF. Although the study failed to report sufficient statistics for their SF-36 outcomes to conduct a fragility analysis, it suffers significantly because of loss to follow-up and underpowered study size, which was acknowledged in their original manuscript. Despite the underlying statistical issues, the Henning et al study remains one of the most highly cited papers on this topic highlighting the potential for misinterpretation of *P* value–only analysis and the utility of fragility analysis.

Fragility Analysis of "Return to Function" Outcomes

The literature consists of conflicting results regarding return to function outcomes in ORIF vs PA when evaluating evidence solely by *P* value. Our fragility analysis found that studies claiming PA leads to worse outcomes than ORIF were statistically fragile. Of the studies supporting no difference between PA and ORIF, some were statistically robust. Sixteen outcomes across 6 studies assessed various return to function outcomes. Many outcomes indicated no significant difference between ORIF- and PA-managed injuries.[3,20,21,23] For example, Hawkinson et al[20] found no significant difference in "return to duty" with a robust FI of 8 and no loss to follow up in a retrospective study of 171 active duty military participants. Henning et al[21] similarly found no statistical difference in requirement of assistive device for ambulation at final follow-up between ORIF and PA with a CFI of 10 and LTF of 8. One must use caution in interpreting these results as the study was severely underpowered with retention of only 32 of 60 patients needed to meet their predetermined benchmark for statistical power. Finally, So et al[40] reported that PA had a significantly longer time to weightbearing than ORIF. However, further analysis revealed that the difference was 0.7 days, which brings into question the clinical relevance of this finding. Taken together, fragility analysis from these studies suggests equivalence between PA and ORIF when assessing return to function.

Fragility Analysis of "Wound Complications" and "Reoperation Rate" Outcomes

Lastly, we investigated various outcomes regarding wound complications and reoperation rates. This category includes 6 outcomes across 5 studies.[3,6,21,23,40] The majority of studies were statistically fragile, and only 1 study supporting no difference between ORIF and PA was found to be statistically robust. In Buda et al,[6] reoperation rates were not statistically significant between ORIF and PA with a CFI of 8 and LTF of 0. However, this robust finding may be driven by the fact that the authors excluded removal of hardware surgeries from this analysis. Studies often reported 2 reoperation rates: one that included and one that excluded planned hardware removal. Focusing on reoperation rates excluding planned removal could more accurately reflect reoperation rates due to true complications, serving as a more clinically relevant measure. Other studies investigating wound healing outcomes demonstrated statistical weakness regardless of their stance of PA vs ORIF (rate of infection: So et al,[40] FI = 4, LTF = 0, study size = 196; Rate of salvage procedure: Henning et al,[21] FI = 4, LTF = 8, study size = 40). Although fragility analysis cannot support either PA or ORIF with regard to wound complications and reoperation rates, it does reveal the absence of robust literature, further demonstrating its usefulness directing future research efforts.

Strengths and Limitations

Our study is one of the first, to our knowledge, that uses fragility analysis to conduct a study-individualized and outcome-specific analysis to support clinical decision making. The majority of statistical fragility papers lump all studies and outcomes into one analysis, or at the most, categorize their analysis by trial type.[12–17,22,25,26,31,32,38,39,46] Our analysis enables readers to directly compare the claims of studies and focus on clinically relevant outcome categories. Furthermore, this is one of the few studies that conducts a comprehensive fragility analysis, investigating both dichotomous and continuous outcomes as well as RCT, retrospective, and case series studies.

Our study does acknowledge multiple weaknesses. Many outcomes for foot and ankle studies are subjective patient-reported outcomes, the most common being the AOFAS. The AOFAS has not been validated, and its use is recommended against by the American Orthopaedic Foot & Ankle Society,[35] presenting a limitation in interpreting the literature. Additionally, given the paucity of high-quality, randomized trials investigating this topic, there are inherent limitations in our ability to account for potential confounding variables from the results included in our fragility analysis. Unfortunately this limitation is compounded by the fact that studies such as the RCT by Ly and Coetzee were also excluded from our analysis because of insufficient distribution statistics reported in original manuscripts. Another limitation from the study is that all variations of ORIF are grouped together and therefore does not account for differences in fragility between techniques such as dorsal plating vs independent screw fixation. Lastly, fragility analysis cannot be the deciding factor when interpreting clinical research and thus limits the applicability of our findings. Additional information such as the minimal clinically important difference, effect size, and study size all provide useful context when interpreting outcomes and must always be considered.

Conclusion

We conducted an outcome-specific and study-specific analysis, investigating the management of Lisfranc injuries with PA vs ORIF. Although the full statistical context of each study must be considered, studies supporting PA superiority with regard to functional scores and return to function metrics were found to have robust indices, while studies supporting no difference between PA and ORIF were found to have fragile indices. The majority of studies investigating wound complications and reoperation rates were statistically fragile, highlighting the need for additional research focused on these outcomes.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article. Disclosure forms for all authors are available online.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Benjamin J. Ahn, BS, https://orcid.org/0009-0005-5185-1096

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. doi:10.1097/CCM.0000000000001976
2. Alberta FG, Aronow MS, Barrero M, Diaz-Doran V, Sullivan RJ, Adams DJ. Ligamentous Lisfranc joint injuries: a biomechanical comparison of dorsal plate and transarticular screw fixation. *Foot Ankle Int*. 2005;26(6):462-473. doi:10.1177/107110070502600607
3. Aneja A, Gal ZT, Dawson AN, et al. Functional outcomes of primary arthrodesis (PA) versus open reduction and internal fixation (ORIF) in the treatment of Lisfranc injuries. *J Orthop Trauma*. 2023;37(8):412-416. doi:10.1097/BOT.0000000000002606
4. Bhandari M, Montori VM, Schemitsch EH. The undue influence of significant p-values on the perceived importance of study results. *Acta Orthop*. 2005;76(3):291-295.
5. Brazier JE, Harper R, Jones NM, et al. Validating the SF-36 Health Survey questionnaire: new outcome measure for primary care. *BMJ*. 1992;305(6846):160-164. doi:10.1136/bmj.305.6846.160
6. Buda M, Kink S, Stavenuiter R, et al. Reoperation rate differences between open reduction internal fixation and primary arthrodesis of Lisfranc injuries. *Foot Ankle Int*. 2018;39(9):1089-1096. doi:10.1177/1071100718774005
7. Caldwell JME, Youssefzadeh K, Limpisvasti O. A method for calculating the fragility index of continuous outcomes. *J Clin Epidemiol*. 2021;136:20-25. doi:10.1016/j.jclinepi.2021.02.023

8. Cassebaum WH. Lisfranc fracture-dislocations. *Clin Orthop Relat Res*. 1963;30:116-129.
9. Clare MP. Lisfranc injuries. *Curr Rev Musculoskelet Med*. 2017;10(1):81-85. doi:10.1007/s12178-017-9387-6
10. Crutzen R, Peters GJY. Targeting next generations to change the common practice of underpowered research. *Front Psychol*. 2017;8:1184. doi:10.3389/fpsyg.2017.01184
11. Dubois-Ferrière V, Lübbeke A, Chowdhary A, Stern R, Dominguez D, Assal M. Clinical Outcomes and Development of Symptomatic Osteoarthritis 2 to 24 Years After Surgical Treatment of Tarsometatarsal Joint Complex Injuries. *J Bone Joint Surg Am*. 2016;98(9):713-20. doi: 10.2106/JBJS.15.00623
12. Ehlers CB, Curley AJ, Fackler NP, et al. The statistical fragility of single-bundle vs double-bundle autografts for ACL reconstruction: a systematic review of comparative studies. *Orthop J Sports Med*. 2021;9(12):23259671211064626. doi:10.1177/23259671211064626
13. 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. doi:10.1177/0363546520969973
14. 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. doi:10.1016/j.spinee.2015.06.004
15. Fackler NP, Karasavvidis T, Ehlers CB, et al. The statistical fragility of operative vs nonoperative management for Achilles tendon rupture: a systematic review of comparative studies. *Foot Ankle Int*. 2022;43(10):1331-1339. doi:10.1177/10711007221108078
16. Forrester LA, McCormick KL, Bonsignore-Opp L, et al. Statistical fragility of surgical clinical trials in orthopaedic trauma. *J Am Acad Orthop Surg Glob Res Rev*. 2021; 5(11):e20.00197. doi:10.5435/JAAOSGlobal-D-20-00197
17. Go CC, Maldonado DR, Go BC, et al. The fragility index of total hip arthroplasty randomized control trials: a systematic review. *J Am Acad Orthop Surg*. 2022;30(9):e741-e750. doi:10.5435/JAAOS-D-21-00489
18. Grewal US, Onubogu K, Southgate C, Dhinsa BS. Lisfranc injury: a review and simplified treatment algorithm. *Foot (Edinb)*. 2020;45:101719. doi:10.1016/j.foot.2020.101719
19. Hardcastle PH, Reschauer R, Kutscha-Lissberg E, Schoffmann W. Injuries to the tarsometatarsal joint. Incidence, classification and treatment. *J Bone Joint Surg Br*. 1982;64(3):349-356. doi:10.1302/0301-620X.64B3.7096403
20. Hawkinson MP, Tennent DJ, Belisle J, Osborn P. Outcomes of Lisfranc injuries in an active duty military population. *Foot Ankle Int*. 2017;38(10):1115-1119. doi:10.1177/1071100717719532
21. Henning JA, Jones CB, Sietsema DL, Bohay DR, Anderson JG. Open reduction internal fixation versus primary arthrodesis for Lisfranc injuries: a prospective randomized study. *Foot Ankle Int*. 2009;30(10):913-922. doi:10.3113/FAI.2009.0913
22. Herndon CL, McCormick KL, Gazgalis A, Bixby EC, Levitsky MM, Neuwirth AL. Fragility index as a measure of randomized clinical trial quality in adult reconstruction: a systematic review. *Arthroplasty Today*. 2021;11:239-251. doi:10.1016/j.artd.2021.08.018
23. Kandil MI, Abouzeid M, Eltaher SM, Eltregy S. Primary fusion versus open reduction internal fixation for purely ligamentous Lisfranc injuries: a prospective comparative study and analysis of factors affecting the outcomes. *Foot Ankle Surg*. 2022;28(7):898-905. doi:10.1016/j.fas.2021.12.006
24. Kelley K, Preacher KJ. On effect size. *Psychol Methods*. 2012;17(2):137-152. doi:10.1037/a0028086
25. 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. doi:10.1177/0363546516674469
26. Khormaee 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):e418-e423. doi:10.1097/BPO.0000000000001201
27. Kirzner N, Teoh W, Toemoe S, et al. Primary arthrodesis versus open reduction internal fixation for complete Lisfranc fracture dislocations: a retrospective study comparing functional and radiological outcomes. *ANZ J Surg*. 2020;90(4):585-590. doi:10.1111/ans.15627
28. Kitaoka HB, Alexander IJ, Adelaar RS, Nunley JA, Myerson MS, Sanders M. Clinical rating systems for the ankle-hindfoot, midfoot, hallux, and lesser toes. *Foot Ankle Int*. 1994;15(7):349-353. doi:10.1177/10711007940 1500701
29. Lin L, Chu H. Assessing and visualizing fragility of clinical results with binary outcomes in R using the fragility package. *PLoS One*. 2022;17(6):e0268754. doi:10.1371/journal.pone.0268754
30. Ly TV, Coetzee JC. Treatment of primarily ligamentous Lisfranc joint injuries: primary arthrodesis compared with open reduction and internal fixation. A prospective, randomized study. *J Bone Joint Surg Am*. 2006;88(3):514-520. doi:10.2106/JBJS.E.00228
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. doi:10.1016/j.arthro.2021.01.049
32. Milto AJ, Negri CE, Baker J, Thuppal S. The statistical fragility of foot and ankle surgery randomized controlled trials. *J Foot Ankle Surg*. 2023;62(1):191-196. doi:10.1053/j.jfas.2022.08.014
33. 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. doi:10.5435/JAAOS-D-17-00636
34. Philpott A, Epstein DJ, Lau SC, Mnatzaganian G, Pang J. Lisfranc fixation techniques and postoperative functional outcomes: a systematic review. *J Foot Ankle Surg*. 2021;60(1):102-108. doi:10.1053/j.jfas.2020.04.005
35. Pinsker E, Daniels TR. AOFAS position statement regarding the future of the AOFAS clinical rating systems. *Foot Ankle Int*. 2011;32(9):841-842. doi:10.3113/FAI.2011.0841
36. Ponkilainen VT, Laine HJ, Mäenpää HM, Mattila VM, Haapasalo HH. Incidence and characteristics of midfoot injuries. *Foot Ankle Int*. 2019;40(1):105-112. doi:10.1177/1071100718799741

37. Qiao YS, Li JK, Shen H, et al. Comparison of arthrodesis and non-fusion to treat Lisfranc injuries: comparative study of Lisfranc injuries. *Orthop Surg*. 2017;9(1):62-68. doi:10.1111/os.12316
38. Reito A, Raittio L, Helminen O. Fragility index, power, strength and robustness of findings in sports medicine and arthroscopic surgery: a secondary analysis of data from a study on use of the Fragility Index in sports surgery. *PeerJ*. 2019;7:e6813. doi:10.7717/peerj.6813
39. Ruzbarsky JJ, Khormaee S, Daluiski A. The fragility index in hand surgery randomized controlled trials. *J Hand Surg*. 2019;44(8):698.e1-698.e7. doi:10.1016/j.jhsa.2018.10.005
40. So E, Lee J, Pershing ML, et al. A comparison of complications and reoperations between open reduction and internal fixation versus primary arthrodesis following Lisfranc injury. *Foot Ankle Spec*. Published online November 28, 2021. doi:10.1177/19386400211058264
41. Stødle AH, Hvaal KH, Enger M, Brøgger H, Madsen JE, Husebye EE. Lisfranc injuries: incidence, mechanisms of injury and predictors of instability. *Foot Ankle Surg*. 2020;26(5):535-540. doi:10.1016/j.fas.2019.06.002
42. Sun C, Miao X, Zhang M, et al. Lisfranc injuries with dislocation the first tarsometatarsal joint: primary arthrodesis or internal fixation (a randomized controlled trial). *Int Orthop*. 2022;46(11):2529-2537. doi:10.1007/s00264-022-05478-y
43. 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. doi:10.1001/jamasurg.2018.4318
44. 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. doi:10.1016/j.jclinepi.2013.10.019
45. Welck MJ, Zinchenko R, Rudge B. Lisfranc injuries. *Injury*. 2015;46(4):536-541. doi:10.1016/j.injury.2014.11.026
46. Xu AL, Ortiz-Babilonia C, Gupta A, Rogers D, Aiyer AA, Vulcano E. The statistical fragility of platelet-rich plasma as treatment for chronic noninsertional Achilles tendinopathy: a systematic review and meta-analysis. *Foot Ankle Orthop*. 2022;7(3):24730114221119758. doi:10.1177/24730114221119758