The Statistical Significance of Orthopaedic Randomized Controlled Trial Results May Be Frequently Reversed by Patients Lost to Follow-up: A Case for a Lost to Follow-up Index

Jacob Francis Oeding, Michael Robert Mazzucco, Robert Nicholas Uzzo, Steven G Persaud, John Eric Lama, Kyle Kunze, Kristian Samuelsson[1]
[1]Sahlgrenska University Hospital

INTRODUCTION:

In a parallel, two-arm, randomized clinical trial, patients are assigned to one of two groups, and their outcomes are observed over an extended period of time. During this time, some patients may be lost to follow-up (LTF), making their outcomes unobservable and potentially biasing trial conclusions. Recently, the orthopaedic literature has seen a rise in the number of studies applying the fragility index (FI) to study the robustness of conclusions made from randomized controlled trials (RCTs). In short, the FI measures how many patient outcomes would need to change to alter the significance of the results (flip the p-value from significant to non-significant). However, without official thresholds for what constitutes a “fragile” vs. “robust” RCT, many authors resort to comparing the FI to the number of patients LTF to highlight concerns about the impact of missing data. Unfortunately, such comparisons may be highly inaccurate and misleading, as the FI changes outcomes without altering patient counts, unlike adding patients LTF back to the trial and considering their effect on the significance of the trial. Thus, there is a need for a statistical measure that can precisely and accurately determine the extent to which an RCT’s conclusions are fragile to its patients LTF. In this study, we propose an LTF index (LTFI) and apply it to study the extent to which RCTs published in high-impact orthopaedic journals are susceptible to having the significance of their conclusions reversed by patients LTF.

METHODS:

A comprehensive search of the literature was performed to identify all RCTs published in the top orthopaedic journals by impact factor between 2018 and 2023. Studies were identified that involved two parallel treatment arms, allocated patients to experimental and control groups in a 1:1 ratio, and had a primary outcome that was significant and based on a dichotomous variable. The LTFI was defined as the number of patients LTF that would be required to flip the significance of each trial from significant to non-significant if added back to the trial in the distribution of maximal significance reversal. The distribution of maximal significance was defined as the distribution of patients LTF that maximally increases the trial’s p-value. Because this distribution is dependent on the distribution of patients included in the original trial, this distribution could be one of three options: 1) all patients LTF are from the control group, 2) all patients LTF are from the experimental group, or 3) patients LTF are approximately equally distributed between control and experimental groups. In the third case, patients are added back to the trial by alternating those added to the experimental and control groups. Next, the number of patients assumed to be LTF was incrementally increased one by one, and these patients were sequentially added back to the trial in the distribution of maximal significance reversal until a reversal of significance was realized. These values were compared to the number of actual patients LTF in each trial as well as each trial’s FI. **Figure 1** (FI) and **Figure 2** (LTFI) illustrate the difference in effect on significance between the FI and LTFI in an example trial.

RESULTS:

Eleven studies met the inclusion criteria and were ultimately included in this review (**Figure 3)**. **Table 1** provides a summary of included studies, of which eight were determined to be at low risk of bias and three were determined to have some concerns for risk of bias based on the Cochrane Risk of Bias version 2 tool (**Figure 4**). The mean (IQR) number of patients LTF was 11 (8). **Figure 5** provides a summary of included studies comparing each study’s number of patients LTF to both the FI and LTFI for that study. Importantly, it was found that up to 45% of studies could have their significance reversed simply by maintaining patient follow-up (**Figure 6**).

DISCUSSION AND CONCLUSION:

The primary findings of the present study are 1) a LTFI may be used to quantify the extent to which RCTs are fragile to patients LTF, which importantly, is different from the extent to which RCTs are fragile to changes in outcomes of existing patients (the FI), and 2) under the three extreme distributions of patients LTF, 36% to 45% of RCTs in this study could have the significance of their results reversed by the patients LTF. This study has significant implications for clinical trial investigators, journal editors, systematic reviewers, and readers of orthopaedic literature. To ensure that results of RCTs are interpreted in their appropriate context, investigators should minimize LTF and provide detailed reporting on its extent, timing, reasons, and the characteristics of those lost, broken down by study arm. In addition, readers must recognize that positive results may be fragile to patients LTF, particularly with small effect sizes and high LTF rates. The LTFI is a tool that may be used by future investigators to study the extent to which RCTs may be fragile to patients LTF.

Table 1

Summary Table of Included Studies

	N = 11[1]
Journal	
Am J Sports Med	3 (27%)
Arthroscopy	2 (18%)
J Bone Joint Surg Am	6 (55%)
Sample Size (#)	235 (150)
Mean age (yrs.)	27 (22)
Females (%)	35 (31)
Number lost to follow-up (LTF)	11 (8)
Total Sample Size Control (#)	117 (78)
Total Sample Size Experimental (#)	118 (78)

[1] n (%); Mean (IQR)

Figure 1

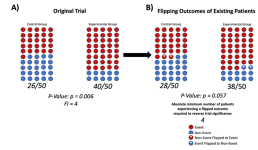

Figure 2

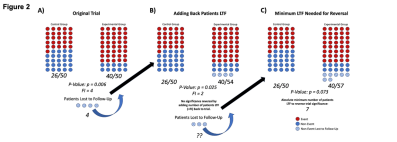

Figure 3

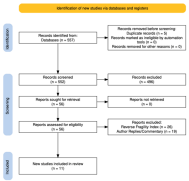

Figure 4

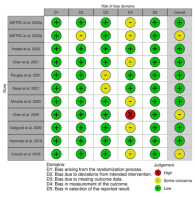

Figure 5

Figure 6