The Fragility of Statistical Findings in the Femoral Head Avascular Necrosis Literature: A Systematic Review of Randomized Controlled Trials

Ara Alexanian1; Christopher Gonzalez2; Avanish Yendluri2; Michael N. Megafu3; Emmanuel C. Megafu4; Katrina Nietsch5; Thomas A. Einhorn5; Robert L. Parisien6

1Harbor-UCLA Medical Center, Torrance, CA, 2Icahn School of Medicine at Mount Sinai, New York, NY, 3A.T. Still University, Kirksville, MO, 4Geisinger Commonwealth School of Medicine, Scranton, PA, 5New York University Langone Medical Center, New York, NY
AAlexanian@dhs.lacounty.gov


INTRODUCTION: Avascular necrosis of the femoral head (AVNFH) is a debilitating condition that may progress to femoral head collapse, requiring hip arthroplasty. Comparative analyses with the reporting of P-values are often used in the orthopaedic trauma literature and inform management of AVNFH. This study utilized the fragility index (FI), reverse fragility index (rFI), and fragility quotient (FQ) metrics to determine the statistical stability of outcomes in randomized controlled trials (RCTs) assessing AVNFH.

METHODS: Pubmed, Embase, and MEDLINE were queried for RCTs published from January 2010 - March 2023 that reported dichotomous, categorical outcomes related to AVNFH. The FI and reverse FI were defined as the number of outcome event reversals required to reverse the significance of the study. The FQ was determined by dividing the FI or reverse FI by the respective sample size.

RESULTS SECTION: Of the 294 studies evaluated, 22 RCTs were included, comprising a total of 82 outcome events. Across the 82 total outcomes the median FI was 4 (IQR 2-5) and the median FQ was 0.058 (IQR 0.034-0.097). There were 15 statistically significant outcomes with a median FI of 3 (IQR 1-11) and associated FQ of 0.031 (IQR 0.021-0.057). For the 67 nonsignificant outcomes the median FI was 4 (IQR 2.5-5) with an associated FQ of 0.063 (0.041-0.099). Clinical improvement and failure were the most fragile outcome categories with median FI's of 1.5 (IQR 1-4) and 1.5 (IQR 1.25-1.75), respectively. Across 20 radiographic outcomes, the median FI was 4.5 (IQR 4-6). For 16 femoral head collapse outcomes, the median FI was 3.5 (IQR 2.75-5). The 13 conversion to THA outcomes had a median FI of 5 (IQR 2-5). Lastly, for progression of AVN, complication/adverse event, and survival/mortality, the median FI's were 4 (IQR 3-5), 5 (IQR 2.5-5.5), and 3 (IQR 2-4), respectively. In 22% of outcomes, the number of patients lost to follow-up was greater than the FI.

DISCUSSION: The outcomes reported in RCTs evaluating treatment of AVNFH are fragile and should be interpreted with caution. Alarmingly, significant outcomes were particularly fragile and the number of patients lost to follow-up was greater than the FI in 22% of outcomes, indicating that only a few outcome event reversals or even maintaining postoperative follow-up may alter outcome significance. We therefore recommend triple reporting of the P-value with FI and FQ metrics to allow clinicians to ensure the implementation of evidence-based interventions.

SIGNIFICANCE/CLINICAL RELEVANCE: This study addresses the reliability of significant reported outcomes in the AVNFH literature. It will help guide clinical management of these patients while recommending future studies implement FI and FQ in their analyses. This technique can also be expanded to numerous other topics in orthopaedics.