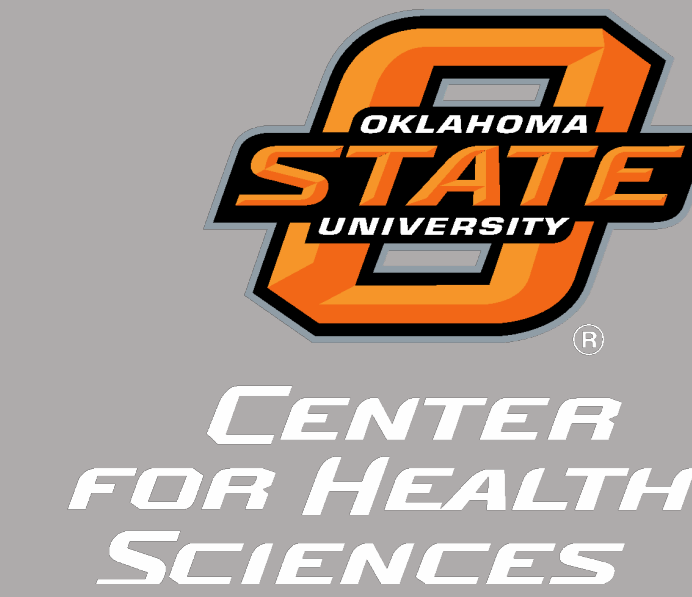




Statistical Significance and Orthopaedic Traumatology



1. Oklahoma State University Center for Health Sciences
2. Orthopaedic & Trauma Services of Oklahoma, Tulsa, Oklahoma.
3. Oklahoma State University Medical Center - Department of Orthopaedics

Austin L. Johnson, OMS I¹, Sheridan Evans, OMS I¹, Jake X. Checketts, OMS III¹, Jared T. Scott, OMS IV¹, Cole Wayant OMS II¹, Mark Johnson, DO³, Brent Norris, MD^{2,3}, Matt Vassar, PhD^{1,3}.

Introduction

A recent proposal suggests changing the threshold for statistical significance from a P value of .05 to .005 to minimize bias and increase reproducibility of future studies^{1,2}. P values less than .05 but greater than .005 would be reclassified as “suggestive”, whereas P values less than .005 would be considered significant. The present study explores how lowering the P value threshold would affect the interpretation of previously published trauma orthopaedic randomized clinical trials (RCTs) and whether outcomes from these trials would maintain statistical significance under the proposed P value threshold.

Methods

All RCTs published between January 01, 2016 and January 31, 2018 in the *Journal of Orthopaedic Trauma*, *Injury*, and *Archives of Orthopaedic and Trauma Surgery*. All RCTs were screened by at least 2 authors. We extracted P value data for primary endpoints, since RCTs are most often powered for these endpoints. If a study had multiple primary endpoints, or evaluated the primary endpoint from multiple domains, all P values for these endpoints were include. Data were extracted blinded and in duplicate fashion. Discrepancies were resolved by consensus. We used Google Forms for data extraction and STATA 13.1 for the data analysis.



Results

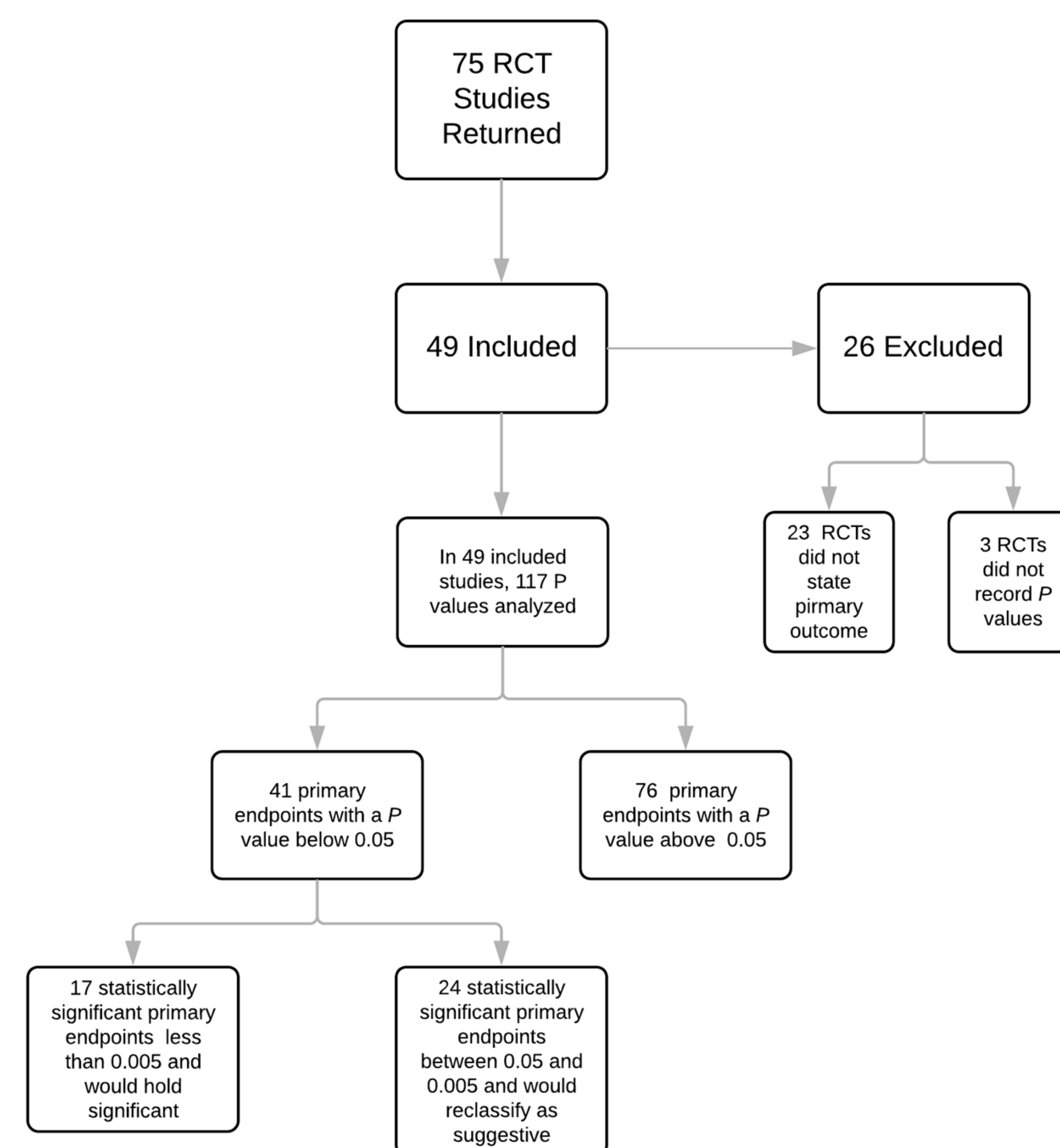
We identified 117 primary endpoints from 49 trials: 41 endpoints (35.0%) had a P value less than .05 and 76 (65.0%) had a P value greater than .05. Overall, 41.5% (17/41) of statistically significant primary endpoints were less than .005, while 58.5% (24/41) would be reclassified as suggestive. Of the 117 primary endpoints, only 17 (14.5%) of the endpoints were less than .005, and would hold significance with the proposed threshold. Only 6.12% (3/49) of the included studies had all primary endpoints that met the new threshold of .005.

Of these 117 total P values, 18.2% (6/33) in *Archives of Orthopaedic and Trauma Surgery*, 26.0% (6/23) in *Injury*, and 8.2% (5/61) in *Journal of Orthopaedic Trauma* were less than .005.

Surgery was the primary intervention type for 59.2% (29/49) of the included trials. The majority of trials did not mention any funding source (25/49, 51.0%). The majority of included trials, 89.8% (44/49), had randomized groups. All included studies were single nation, however only 71.4% were single center (35/49). 65.3% (32/49) of studies included a power analysis. Of the 3 studies where all primary endpoints met the new threshold, none mentioned a power analysis.

Characteristic	No. (%)
Journal (n = 49)	
Archives of Orthopaedic and Trauma Surgery	30.6%
Journal of Orthopaedic Trauma	44.9%
Injury	24.5%
Intervention (n = 49)	
Drug	10.2%
Procedure	10.2%
Anesthesia/Analgesia (Nerve Blocks/Pain Management)	10.2%
Surgery	59.2%
Education	2.0%
Other	8.2%
Funding Source (n = 49)	
Industry	4.1%
Public	4.1%
Private	16.3%
Hospital	2.0%
Mixed (no Industry)	6.1%
Mixed (with Industry)	2.0%
Other	2.0%
Not Mentioned	51.0%
None	12.2%
Number of trial centers (n = 49)	
Multicenter	28.6%
Single center	71.4%
Location (n = 49)	
Multinational	0%
Single country	100%
Type of endpoint (n = 117)	
Subjective	14.3%
Objective	71.4%
Mixed	14.3%
Sample size (median, [IQR])	
	76

P Value CONSORT Flowchart



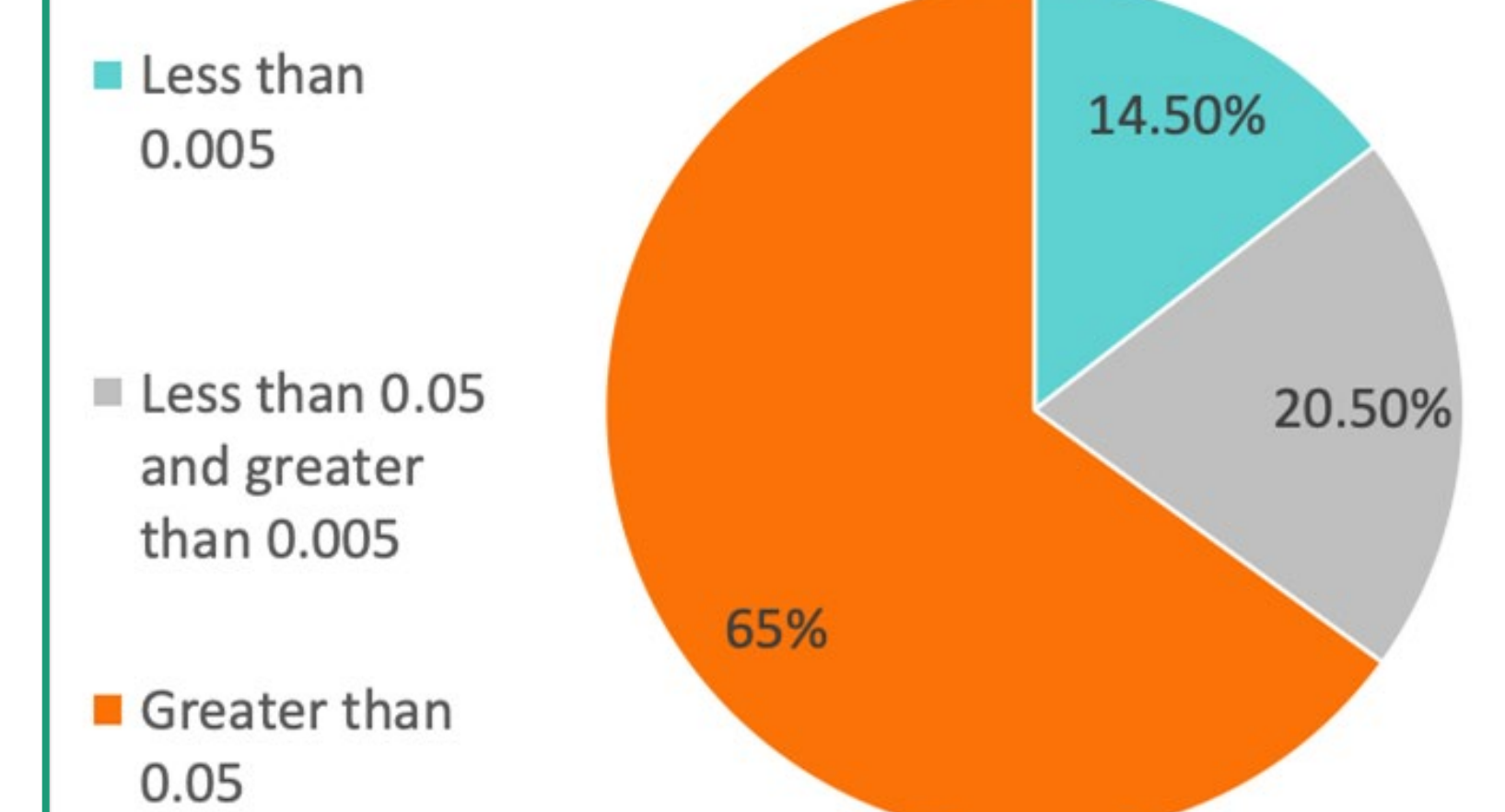
Conclusion

Based on our results, adopting a lower threshold of significance would heavily alter the interpretation of orthopaedic trauma RCTs and should be further evaluated and cautiously considered in regards to the impact it may have on orthopaedic traumatology practice.

Limitations

One limitation of this study is we included only 3 high impact factor Orthopaedic Trauma medical journals over a 3 year period; thus, the results may not be generalizable to RCTs in other orthopaedic traumatology journals. Furthermore, we only included 2 years of studies, and thus the results of our study may not be generalizable to past or future years.

P Values



References

1. Benjamin DJ, Berger JO, Johannesson M, et al. Redefine statistical significance. *Nature Human Behaviour*. 2017;2(1):6-10.
2. Ioannidis JPA. The Proposal to Lower P Value Thresholds to .005. *JAMA*. 2018;319(14):1429-1430.