The Fallacy of Patient-Reported Outcome Measures: Why Patient-Reported Outcome Measures are Bad for Clinical Trials

Nathan O'Hara, PhD, MHA*; Dane Brodke, MD; Ashley E. Levack, MD; Gerard P. Slobogean, MD, MPH*

Purpose: Patient-reported outcome measures are increasingly common in orthopaedic trauma research, yet they routinely fail to detect clinically meaningful differences in clinical trials. We assessed whether orthopaedic trauma studies are more likely to detect clinically important differences with a binary outcome, such as nonunion, or a continuous patient-reported outcome sensitive to important clinical differences.

Methods: We constructed a hypothetical clinical trial comparing two treatments of tibial shaft fractures. We assumed the true nonunion rate was 10% in the treatment group and 20% in the control group. Based on recent publications, we assumed patients with nonunion had a mean PROMIS physical function score of 40 (SD: 10) and patients with union had a mean score of 50 (SD: 10), a difference that is double the minimum clinically important difference. Based on these expected distributions, we simulated 10,000 trials with a sample size of 400 patients to assess the probability of detecting a significant difference between treatment groups at $p<0.05$. The trial results were analyzed using the binary outcome of nonunion surgery and PROMIS physical function.

Results: Of the 10,000 simulated trials with a 400-patient sample, the observed difference in nonunion surgery between the treatment groups was statistically significant 80.9% of the time. In contrast, the 400-patient sample detected a statistically significant difference between groups in PROMIS physical function only 4.2% of the time.

Conclusion: Even if patient-reported outcome measures are sensitive to important clinical events, most studies will fail to detect statistically significant differences between the treatment groups when the important clinical outcome occurs in a minority of the overall sample. Orthopaedic traumatologists should be aware of this limitation when designing and appraising studies that compare treatments with patient-reported outcomes.