January 3, 2018

Are the Conclusions Supported by the Results and Study Design?

Aleksi Reito, MD, PhD
Resident in Orthopaedic Surgery
Central Finland Hospital

I read the study by Helmerhorst et al. with great interest. There is no doubt that the opioid crisis has hit catastrophic heights. However, I have some concerns related to this study, because research addressing this problem of postoperative analgesia must be robust.

Relatively few patients were enrolled into the study, considering the heterogeneity of the baseline injuries and the long recruitment period. I am concerned that the small sample may not warrant the conclusion of the study, which states that “acetaminophen should be considered the mainstay for pain relief in patients recovering from extremity fracture surgical procedures.”

The sample size number was based on a priori power analysis. The authors helpfully disclose alfa and beta levels and the noninferiority (NI) margin that they used. The authors used RStudio software, but which function and package in RStudio did they use for this analysis? Most importantly, the uncertainty associated with the primary outcome, ie. standard deviation, is crucial and should be always reported for the purposes of a priori power analysis.

The primary outcome used in the power analysis, self-reported satisfaction with pain relief, was measured using an 11-point ordinal scale. The authors state that a 2.0 point NI margin was used based on a previous study, but in their previous study, the authors used a 5-point Likert scale. How was this 2.0 point NI margin derived, and is it meaningful in the 11-point scale? Has the 11-point scale been externally validated, and has the minimal clinically important difference been assessed prior to this study?

In the previous study, the authors assumed an effect size (Cohen´s d) of 1.0 since both the mean difference and deviation was 1. An effect size of 1 is extremely large. In their previous study, the authors also state that “we felt that only very large differences between the cohorts were clinically relevant.” Was this same rationale used in this study? Should not Level I studies rely on outcome measures that have been

externally validated and for which clearly defined minimal clinically important differences have been derived?

To conclude, it seems that this study may have been underpowered to reliably rule out a meaningful difference between the analgesic regimens. Based on the methods, only a very large difference in the effects of these regimens could be detected. Therefore I feel that the authors' conclusions are not supported by the results.

Conflict of Interest: