Raters associated with erratic ratings differ in their scoring technique

Alan Kott, MUDr; Xingmei Wang, MS; David Daniel, MD

Signant Health

THE METHODOLOGICAL ISSUE BEING ADDRESSED

Do raters who are associated with erratic ratings differ in their scoring technique of the PANSS scale?

INTRODUCTION

- Erratic ratings represent large visit to visit score changes in opposite direction across multiple consecutive visits.

- In prior sensitivity analyses, the removal of erratic data in negative symptom schizophrenia trials made the 10mg of bitopertin significantly outperform placebo (Umbricht et al, 2020), and removing erratic data in a successful phase 2 trial in acute schizophrenia increased the effect size (Kott et al, 2021).

- Erratic ratings could be at least partially explained by the increased use of peripheral (lowest or highest) scores of the item score ranges.

- Peripheral scoring tendency represents an incorrect scoring technique where raters exhibit a bias toward use of peripheral scores in preference to scores from the middle of the scoring range.

- In the current post-hoc analysis we assessed whether raters associated with erratic ratings use peripheral scores significantly more than raters not associated with erratic ratings.
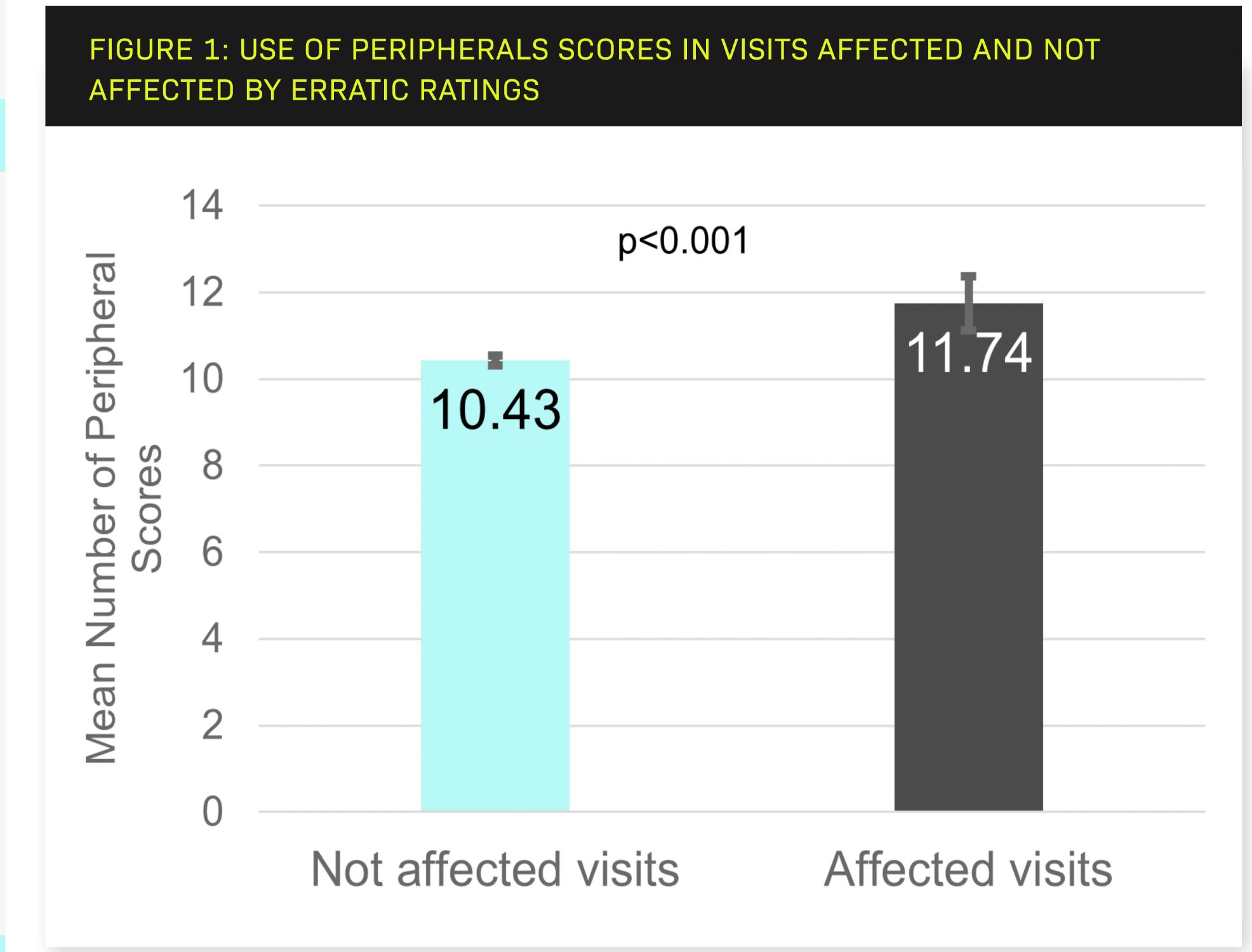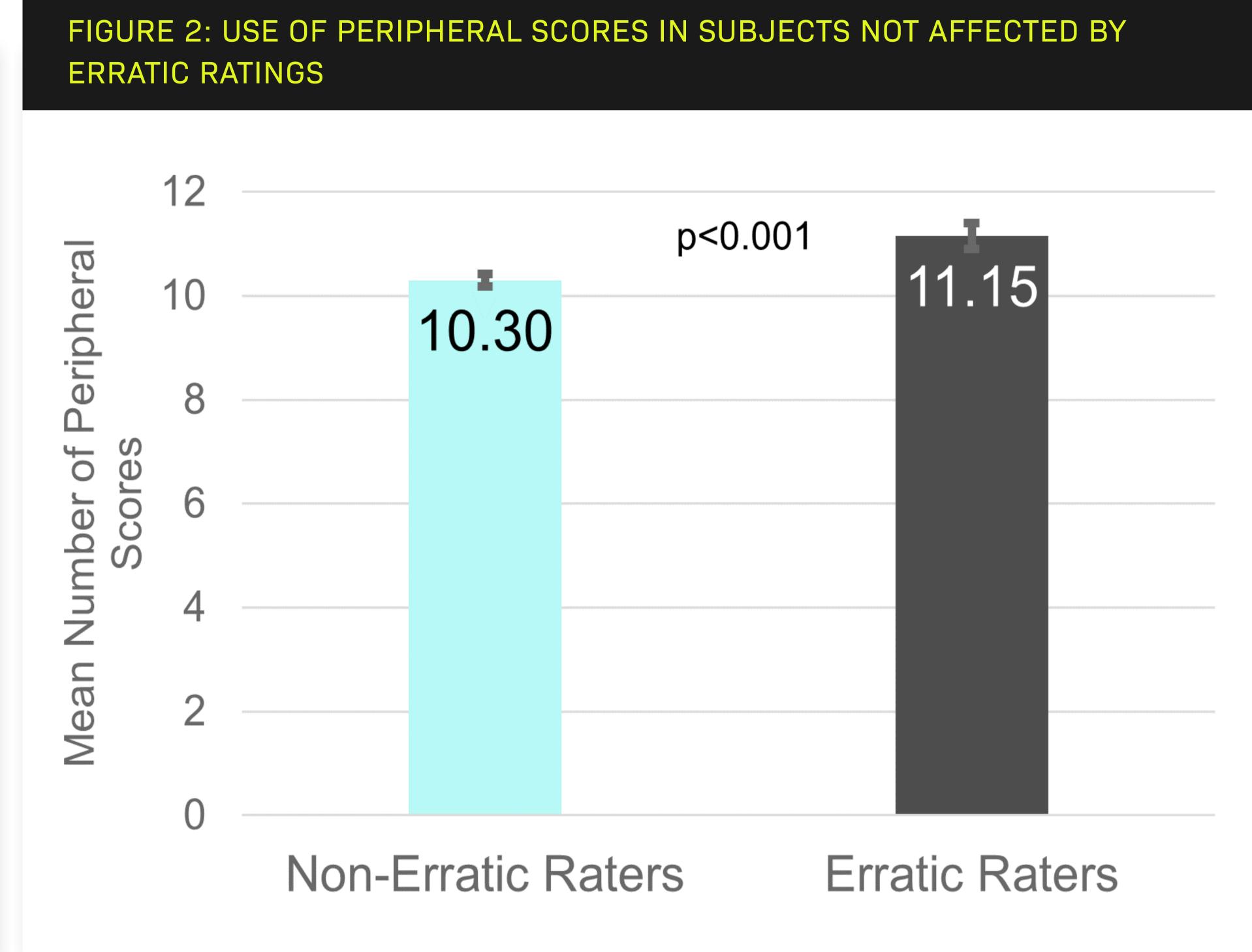
METHODS

- Individual PANSS item level data were obtained from a database of 15 acute schizophrenia clinical trials.

- Erratic ratings were operationally defined as at least 3 consecutive visits where the PANSS total score changed by at least 15 points in opposite directions.

- Peripheral scores were operationally defined as item scores of 1 (absent), 2 (minimal), 6 (severe), or 7 (extreme).

- Erratic-raters were operationally defined as those PANSS raters who were associated with at least 1 erratic rating.

- In the first step we tested whether erratic ratings were associated with an increased use of peripheral scores using generalized linear model.

- In the second step we tested whether erratic-raters tend to use more peripheral scores compared to the other raters. To do so, we analyzed data only from subjects not affected by erratic ratings using generalized linear mixed model.

RESULTS

- Our dataset consisted of 32,871 PANSS assessments that had the potential to be affected by an erratic rating.

- Of these, 323(0.98%) were affected by erratic ratings.

- A total of 1,524 raters rated the PANSS in our dataset, of these 220(14%) were erratic-raters.

- Peripheral scores were significantly more frequent in erratic ratings (p<0.001). Figure 1

- Raters associated with erratic ratings tended to use peripheral scores significantly more even in those subjects, who were not affected by erratic ratings (p<0.001).   Figure 2

CHARTS AND FIGURES

FIGURE 1: USE OF PERIPHERALS SCORES IN VISITS AFFECTED AND NOT AFFECTED BY ERRATIC RATINGS

FIGURE 2: USE OF PERIPHERAL SCORES IN SUBJECTS NOT AFFECTED BY ERRATIC RATINGS

CONCLUSION

- Erratic ratings with their demonstrated impact on signal detection represent a significant data quality concern.

- Erratic ratings are likely multifactorial in etiology. Some erratic ratings could accurately represent unusual symptom changes in a subpopulation of schizophrenic study participants while other erratic ratings could reflect correct scoring techniques in a subgroup of PANSS raters.

- Our post-hoc analysis of a large dataset identified that raters associated with erratic ratings tend to overuse peripheral scores even when rating subjects not affected by erratic ratings.

- We thus hypothesize that erratic ratings in part, at least, represent artificially exaggerated ratings of subtle symptom fluctuations, a hypothesis we plan to test further.

- Raters involved in erratic ratings should be remediated on their scoring technique and carefully monitored throughout the study.

- In instances of accumulation of erratic ratings with a single rater across multiple subjects, these raters may be considered to be replaced.

REFERENCES

- Umbricht D, Kott A, Daniel DG. The Effects of Erratic Ratings on Placebo Response and Signal Detection in the Roche Bitopertin Phase 3 Negative Symptom Studies—A Post Hoc Analysis. Schizophrenia Bulletin Open. 2020;1(1):203. doi:10.1093/schizbullopen/sgaa040.

- Kott A, Brannan S, Wang X, Daniel D. The Impact of Aberrant Data Variability on Drug–Placebo Separation and Drug/Placebo Response in an Acute Schizophrenia Clinical Trial. Schizophrenia Bulletin Open. 2021;2(1):295. doi:10.1093/schizbullopen/sgab037.

SIGNANT HEALTH
THE EVIDENCE GENERATION COMPANY™