Estimating the Likelihood of Two Raters Assessing the Same Patient Exactly the Same Using the PANSS

Kott, A1; Daniel, DG2
1Bracket, Prague, Czech Republic; 2Bracket, McLean, VA, USA

INTRODUCTION

We have recently identified an unusually high proportion of identical scorings across consecutive visits (all 30 PANSS items identical) in clinical trial data from schizophrenia studies. (Daniel and Kott, 2014)

The proportion was similarly high (approximately 5% of visits) for the situation when the same rater rated the subject across consecutive visits as for the situation when the subject was rated by two different raters.

In an analysis of 11 clinical trials and associated certification datasets, we identified the proportion of identical scorings in the trial dataset associated with a rater change to be dramatically higher compared to the proportion of identical scoring of videotaped interviews (required for certification to rate) (4.3% vs. 0.016%).

In the current analysis we estimated the likelihood two random raters would agree on all 30 PANSS items rating the same subject from available schizophrenia certification datasets.

METHODS

The likelihood of two random raters scoring the same subject exactly the same on all 30 PANSS items can be estimated from certification data. Raters intending to rate PANSS in a clinical trial are required to demonstrate inter-rater reliability with other raters in the study. This is typically demonstrated by rating a single videotaped patient interviewed utilizing the PANSS scale. As all raters rate the same interview, we can calculate the proportion of identical ratings between individual raters. Our dataset consisted of 73 certification datasets across 69 studies. Overall the data consisted of 18,308 datasets, we identified the proportion of identical scorings in the trial dataset associated with a rater change to be dramatically higher compared to the proportion of identical scoring of videotaped interviews (required for certification to rate) (4.3% vs. 0.016%).

In the current analysis we estimated the likelihood two random raters would agree on all 30 PANSS items rating the same subject from available schizophrenia certification datasets.

RESULTS I

In the whole dataset of 4,316,358 rating pairs there were 1,771 (0.04%) rating pairs where all 30 PANSS items were 100% identical

Across the 73 certification datasets the mean proportion of identical scorings was 0.041% (SD= 0.05) [Figure 1]

CONCLUSIONS

The PANSS is a complex instrument requiring integration of information obtained from patient report, direct observation of behavior and informant report.

We have previously identified identical scorings within the same subject across consecutive visits associated with rater change to be unexpectedly high in study data compared to certification data.

With respect to the certification data, our current analysis expands on our previous findings and confirms that the likelihood that two random raters would agree on the same PANSS scores in the ideal situation of no information variance (e.g. rating the same videotaped interview) is extremely small.

Our certification data as well indicate that the vast majority of identical scorings occur within individual countries. This finding is consistent with sharing of information in the process of scoring the certification videotape. In contrast, identical scoring patterns in the certification process are exceedingly rare among raters from different countries.

The implication of our results for quality monitoring of clinical trial data are that occurrences of identical scoring of the PANSS across visits, especially when associated with a change in raters should be investigated as a potential marker of poor data quality. In the certification process, identical scoring should be investigated as a potential sign of non-independent scoring.

REFERENCES


Disclosure: Dr. Kott and Dr. Daniel are employed by Bracket Global, LLC.