Exploring the association of PANSS rater change with extreme within-subject variability in schizophrenia clinical trials

Kott, A.; Wang, X.; Daniel, D.

CRF Bracket

The Methodological Question Being Addressed

In schizophrenia trials does a change of rater result in increased variability in scoring the PANSS?

Introduction

Measurement of subjective endpoints requires that multiple raters of the same patient be well synchronized in interview and measurement technique. We hypothesized that within person variance (WPV) and mean square successive difference (MSSD) could provide valuable information on within-subject variability when multiple raters are used. We have previously identified increased MSSD or WPV as markers of potentially compromised data quality. Additionally, sites with accumulation of subjects with increased WPV were shown to have an increased placebo response and decreased drug placebo separation. In the current analysis we wanted to explore whether extreme values of MSSD or WPV were associated with rater change.

Methods

Data from 24 clinical trials in schizophrenia were pooled together. Mean square successive difference and within person variance was calculated for every of the 9,661 subjects for each of the PANSS factors. Within each study, subjects with the MSSD above the 90^{th} percentile were identified as subjects with high MSSD and subjects with MSSD below 10^{th} percentile as subjects with low MSSD. Similarly subjects were stratified on the WPV. The association between the presence of rater change and the extremes in MSSD or WPV were explored using chi-square test.

Results

The presence of either low MSSD or low WPV was significantly more likely to be observed in the same rater group than in the group where there was a rater change at the level of alpha = 0.05.

High MSSD was significantly more likely to be observed in the group with rater change for the negative, positive, disorganized, and anxiety factor scores but not for the hostility factor score. The presence of high WPV was significantly associated with the presence of rater change for the negative, disorganized and anxiety factor scores but not the positive and hostility factor scores.

Discussion

The results strongly support the need for rater training to specifically focus on synchronization of interview and measurement technique among raters who cross-cover for each other at the same research site. Moreover, several additional important observations can be inferred from our results. The presence of rater change is associated with an increase in within-subject variability, but this increase is not seen consistently across all 5 PANSS factors. The fact that the hostility and to a lesser degree the positive factors are not associated with significantly increased variability in the presence of rater change indicate these factors are relatively immune to the differences between the individual raters. There may be numerous reasons for the observed differences between the factors such as the clarity of the individual items offering clear guidance on scoring or the fact that the positive symptoms are easier to identify and score compared to the negative or disorganized

symptoms. The association of extremely low within subject variability with the absence of rater change may indicate that raters tend to be more influenced by their previous ratings and are therefore reluctant to change their scores even in the presence of true symptom changes. Further analyses need to happen to fully understand these findings.

Disclosure

All authors report potential conflicts, all are full time employees of CRF Bracket.