Subjective Interview Quality Metrics Predict Scoring Performance in the Administration of the PANSS

Selam Negash PhD1, Barbara Echevarria PhD1, Lisa Stein PhD1, Elisabeth Prochnik MSW1, Michael T. Ropacki PhD1
1 MedAvante-ProPhase, Inc.

INTRODUCTION

- With the increasing globalization of clinical trials involving a large number of trial sites and raters, standardization of clinical assessments is a significant challenge.
- Inconsistent administration of clinical outcome assessments (COA) leads to scoring variability and resultant error variance in collected data, as well as potentially contributing to high placebo response rates and the possibility for inconclusive results.
- Quality control metrics that evaluate raters’ interview quality have been shown to be associated with drug-placebo separation.
- The extent to which other subjective metrics intended to improve data quality in psychiatry trial COA administration is less well-studied.
- The Positive and Negative Syndrome Scale (PANSS), one of the most widely used instruments in psychiatry trials, is a complex instrument to administer, with 30 different items, each with its own unique description, rating guidelines, scoring anchors and conventions.
- PANSS raters are required to consult several sources during administration and scoring, including informant reports, resulting in a cumbersome process that can impede the interview and is prone to errors.
- The goal of the present study was to investigate whether subjective interview quality metrics help minimize errors in PANSS administration and scoring in schizophrenia trials, and the extent to which the use of Virgil – an electronic clinical outcome assessment (eCOA) platform with real-time clinical guidance to standardize assessments – helps improve subjective interview quality.

METHODS

- A total of 1,651 PANSS assessments in three randomized, double-blind, multisite and multinational clinical trials of treatments for schizophrenia were reviewed. A subset of these assessments was scored using paper and pen (n = 288) and the rest utilized Virgil eCOA.
- After conducting the subject interview, raters completed PANSS scoring based on standard individual scoring anchors and conventions for each of the 30 items.
- A cohort of expert calibrated Central Review clinicians examined video/audio recordings and source documents to identify raters’ scoring discrepancies on the PANSS and assess their overall interview quality.
- Interview quality assessment was based on defined guidelines associated with good interview skills: adequate follow-up, clarification, neutrality, comprehensiveness, adherence to scale questions, and rapport; each interview was evaluated as “meets criteria” or “does not meet criteria.”
- The percentages of reviews with two or more discrepancies as a function of interview quality were analyzed.

RESULTS

- 113 (7%) of the assessments did not meet criteria for interview quality (Figure 1). Of these assessments, 101 (89%) had two or more scoring errors.
- In contrast, of the 1,538 assessments that met criteria for interview quality, 473 (31%) had scoring errors.
- There was a significant association between interview quality and scoring errors, [estimate = 8.2, SE = 0.56; p < 0.0001].
- The percentages of reviews that did not meet criteria on interview quality also increased as the number of errors per assessment increased (Figure 2).
- In a comparison of paper versus electronic administration (Figure 3) the proportion of assessments that did not meet interview quality criteria was significantly higher for paper and pen than Virgil eCOA; X² = 141.3, p < 0.0001.

CONCLUSION

- This study showed that subjective interview quality metrics are predictive of in-study scoring performance, and can serve as a useful tool in ongoing rater monitoring in PANSS administration and scoring.
- Rater feedback and remediation on assessments that do not meet quality criteria also help improve data quality.
- The use of Virgil eCOA significantly improved subjective interview quality, thereby reducing scoring errors and ultimately the error variance in PANSS administration and scoring.
- These findings highlight the value of subjective interview quality metrics and Virgil eCOA administration to improve standardized assessment and scoring of the PANSS, thereby increasing signal detection in global schizophrenia clinical trials.

References