Improvement in Ratings Behavior Over Time in Four Separate Placebo Controlled MDD Trials

Busner, J, Randall, P, Wilson, E, Tummala, R

ABSTRACT

The Methodological Question Being Addressed: Can rater errors improve over time in clinical trials and is such improvement dependent upon visit type (baseline vs. non-baseline)?

INTRODUCTION (AIMS): Recent work has shown baseline to later visit improvements in internal consistency (Cronbach's alpha) of efficacy measures in MDD, schizophrenia, and bipolar disorder clinical trials that included external rater oversight and remediation interventions. This is consistent with work showing decreased rater scale usage errors in initial (i.e., screening or baseline) compared with final subject visits. It is unclear whether and to what degree such improvements are due to factors unique to problematic baseline/screening visits as has been suggested or due to time, independent of particular visit. We conducted a retrospective analysis of rater errors pre and post chronologic midpoints for four separate multinational clinical trials. The methodology allowed us to examine errors as a function of time rather than visit type.

METHODS: Four large separate industry sponsored double-blind placebo-controlled clinical trials were examined retrospectively. Each trial had made use of daily computerized clinical data quality checks for potential rater scale misuse, with external oversight and clinical remediation of site raters provided throughout the trials as errors occurred.

RESULTS: The proportion of baseline visits in the first chronological halves of the four studies was 48%, 45%, 50%, and 56%, respectively, and increases in internal consistency (Cronbach's alpha) of efficacy measures were shown baseline to later visit improvements in internal consistency of ratings at post-baseline visits compared with ratings at screening/baseline visits. Others have noted better clinician agreement with patient reports at post-baseline compared with baseline visits. Proportionately more baseline visits may occur in the initial months of trial enrollment. It is unknown if rater improvement in post-baseline visits is due to unique properties of baseline visits or increased time/experience in the trial. The question is important methodologically in our understanding of the role of education/remediation throughout trials: if education/remediation was effective, one would expect to see improved performance as a function of time rather than a simple decrease of problematic baseline visits.

CONCLUSIONS: In four individual industry sponsored trials, clinical data errors lessened significantly as a function of time in surveillance program. As baseline visits were well distributed across time periods, the findings suggest that time in program improves performance, independent of visit type. The findings lend support to the provision of ongoing rater oversight and remediation throughout the course of large multinational trials as a means of improving ratings behavior.

REFERENCES


DISCLOSURES

Dr. Busner is a full time employee of Bracket; Dr. Randall is a full time employee of Quintiles; Mr. Wilson and Dr. Tummala are full time employees of AstraZeneca.