How Do US Clinical Research Sites Compare with Rest of World in Interview and Ratings Quality?

David G. Daniel¹ and Alan Kott²

¹Bracket, McLean, VA, USA; ²Bracket, Prague, CZ
How Do US Clinical Research Sites Compare with Rest of World in Interview and Ratings Quality?

David G. Daniel1 and Alan Kott2
1Bracket, McLean, VA, USA; 2Bracket, Prague, CZ

INTRODUCTION

As drug-placebo differences have diminished over time clinical trials sites in the United States have come under increasing scrutiny (1,2). We compare the quality of interview and ratings procedures delivered by North American (NA) vs. rest of world (ROW) sites using pooled results of video/audio surveillance measures from ten international schizophrenia clinical trials.

METHOD

A proprietary video/audio recording system was utilized to record PANSS rating procedures in over time clinical trials sites in the United States utilized to record PANSS rating procedures in clinical trials conducted in North America, Europe, South America and Asia. Sites uploaded video or audio recorded ratings assessments for review by calibrated external reviewers. External reviewers provided feedback on ratings quality on an ongoing basis to the site and sponsor. Prior to study initiation, raters were trained at investigator meetings by highly inter active procedures, including slide presentations, rating of videotaped patient interviews, and in some cases, interview and rating of live actors trained to portray schizophrenia symptoms. Instruments used to assess ratings and interview quality included the Research Interview Assessment Scale (RISA) (3) and a new tool, the Rater Quality Questionnaire (RQQ) (4).

RESULTS

SUMMARY OF RESULTS I: 1901 videotaped PANSS administrations at sites were graded for interview quality by external reviewers using the RISA. Mean total RISA scores of the North American (NA) raters (27.9 +/- 3.00) were moderately, but statistically significantly lower (indicating lower quality) than rest of world raters (ROW) (28.3 +/- 2.54) (t(1899)=2.94, n=1901, p<0.01).

SUMMARY OF RESULTS II: 1616 videotaped PANSS administrations at sites were rated for ratings quality consisting of two domains addressing: 1) the overall quality of the patient and/or informant material collected for this subject visit? and 2) the overall quality of the ratings or diagnostic evaluation (including proper application of the rules and anchor points of the rating scale or structured interview) for the current subject visit. Each domain is evaluated on a Likert-like scoring. A score of 2 represents deficiencies that have a minor impact on interview quality or scoring. A score of 3 reflects serious deficiencies in either collecting the information or rating the symptoms observed.

SUMMARY OF RESULTS III: 1616 videotaped PANSS administrations at sites were rated for proper application of the anchor points and other instructions of the measurement tool by external reviewers using the RQA global ratings quality axis. Mean total RQQ scores of the North American (NA) raters (1.27 +/- 0.47) were not statistically significantly different than rest of world raters (ROW) (1.30 +/- 0.49) (t(1614)=1.26, n=1616, p=NS).

DISCUSSION

North American raters scored as well or better than their ROW counterparts on the RQQ which evaluates the quality of information collected by interview and adherence to rating scale rules.

On the RQA total score which evaluates a broad range of interview behaviors (eg, ability to clarify, adherence to structured interview guidelines, handle questions that may induce placebo response) NA raters scored modestly worse than their ROW counterparts.

This report is preliminary as data continues to be collected in ongoing clinical trials.

REFERENCES