A Post-hoc Examination of Patient Characteristics: Implications for Enhanced Medical Surveillance

Bethanne Friedmann, PsyD[1], Henry J. Riordan, Ph.D. [1], Erin B. Kornsey, MS[1], Kenneth Williams, MD[1], Kevin Cox, MD[2], Neal R. Cutler, MD[2], Michael F. Murphy, MD, PhD[1]

[1]Worldwide Clinical Trials, King of Prussia, PA, [2]Worldwide Clinical Trials, Beverly Hills, CA

Abstract

Introduction

Recent studies in psychiatry have been marked by failed as well as negative trials. One contributory factor may be the enrollment of subjects that technically meet eligibility requirements including threshold scores for disease severity, who nevertheless are suboptimal candidates for study participation when viewed through a prism of historical and clinical information not captured by a study database.

Methods

The appropriateness of acutely exacerbated schizophrenic patients randomized to a completed study assessing the safety and efficacy of a novel antipsychotic was retrospectively, and independently, evaluated by three clinicians, blinded to country, site, patient identity, and treatment. During this process, medical history available through the study data base was amplified by examination of source documents addressing the subject's background according to predefined criteria (e.g., living situation, employment history), clinical history (treatment, diagnosis, trends, medications), as well as information on trial procedures (e.g., PANSS interview process). Approximately 13% (26) of all randomized subjects were selected from participating centers. Two psychiatrists and one psychologist each with over ten years of psychiatric trial experience independently rendered an opinion regarding the subject's appropriateness for inclusion, and all opinions were aggregated for final review.

Results

The three raters agreed on 88% (23/26) of subjects with an overall rater agreement of 92.3% for the 26 subjects for study appropriateness. A kappa value for multiple raters was considered to have "Very Good Agreement" according to Altman's conventions (Fleiss kappa = .7912, SE = .11, 95% CI .56-1.0) reflecting high concordance. Five subjects were deemed to be inappropriately randomized by all three evaluators although all technically met inclusion/ exclusion criteria.

Conclusions

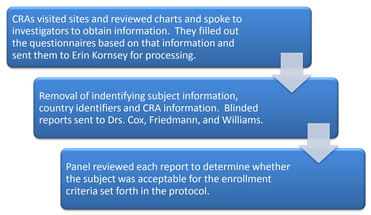
This exercise suggested that an investigator's judgment regarding subject suitability was consistent with third party blinded reviewers presented with an enriched database. Nevertheless, inappropriate subjects can be identified in spite of nominal compliance with protocol eligibility requirements, and independently of conventional efforts to assure congruency in diagnostic and disease severity assessments. A more finely detailed suite of inclusion/ exclusion criteria coupled to "real-time" monitoring of subject characteristics by blinded clinicians would provide an index of site sophistication that could identify problematic sites where enhanced medical monitoring activity is warranted.

Background

There are many possible reasons for the decreasing drug-placebo differences in schizophrenia trials which seem to be growing larger over time. One reason involves diagnostic accuracy and the certainty that the subjects randomized actually reflect the spirit and not just the letter of the protocol. It is important not just for subjects to meet all relevant inclusion/exclusion criteria but also have requisite characteristics regarding remote clinical history, past participation in other trials, and full treatment history not typically covered by screening investigators.

Methods

- 200 subjects at 20 sites in Eastern Europe and Russia participated in a double-blind inpatient trial to measure the effectiveness of a novel antipsychotic agent versus an active comparator.
- Twenty six of 200 possible subjects were selected from 8 participating centers in Croatia, Russia, Serbia and Ukraine reflecting a range of PANSS outcome scores.
- Study monitors retrieved source documents including medical history, subject's background according to predefined criteria (e.g., living situation, employment history), clinical history (treatment, diagnosis, trends, medications), as well as information on trial procedures (e.g., PANSS interview process).
- Subjects were not interviewed as part of this exercise but site staff augmented existing data according to a predefined questionnaire covering information related to socioeconomic, work/living status, medication, overall behavior on the treatment unit, change in level of psychosis, remote drug abuse, and various information regarding parameters of the PANSS interview (e.g., time of day, structure, location, length of interview, documentation and collateral sources).
- Three mental health clinicians (two psychiatrists and one psychologists), who were blinded to subject identity, country of origin, and treatment independently evaluated whether the subject was appropriate for the trial making a binary decision.
- Inter-rater agreement of independent reviewers was calculated using Fleiss kappa which measures the reliability of agreement between multiple raters when assigning categorical ratings.

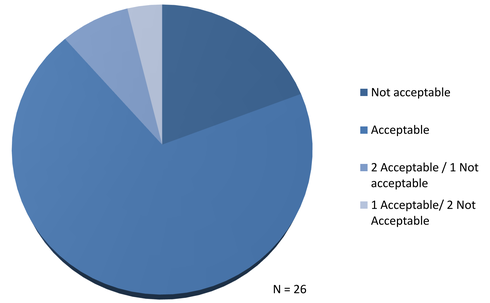


Site Number	Subject Number	Rater 1	Rater 2	Rater 3
2	00001	Acceptable	Not acceptable	Acceptable
2	00005	Acceptable	Acceptable	Acceptable
2	00006	Acceptable	Acceptable	Acceptable
2	00009	Acceptable	Acceptable	Acceptable
4	00009	Acceptable	Acceptable	Acceptable
4	00007	Acceptable	Acceptable	Acceptable
4	00003	Not acceptable	Not acceptable	Not acceptable
7	00003	Not acceptable	Not acceptable	Not acceptable
7	00004	Not acceptable	Not acceptable	Not acceptable
12	00017	Acceptable	Acceptable	Acceptable
12	00002	Acceptable	Acceptable	Acceptable
12	00008	Acceptable	Acceptable	Acceptable

Site Number	Subject Number	Rater 1	Rater 2	Rater 3
15	00002	Acceptable	Acceptable	Acceptable
15	00003	Acceptable	Not acceptable	Acceptable
15	00009	Acceptable	Acceptable	Acceptable
15	00015	Acceptable	Acceptable	Acceptable
17	00015	Acceptable	Acceptable	Acceptable
17	00011	Acceptable	Acceptable	Acceptable
17	00013	Not acceptable	Not acceptable	Not acceptable
18	00010	Acceptable	Acceptable	Acceptable
18	00012	Acceptable	Acceptable	Acceptable
18	00013	Acceptable	Acceptable	Acceptable
22	00002	Not acceptable	Not acceptable	Acceptable
22	00007	Not acceptable	Not acceptable	Not acceptable
22	00014	Acceptable	Acceptable	Acceptable
22	00009	Acceptable	Acceptable	Acceptable

Results

The three raters agreed on 88% (23/26) of subjects with an overall rater agreement of 92.3% for the 26 subjects for study appropriateness. However, 5 of these 26 subjects were felt to be inappropriate for randomization. A kappa value for multiple raters was considered to have "Very Good Agreement" according to Altman's conventions or Excellent Agreement beyond chance" according to Fleiss (Fleiss kappa = .7912, SE = .11, 95% CI .56-1.0) both reflecting high concordance.



Conclusions

- This retrospective surveillance exercise suggests that is important to independently ascertain the appropriateness of subjects for randomization who may meet all protocol mandated inclusion/exclusion criteria but may still not be optimal for the study when an expanded data set is considered using all source documentation.
- Multiple independent blinded clinicians with access to source documentation but not other expert opinion facilitates this process.
- Despite limited access and remoteness from sites, independent clinical evaluations in an international trial exhibit very high concordance.
- Extrapolation to all patients randomized suggest that > 81% of 210 subjects would be considered optimal for inclusion.
- Prospective confirmation of patient eligibility prior to randomization using expert opinion may enhance assay sensitivity; retrospective review of patients already randomized can facilitate remediation efforts at a site for future subjects.

No conflicts of interest exist in the research and development of this poster