

Use of Remote Centralized Raters Via Live 2-Way Video in a Multicenter Clinical Trial for Schizophrenia

Joan Shen, MD, PhD,* Kenneth A. Kobak, PhD,† Yonggang Zhao, PhD,*
Madeline M. Alexander, PhD,† and John M. Kane, MD‡

Abstract: Factors associated with clinician assessment may play a role in the increasing rate of failed trials. The use of centralized raters, a small group of highly skilled, tightly calibrated, and continuously monitored raters linked to the study sites through videoconferencing can address these issues by reducing the sheer number of raters involved, using rigorous calibration procedures not logistically feasible with a larger dispersed group of raters, and by blinding raters to visit and protocol. This phase 2 study was the first randomized controlled trial to use centralized raters in a study of treatments for schizophrenia. Subjects (N = 313) from 32 sites were randomly assigned to 6 weeks of treatment with 1 of 2 doses of an investigational antipsychotic, olanzapine 15 mg, or placebo. Subjects were evaluated weekly using the Positive and Negative Syndrome Scale. Data from the olanzapine (n = 68) and placebo (n = 68) arms were provided by the sponsor. The mean Positive and Negative Syndrome Scale change was significantly greater with olanzapine (-15.2) than placebo (-4.43), $P = 0.002$. The significant difference was apparent at week 1. The effect size was 0.48. Internal consistency was high throughout the study. Scores at screening were normally distributed and not skewed toward the cutoff score. Results found that hospitalized patients with schizophrenia were willing and able to participate in clinical trials using remote interviews conducted via videoconference. This methodology shows enormous promise for use in clinical trials, even with acutely psychotic patients.

(*J Clin Psychopharmacol* 2008;28:691–693)

The growing rate of failed and negative central nervous system (CNS) trials has led to increased examination of clinical trial methodology.¹ Factors associated with clinician assessment, such as expectancy bias, rating inflation due to enrollment pressure, poor interrater reliability, and poor interview quality, may play a role in the failure of marketed and effective novel compounds to separate from placebo.² One proposed solution has been the use of centralized raters,

that is, a small group of highly skilled and tightly calibrated raters who are linked to the various study sites through videoconferencing. The use of centralized raters to remotely administer the primary outcome measure can improve reliability by simply reducing the number of raters involved (eg, 8–10 vs 60–75 raters). They can be rigorously calibrated using procedures not logistically feasible with a larger group of raters at diffuse study sites. Being divorced from the study site eliminates the pressure to enroll patients or inflate baseline scores. Blinding to study visit reduces expectancy bias.

The use of centralized raters in schizophrenia trials raises several issues, such as whether acutely psychotic patients would be willing or able to participate and whether raters using videoconferencing would be able to detect a significant drug effect.

The current phase 2 study is the first randomized controlled trial to use centralized raters in a study of treatments for schizophrenia.

METHODS

Subjects (N = 313) from 32 sites, aged 18 to 65 years, with an acute exacerbation of schizophrenia (mean duration of current episode = 103 days) were included. After a 1-week (± 3 days) washout, subjects were randomly assigned to 6 weeks of treatment with 1 of 2 doses of an experimental compound, active comparator (olanzapine), or placebo. (Use of antipsychotic medication was stopped 1 day before baseline.) Subjects were evaluated weekly using the Positive and Negative Syndrome Scale (PANSS) by a centralized rater who was connected to the study site by high-speed videoconference. Raters were blinded to the study visit and protocol and were provided with data on the subject's behavior from an informant who observed the subject during the previous week. A different rater typically saw the subject at each visit. Results from the central raters' PANSS interviews were faxed back to the sites within 30 minutes. The site principal investigator managed the patient clinically and was responsible for all study-related clinical and protocol decisions. Whereas the site (not the central) rater was responsible for monitoring adverse events, if any were disclosed during the central rater's PANSS interview, the central rater forwarded this information to the sites. The sites then followed up with the patient if necessary for clinical management and to make a determination regarding classifying this information.

*Wyeth Research, Collegeville, PA; †MedAvante Inc, Hamilton, NJ; and ‡The Zucker Hillside Hospital, Glen Oaks, NY; and The Albert Einstein College of Medicine, Bronx, NY.

Received March 24, 2008; accepted after revision September 3, 2008.

This project was funded by Wyeth Pharmaceuticals.

Address correspondence and reprint requests to Kenneth A. Kobak, PhD, MedAvante Inc, 100 American Metro Blvd, Suite 106, Hamilton, NJ, 08619. E-mail: kkobak@medavante.net.

Copyright © 2008 by Lippincott Williams & Wilkins

ISSN: 0271-0749/08/2806-0691

DOI: 10.1097/JCP.0b013e31818c9ba3

Study subjects were scheduled as usual by the sites; the sites then simply called a toll-free number and informed the central raters of the date and time of the next visit. Although a 1-day advance notice was requested when possible for screening visits, and 2 to 3 days' notice for follow-up visits, in fact, 8% were conducted on the same day they were scheduled, 34% were conducted within 2 days of scheduling, and 58% were conducted with 3 or more days' advance notice.

The study was reviewed and approved by an institutional review board. In addition, an independent Data and Safety Monitoring Board regularly reviewed study data for significant trends or findings.

Remote Videoconferencing Technology

Videoconference interviews were conducted using 2 H.323 IP standard-based Polycom iPower (Polycom, Pleasanton, Calif) videoconferencing systems. The 2 devices were connected via T1 line over a virtual private network (VPN), run at a minimum of an industry standard bit rate of 384 kilobits per second. Use of a private T1 line over a VPN guarantees both availability and high bandwidth to ensure speed and picture quality. The system was Health Insurance Portability and Accountability Act-compliant and occurred via an encrypted connection over the VPN, thus protecting patient confidentiality.

The remote centralized rater had the ability to control all functions of both the rater-side and the patient-side cameras, including zoom-out, and up/down/left/right movements. This allowed the centralized rater to observe both whole-body views to assess motor movements and close-ups to assess facial expression and affect. Camera presets afforded the rater instant, 1-button camera relocation to a predetermined position, such as patient head-and-shoulder or patient full-body views. Camera zoom is internal, so no camera movement was visible to the subjects.

Central Raters

The central raters ($n = 18$) all had experience working with patients with schizophrenia (mean, 8.4 years) and

administering the PANSS (mean, 4.3 years). Twelve (66%) of the raters had doctoral degrees, and 6 (33%) had master's degrees. Before rating patients in this study, all raters went through a rigorous training and calibration procedure involving both didactic training on scale administration and scoring conventions, and applied training on clinical administration of the scale.³ Interrater reliability (intraclass correlation) was 0.91.⁴ Raters were monitored at regular intervals throughout the study to ensure ongoing calibration and maintenance of good clinical skills. Previous studies have found good correlations between PANSS and Brief Psychiatric Rating Scale ratings done remotely via videoconferencing and those done face-to-face with acutely psychotic patients.^{5,6} Administratively, central raters are viewed by regulatory bodies in the same way a central laboratory is viewed and are included as such on the study 1572 form.

Due to the confidential nature of data associated with the experimental compound, the study sponsor provided for publication only data from the active comparator (olanzapine) ($n = 68$) and placebo ($n = 68$) arms.

RESULTS

A significant difference in changes from baseline in PANSS total score was found between olanzapine and placebo starting at week 1, which continued throughout the study (Fig. 1). At end point, the adjusted mean change seen with olanzapine (-15.2 points; SE, 2.52) was significantly greater than the mean change for placebo (-4.43 points; SE, 2.51), $P = 0.002$, from an analysis of covariance model with last observation carried forward. A significant difference ($P = 0.021$) was also observed from the mixed model with observed cases. The effect size at end point was 0.48. Internal consistency was high throughout the study (0.92 and 0.89 at end point for olanzapine and placebo, respectively). Scores at screening were normally distributed and not skewed toward the cutoff score. Overall, 1993 remote PANSS assessments were completed by centralized raters ($n = 18$) during the 13-month study. Of the 1993 assessments, 15 (0.76%) experienced temporary interruptions because of technical issues, which were resolved, and the interviews were completed. In 10 cases (0.5%), the interview could not be completed and had to be rescheduled.

DISCUSSION

Results indicate that acutely psychotic patients with schizophrenia are willing and able to participate in clinical trials using remote interviews conducted via videoconference. Internal consistency and reliability were high, and scores were not inflated at baseline by blinded raters. This in itself could reduce the placebo response that has been seen in recent randomized controlled trials. Centralized raters detected significant differences between active drug and placebo, with an effect size similar to or better than those reported in the literature for similar dose with the active comparator. This methodology shows enormous promise for use in clinical trials. Recent reports have noted an increasing placebo response rate in schizophrenia trials,⁷ as well as evidence of baseline inflation,⁸ similar to that seen in depression and anxiety trials.

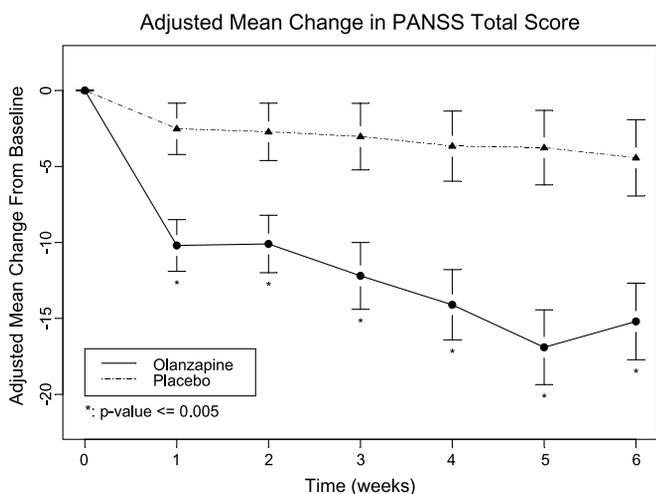


FIGURE 1. Adjusted mean change in PANSS total score by study week (analysis of covariance).

This makes the search for new methodologies to overcome these phenomena critical. In this study, subjects were screened and diagnosed for inclusion into the study by the site raters. Because this procedure is also subject to bias, the use of centralized raters for screening and diagnosing patients before they enter studies is another potential use of this methodology and could help ensure the enrollment of appropriate patients in large-scale multisite studies.

AUTHOR DISCLOSURE INFORMATION

Dr Kobak is an employee of MedAvante, which provides centralized rating services. Dr Alexander is a consultant to MedAvante, and Dr Kane is Chief Scientific Advisor to MedAvante. Drs Shen and Zhao are employees of Wyeth, who funded the study research.

REFERENCES

1. Khan A, Kolts RL, Rapaport MH, et al. Magnitude of placebo response and drug-placebo differences across psychiatric disorders. *Psychol Med.* 2005;35:743–749.
2. Kobak KA, Kane JM, Thase ME, et al. Why do clinical trials fail? The problem of measurement error in clinical trials: time to test new paradigms? *J Clin Psychopharmacol.* 2007;27:1–5.
3. Kobak KA, Engelhardt N, Williams JB, et al. Rater training in multicenter clinical trials: issues and recommendations. *J Clin Psychopharmacol.* 2004;24:113–117.
4. Kane J, Schmidt J, Kobak KA. The use of centralized raters in multicenter clinical trials for schizophrenia. Presented at the International Society for CNS Clinical Trials and Methodology, 3rd Annual Scientific Congress; February 20–21, 2007; Washington, DC.
5. Zarate CA Jr, Weinstock L, Cukor P, et al. Applicability of telemedicine for assessing patients with schizophrenia: acceptance and reliability. *J Clin Psychiatry.* 1997;58:22–25.
6. Yoshino A, Shigemura J, Kobayashi Y, et al. Telepsychiatry: assessment of televideo psychiatric interview reliability with present- and next-generation internet infrastructures. *Acta Psychiatr Scand.* 2001;104:223–226.
7. Alphs LD. Historical perspective on drug-placebo differences: implications for future trial design. Presented at the International Society for CNS Methodology Mid-Year Conference; September 17–18, 2007; Brussels, Belgium.
8. Schoemaker JH, Gaur R, Berkowitz L, et al. Expert rater assisted score evaluation (ERASE): a new method to enhance signal detection in randomized, placebo-controlled, clinical trials. Presented at the International Society for CNS Trials Methodology 4th Annual Meeting; February 25–26, 2008; Washington, DC.