BACKGROUND

The ultimate goal of any clinical trial is drug-placebo separation at study endpoint. This can be challenging in Alzheimer’s disease (AD) clinical trials – particularly those that assess disease modifying drugs - as they tend to be longer, multi-national, and have more rater turnover and drift. Additionally, there is a greater likelihood of encountering potential raters whose experience with the primary efficacy instruments falls below the desired level. Previous research has shown that an enriched training curriculum coupled with an in-study ratings reliability program can enable these raters to both certify to participate, and perform comparably to their more experienced colleagues. The in-study ratings reliability program ensures that raters adhere to scale administration and study conventions at select visits. When errors were detected by calibrated clinicians, the raters in question were remediated and the scoring was corrected. Additionally, such a program has been shown to significantly decrease the frequency of rater errors over the course of a trial. (1)

OBJECTIVE

To determine if these programs (ie – enriched training and in-study ratings reliability), when combined, result in raters who provide data of sufficient quality to allow separation of drug effect from placebo control at study endpoint.

METHODS

Data from 582 subjects in 2 of 4 AD trials previously discussed were available for analysis. The quality of the co-primary endpoint (ADAS-Cog) administration and scoring in the “IDENTITY” study (LY450139) was evaluated at Baseline and Week 52 as part of an in-study ratings reliability program. Calculation of LSMeans was performed to determine if raters were able to separate drug from placebo at study endpoint (Week 76). (2) A mixed model where the change from baseline ADAS-cog score is the dependent variable and the baseline ADAS-cog score, Visit (Week 52, Week 76), Treatment (Drug, Placebo) and the interaction between treatment and visit are the independent variables. Country is also accounted for as a random variable. The LSMeans (and p-values) reported are the difference in change from baseline between Drug and Placebo at week 52 and then at week 76 (the interaction term in the model). The data were further analyzed to only compare ADAS-Cog14 scores from subjects for whom we had data at both the week 52 and 76 time points.

RESULTS

As previously reported, at 52 weeks, all raters, regardless of whether they were inexperienced and required enriched training or were experienced prior to the study, tended toward separating drug from placebo treatment on the ADAS-Cog. (2) An additional analysis examining drug-placebo separation only at week 76 via a difference of LS means methodology, also found a trend (p = 0.0864) toward separation, with drug treated subjects performing worse on the ADAS-Cog compared to placebo treated subjects. This was not an unexpected finding, as subjects in this study tended to do worse on drug at 18 months. (3) When both the week 52 and 76 visits were accounted for in a mixed model, the p-values for difference in LSMeans for treatment (Drug vs. Placebo) are (p=0.947) at 52 weeks and (p=0.022) at 76 weeks. (Fig 1).

CONCLUSION

In this study, while there was a strong trend toward separation of drug from placebo at 52 weeks, this difference became statistically significant at study end (week 76). This is particularly notable given that, for a variety of reasons (eg – study stopped dosing, patients were not doing well and withdrew from study participation) there were fewer patients at week 76. In this analysis, only half of the subjects participating at week 52 remained in the study at week 76.

A customized In-Study ratings reliability program can both identify instances where raters deviate from proper ADAS-Cog administration and/or scoring technique, and by remediating raters in question, can effectively decrease the likelihood of errors recurring over the course of an AD clinical trial. Enriched training enabled less experienced raters to perform comparably to their more experienced peers. Therefore, both programs warrant consideration in AD trials, particularly given their ability to produce raters who, in multi-national trials, were able to separate drug from placebo.

REFERENCES


The authors wish to acknowledge the assistance of Cynthia McNamara, Ph.D.1 in the editing of this poster.