Trials, Errors, and Placebo Prediction: Tradeoffs between effect size and sample size in machine learning models to mitigate the placebo response

Ariana E. Anderson [a], Richard Keefe [bc], Dong Jing Fu[d], Adam J. Savitz[de], Robert M. Bilder [a]
[a] Department of Psychiatry and Biobehavioral Sciences, University of California, Los Angeles
[b] Department of Psychiatry and Biobehavioral Sciences, Duke University, Durham, NC
[c] VeraSci, Durham, NC
[d] Janssen R&D, NY, NY
[e] Weill Cornell Medical College, White Plains, NY

The Methodological Question Being Addressed:
The placebo response may be predictable across studies using machine learning models gleaned from similar trials. Here, we assess how incorporating cross-study placebo models predicting the placebo response may change the effect size estimates and sample size requirements for subsequent trials.

Introduction (Aims)
To counteract the placebo response, some trial designs eliminate subjects who previously responded to placebo treatment. Here, we assess whether predictive modeling of placebo response and medication response using data from previous trials would allow smaller targeted trials of purified patient subgroups, by screening out those patients for whom the medication is *unlikely* to be beneficial thereby eliminating the need for a lead-in placebo intervention. We hypothesize that targeted screening would lead to an increase in statistical power in assessing whether medications are effective.

Methods
In 3,647 patients with schizophrenia covering 11 different clinical studies, we trained machine learning models to predict the expected placebo response (placebo quantified response score [PQRS]) and the expected medication response given the PQRS. These models were trained and tested across studies using leave-one-study-out cross validation. We assessed (1) whether the predicted placebo risk (PQRS) predicted the actual placebo response; (2) whether the PQRS predicted a patient's response to an active medication; and (3) whether pre-trial screening of predicted placebo responders and predicted medication non-responders would increase the effect size seen in the remaining purified sample. Several different trial designs were compared, which combined simulated screening of predicted placebo non-responders ($< 75^{th}$ percentile PQRS), screening out predicted drug non-responders ($>25^{th}$ percentile), and adjusting for a subject's total treatment response by their placebo risk (PQRS), resulting in the comparison of five alternate trial designs. The trial analyses were run for each modified design predicting the treatment response using baseline PANSS score, gender, illness duration, patient age, and the intervention. For a subset of the models, the PQRS was also included as a covariate to assess whether this factor increased separation of the drug and the placebo interventions. Omega-squared, eta-squared, and partial eta-squared were used to compare the effect sizes of the treatments for the original analyses and the 5 modified analyses, across the 10 different studies for which there was a placebo arm.

Results

On average, about 40% of subjects were removed from the study when using the dual screening procedure of omitting predicted drug non-responders and placebo responders, and 25% were removed when screening out only predicted placebo responders. The PQRS was highly predictive of the actual placebo response and the overall treatment response across studies. The average gain in effect size from the modified trial designs ranged from .3% to 7%. The larger effect sizes in the modified trials permitted a reduction in sample size to obtain equivalent study results. For some studies with different demographic populations (e.g. geriatric) than the studies used to train the models, the placebo models showed a weakened ability to predict the actual placebo response. These results suggests that training and testing populations must be closely matched to predict accurately how someone will respond to placebo interventions.

Conclusions

The placebo response can be anticipated based upon machine learning models learned in other studies, and modified trial designs which incorporate screening procedures and adjustment of treatment response for placebo risk showed modest gain in effect sizes in simulated trials. The primary benefit to these study designs may be in their ability to reduce the sample sizes needed to observe clinical efficacy.