Evaluating Drug Efficacy: Leveraging Machine Learning Insights For Placebo Response Modeling

Joseph Geraci[1,2,3,4], Bessi Qorri[1], Luca Pani[5,6], Larry Alphs[1]

[1]NetraMark Corp, [2]Department of Pathology and Molecular Medicine, Queen's University, [3]Centre for Biotechnology and Genomic Medicine, Medical College of Georgia, Augusta University, [4]Arthur C. Clarke Center for Human Imagination, School of Physical Sciences, University of California San Diego , [5]Miller School of Medicine, University of Miami, [6]University of Modena and Reggio Emilia

INTRODUCTION

PLACEBO RESPONSE CHALLENGE IN PSYCHIATRIC DISORDER CLINICAL TRIALS

Placebo response in clinical trials often leads to trial failure, necessitating larger sample sizes, prolonged recruitment time, and high study costs.

AIM

Leverage machine learning (ML) algorithms to identify characteristics of drug and placebo response across psychiatric clinical trials, spanning bipolar disorder, anxiety, and schizophrenia.

METHODOLOGICAL ISSUE BEING ADDRESSED

We explore the impact of ML-generated placebo response modeling on clinical trial design and patient enrichment strategies.

METHODS

DATASETS

- **Bipolar 1 Disorder Trial:** 378 patients evaluating the therapeutic efficacy of an investigational drug for acute depressive disorder using clinical scales (NCT01467700).
- **Phase III Anxiety Trial:** 171 active patients and 161 controls featuring ~100 independent variables per subject.
- **Phase IIa Schizophrenia Trial:** 87 patients randomized to the placebo arm and 48 to the study drug treatment, with 128 variables per subject including clinical scales and physiological measurements.

MACHINE LEARNING APPROACH

A novel ML algorithm was used to analyze efficacy, demographics, and safety data for predicting placebo responses. Independent variables were extracted from initial patient assessments, while dependent variables were based on trial outcomes. The algorithm used has the capability to deconstruct the patient population into explainable and unexplainable subpopulations.

- Explainable subpopulations are used to generate 'personas' characterized by 3-12 statistically significant variables.
- Drug effects are distinguished from placebo by applying the same model to the active and placebo arms of a trial. Comparative analysis of predicted response rates elucidates the degree of drug efficacy relative to placebo.

RESULTS

Identifying markers of treatment responders and non-responders in a bipolar disorder trial (378 Patients)

Clinical scale data (MADRS, HAM-A, YMRS, CTSS-M) with primary endpoint of 50% improvement in MADRS from baseline. Study drug did not show separation from placebo.

- Training data: A 115 patient training set allowed for the identification of an explainable subpopulation of 71 placebo patients using 6 CTSS clinical scale items and 2 YMRS items
- Models tested on 239 independent drug participants with baseline CTSS-M due to no significant differences from placebo
- ***ML model correctly predicted placebo responders 87% of the time***
- ***Accurately identified 39/44 drug non-responders; falsely identified 5/44 non-responders***

Variable	Explanation
↓CTSS 49	NR had less desire to get better
↓CTSS 15	NR think less of meds
↓CTSS 38	NR struggled for less time
↓CTSS 01	NR forget more often
↓CTSS 41	NR felt worse about signing up
↑CTSS 50	NR had more faith in talk therapy
↑YMRSSC	Total NR high
↑YMRS 04	NR slept slightly better

Identifying markers of responders and non-responders in a failed phase III anxiety trial (332 Patients)

Placebo Response

8 variables (CTSS-Q16, Q18, Q20, Q21, Q25, Q26, Total CTSS, and HAMD1) captured:

- 55/73 PR and 50/88 PNR (mixed class)
- 42% of PNR were highly explainable with an accuracy of 74%

Drug Response

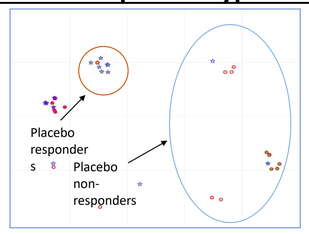
- Drug response was very poor and acted much like a placebo except on a small class of patients. ***Specifically, 6 variables (CTSS-Q6, Q7, Q9, Q18, Q21, and HAMD – Work and Activities) explained 10 DNR***

Inclusion Criteria	Expected Effect on Study Population
CTSS Q18 ≥ 4	Removes most DNR but keeps many PR, eliminating PNR
CTSS Q25 ≥ 2	Removes many PR and keeps many PNR Removes most DR and DNR with high screen fail rate
CTSS Q26 ≥2	Removes many PR and keeps most PNR; more DNR
CTSS Q9 ≥4	Removes many DNR and keeps most DR; more PR remain
CTSS Q7 ≥7	Removes many DNR and keeps most DR No cut off score significantly reduces relative number of PR
Total Score CTSS ≥40	Any cut off seems to remove PR, PNR, DR, and DNR similarly

Identifying markers of responders and non-responders in a phase IIa schizophrenia trial (135 Patients)

Primary Endpoint: 10% PANSS improvement over placebo

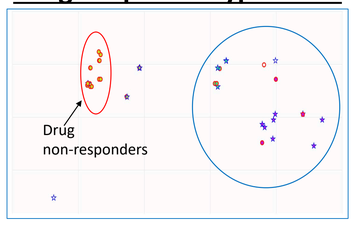
Placebo Response Hypothesis



~50% of PR characterized by:

- Total baseline depression scale <1
- Supine respiration rate ≤16.5
- Emotional withdrawal (PANSS) ≥2
- Disorientation (PANSS) ≥2

Drug Response Hypothesis



~37.5% DNR characterized by:

- Attention (PANSS) =1
- Judgement and intuition (PANSS) =1
- Cognition at baseline <5

Inclusion Criteria	Expected Effect on Study Population
Depression Scale ≥1	Removes most PR; Little effect on drug treatment group, but removes more DNR
Supine Respiration >16	Removes most PR; removes a few more DNR; High screen failure rate
Emotional Withdrawal ≤2	Removes most PR; removes more DNR; Increases screen failure rate
Disorientation ≤1	Removes many PR; Removes most DNR; Increases screen failure rate
Attention > 0	Removes DNR and many PR; Some increase in screen failure rate
Judgement and Intuition ≥1	Removes most DNR and PR; Increases screen failure rate
Cognitive Total ≥6	Removes most DNR and PR; Big increase in screen failure rate

CONCLUSIONS AND SIGNIFICANCE

By fracturing data sets into explainable subsets, this AI avoids overfitting and provides powerful insights:

- This methodology can be used to discover precision I/E criteria to enrich subsequent trials.
- The data from the schizophrenia trial revealed actionable insights that led to I/E criteria that could greatly improve p-values in subsequent trials.
- The anxiety drug lacked the efficacy to be improved and these techniques reflected this shortcoming while avoiding overfitting

REFERENCES

1. Smith *et al.* Using artificial intelligence-based methods to address the placebo response in clinical trials. *Innov Clin* Neurosci (2022) 19(1-3):60-70.
2. Hafliðadóttir, S. H. *et al.* Placebo response and effect in randomized clinical trials: meta-research with focus on contextual effects. *Trials* **22**, (2021)
3. Smith, E. A. *et al.* Using Artificial Intelligence-based Methods to Address the Placebo Response in Clinical Trials. *Innov Clin Neurosci* **19**, 60 (2022).
4. Tsay, Mi., Geraci, J. & Agrawal, A. Next-Gen AI for Disease Definition, Patient Stratification, and Placebo Effect. doi:10.31219/OSF.IO/PC7AK.
5. Morlock, R. & Geraci, J. Placebo Response Propensity Scale (PRPS) Scores and Probability of Placebo Response by Disease Severity.

DISCLOSURES

Joseph Geraci is the founder of NetraMark and is a significant shareholder of NetraMark Holdings, which is a publicly traded company. Bessi Qorri, Luca Pani, and Larry Alphs are employed by NetraMark

Luca Pani's disclosures (past 3 years): AbbVie, USA; Acadia, USA; Alexion, Italy; BCG, Switzerland; Boehringer Ingelheim International GmbH, Germany; Compass Pathways, UK; EDRA-LSWR Publishing Company, Italy; Ferrer, Spain; Gedeon-Richter, Hungary; GLG-Institute, USA; Immunogen, USA; Inpeco SA, Switzerland; Ipsen-Abireo, France; Johnson & Johnson USA; NeuroCog Trials, USA; Novartis-Gene Therapies, Switzerland; Sanofi-Aventis-Genzyme, France and USA; NetraMark, Canada*; Otsuka, USA; Pfizer Global, USA; PharmaMar, Spain; Relmada Therapeutics, USA*; Takeda, USA; Vifor, Switzerland; WCG-VeraSci/Clinical Endpoint Solutions, USA (*options/shares)

NETRAMARK