Principal Investigator

First Name: Mariam
Last Name: Khayretdinova
Degree: MSc
Primary Affiliation: Brainify.AI
E-mail: mariam@brainify.ai
State or Province: Delaware
Country: United States

General Information

Key Personnel (other than PI):
First Name: Andrey
Last name: Zhdanov
Degree: PhD
Primary Affiliation: Brainify.AI
SCOPUS ID:
Requires Data Access? Yes

First Name: Polina
Last name: Pshonkovskaya
Degree: MSc
Primary Affiliation: Brainify.AI
SCOPUS ID:
Requires Data Access? Yes

Are external grants or funds being used to support this research?: No external grants or funds are being used to support this research.
How did you learn about the YODA Project?: Data Holder (Company)

Conflict of Interest

https://yoda.yale.edu/wp-content/uploads/2024/06/YODACoI.pdf
https://yoda.yale.edu/wp-content/uploads/2024/07/Khayretdinova-M.-CoI.-YODA-Project.pdf
https://yoda.yale.edu/wp-content/uploads/2024/07/Zhdanov.A_YODA_CoI.pdf

Certification

Certification: All information is complete; I (PI) am responsible for the research; data will not be used to support litigious/commercial aims.
Data Use Agreement Training: As the Principal Investigator of this study, I certify that I have completed the YODA Project Data Use Agreement Training

1. NCT02417064 - ESKETINTRD3001 - A Randomized, Double-blind, Multicenter, Active-controlled Study to Evaluate the Efficacy, Safety, and Tolerability of Fixed Doses of Intranasal Esketamine Plus an Oral Antidepressant in Adult Subjects With Treatment-resistant Depression
2. NCT02418585 - ESKETINTRD3002 - A Randomized, Double-blind, Multicenter, Active-controlled Study to Evaluate the Efficacy, Safety, and Tolerability of Flexible Doses of

Intranasal Esketamine Plus an Oral Antidepressant in Adult Subjects With Treatment-resistant Depression
3. NCT02422186 - ESKETINTRD3005 - A Randomized, Double-blind, Multicenter, Active-controlled Study to Evaluate the Efficacy, Safety, and Tolerability of Intranasal Esketamine Plus an Oral Antidepressant in Elderly Subjects With Treatment-resistant Depression
4. NCT03039192 - 54135419SUI3001 - A Double-blind, Randomized, Placebo-controlled Study to Evaluate the Efficacy and Safety of Intranasal Esketamine in Addition to Comprehensive Standard of Care for the Rapid Reduction of the Symptoms of Major Depressive Disorder, Including Suicidal Ideation, in Adult Subjects Assessed to be at Imminent Risk for Suicide
5. NCT03097133 - 54135419SUI3002 - A Double-blind, Randomized, Placebo-controlled Study to Evaluate the Efficacy and Safety of Intranasal Esketamine in Addition to Comprehensive Standard of Care for the Rapid Reduction of the Symptoms of Major Depressive Disorder, Including Suicidal Ideation, in Adult Subjects Assessed to be at Imminent Risk for Suicide

What type of data are you looking for?: Individual Participant-Level Data, which includes Full CSR and all supporting documentation

Research Proposal

Project Title

Placebo Prediction in Major Depressive Disorder Using AI/ML

Narrative Summary:

This project aims to use advanced technology to predict how patients with Major Depressive Disorder will respond to a placebo treatment. By employing machine learning, a type of artificial intelligence, the goal is to analyze the clinical data to foresee which patients might show improvements simply because they believe they are receiving actual treatment, when in fact they have been receiving placebo. We plan on learning, which clinical features can potentially have impact on higher placebo responses (e.g. age, sex, etc).This predictive ability can help doctors make better decisions about patient care, design overall treatment plans more effectively, and ultimately improve patient outcomes.

Scientific Abstract:

Background: Depression affects over 350 million individuals globally, with 30% of those diagnosed with Major Depressive Disorder (MDD) experiencing treatment-resistant depression (TRD). The placebo effect complicates clinical trials by masking the true efficacy of new antidepressants.
Objective: To develop and validate a machine learning model using clinical data to predict placebo responders in MDD patients with an accuracy of over 75% balanced accuracy.
Study Design: Retrospective study using clinical data from vivli.org, focusing on placebo-controlled trials involving MDD patients.
Participants: Individuals diagnosed with MDD, aged 18 to 80 years.
Primary and Secondary Outcome Measures: Prediction of placebo response, defined as a significant reduction in symptoms (50% response) or remission, based on clinical measures like HAM-D and MADRS scores. Identification of demographic and clinical features correlating with placebo response.
Statistical Analysis: Data acquisition, preprocessing, feature selection, and machine learning model development using clinical data to predict placebo responses in MDD patients.

Brief Project Background and Statement of Project Significance:

Depression is a multifaceted condition that impacts over 350 million individuals globally. Its diagnosis and treatment are challenging due to the heterogeneous nature of its manifestations,

complicating the identification of effective therapeutic interventions. Notably, approximately 30% of individuals diagnosed with Major Depressive Disorder (MDD) exhibit treatment-resistant depression (TRD), emphasizing the necessity for treatments that are customized to the specific needs of each patient.

The placebo effect is a significant phenomenon in this context, wherein individuals experience symptomatic improvement following the administration of a treatment lacking active therapeutic ingredients, driven solely by the belief that they are receiving an actual treatment. This phenomenon underscores the powerful influence of psychological factors on health outcomes and presents a substantial challenge in clinical trials for novel antidepressants, as it can obscure the true efficacy of the investigational drug.

To address these challenges, researchers are increasingly utilizing clinical data to identify biomarkers predictive of this responses. By employing machine learning techniques to analyze clinical data, we aim to uncover patterns indicative of a patient's potential response to placebo treatments. By elucidating the mechanisms underlying the placebo effect and recognizing the unique clinical profiles of patients, we strive to advance the development of more efficacious and personalized therapeutic options for depression. This research is pivotal as it holds the potential to significantly improve treatment outcomes for individuals who have not achieved relief with existing therapies, thereby enhancing patient care and outcomes in the treatment of depression.

Specific Aims of the Project:

Main Hypothesis:

Placebo response can be predicted in MDD patients using machine learning models on a clinical data.

Specific Aims and Objectives:
1. To develop and validate a machine learning model using clinical data for predicting placebo responders
2. To predict placebo response in MDD with the accuracy of over 75% Balanced Accuracy (the average between sensitivity and specificity)

Study Design:

Methodological research

What is the purpose of the analysis being proposed? Please select all that apply.

New research question to examine treatment effectiveness on secondary endpoints and/or within subgroup populations

Confirm or validate previously conducted research on treatment effectiveness

Research Methods

Data Source and Inclusion/Exclusion Criteria to be used to define the patient sample for your study:

Data Source: The primary data for this study will be obtained from clinical trials available through the YODA Project, supplemented with data from vivli.org, focusing on placebo-controlled trials involving Major Depressive Disorder (MDD) patients.
Inclusion Criteria:
-Patients diagnosed with MDD
-Participants aged 18 to 80 years
-Participants enrolled in placebo-controlled trials.

-Availability of baseline and follow-up clinical measures of depression severity
Exclusion Criteria:
-Incomplete Data
-Non-Standard Treatment
We plan to pool data from the YODA Project with the following studies from Vivli: NCT00672620, NCT00073411, NCT01140906, NCT00635219, NCT00863798, NCT00369343.
Additionally, we have already been granted access to one dataset on Vivli: https://search.vivli.org/studyDetails/32981703-7602-41ef-bbde-68f002d4cff8.
We will be conducting the IPD analysis within the secure platform provided by the vivli. All work on participant-level data will take place within this secure platform.

Primary and Secondary Outcome Measure(s) and how they will be categorized/defined for your study:

Prediction of Placebo Response: A significant reduction in depression symptoms, quantified as a 50% or greater reduction in HAM-D or MADRS scores from baseline to follow-up.
Categorization: Participants will be categorized as responders (≥50% reduction) and non-responders (<50% reduction) based on HAM-D or MADRS scores.
Secondary Outcome Measures:
Identification of Demographic and Clinical Features Correlating with Placebo Response: Analysis of demographic (age, gender) and clinical characteristics (baseline severity of depression, prior treatment history) correlating with placebo response.Categorization: Features will be categorized based on their distribution and relevance to placebo response using logistic regression and machine learning model feature importance rankings.
Treatment Response Status: Classification based on treatment response status, including response (≥50% reduction in symptoms), partial response (25%-49% reduction), and non-response (<25% reduction).

Main Predictor/Independent Variable and how it will be categorized/defined for your study:

Main Predictor/Independent Variable:
Change in Clinical Questionnaires Data: The main predictor variable will be the change in clinical questionnaire scores from baseline to follow-up.
Categorization:
The change will be quantified using the Hamilton Depression Scale (HAM-D) and Montgomery–Åsberg Depression Rating Scale (MADRS) scores.
Participants will be categorized based on their response:
Responders: Those with a 50% or greater reduction in HAM-D or MADRS scores.
Partial Responders: Those with a 25%-49% reduction in HAM-D or MADRS scores.
Non-Responders: Those with less than a 25% reduction in HAM-D or MADRS scores.

Other Variables of Interest that will be used in your analysis and how they will be categorized/defined for your study:

Demographic Variables:

Age: Continuous variable or grouped (e.g., 18-30, 31-50, 51-70, 71-80).
Gender: Categorical variable (e.g., Male, Female, Other).
Clinical Variables:

Baseline Severity of Depression: Continuous variable or grouped (e.g., mild, moderate, severe).
Prior Treatment History: Binary variable (e.g., Yes/No) or detailed categories based on treatment types.
Treatment Variables
Type of Treatment Received: Categorical variable (e.g., Placebo, Active Treatment).
Duration of Treatment: Continuous variable or grouped (e.g., 8 weeks).
Symptom Scores

HAM-D Scores: Continuous variable.
MADRS Scores: Continuous variable.

Statistical Analysis Plan:

The analysis will incorporate a multi-step approach:

Descriptive Statistics to summarize the dataset's central tendencies, dispersion, and distribution shape.
Correlation Analyses to identify relationships between clinical level data and placebo responses.
Machine Learning Models including support vector machines (SVM) and neural networks (NN), chosen for their ability to handle high-dimensional data and complex nonlinear relationships.

We plan to test variables that could influence placebo response, including:

Demographic data: Age and gender
Clinical data: Baseline depression severity and prior treatment history

Type of Model:
The proposed model will be a hybrid, combining SVM and NN, to leverage SVM's efficacy in high-dimensional space and NN's capability for pattern recognition and prediction.

Model validation and training will involve:

Cross-Validation Techniques. To ensure the model's generalizability, we'll use k-fold cross-validation, dividing the data into k subsets and training the model k times, each time using a different subset as the test set and the remaining data as the training set.
Performance Metrics. Balanced accuracy, sensitivity, specificity, and area under the curve (ROC) will assess model performance, providing a comprehensive evaluation of its predictive capabilities.
Adjustments for Confounders. Statistical methods will be employed to adjust for potential confounders such as age, gender, and baseline depression severity, ensuring that the model's predictions are as accurate and unbiased as possible.

Our approach to handling missing values in the dataset will be twofold, depending on the nature and extent of the missing data. For variables with a minimal amount of missing data (less than 10%), we will employ imputation techniques to preserve the dataset's integrity and maintain statistical power. Common imputation methods, such as mean imputation for continuous variables and mode imputation for categorical variables, will be considered. For cases where the missing data is extensive or not random, those observations may be excluded from the analysis to prevent bias.

We plan to handle the outputs of our research to cover both: commercial application and academic dissemination. Whilst we do not plan on making our code public, there is a strong commitment to contributing to the scientific community through the publication of research findings in high-impact journals. The intention to publish in journals such as NeuroImage, Human Brain Mapping, and Journal of Affective Disorders reflects a strategy to ensure the research reaches a wide audience of researchers, clinicians, and scholars interested in the intersections of neuroscience, mental health, and machine learning.

Software Used:

Python

Project Timeline:

1.Project Preparation and Initial Setup - 1st month
September 1, 2024. Anticipated project start date after YODA and vivli approvals.
2. Data Acquisition and Initial Review - 1 month
3. Data Cleaning and Preparation- half month

4. Preliminary Data Analysis - 1 month
5. Main Data Analysis - 2 months
6. Results Interpretation and Validation- 1 month
7. Manuscript Drafting - 2 months
8. Manuscript Submission for Publication - April 2025
9. Report Results to YODA Project - May 2025
10. Possible Extension - June 2025

Dissemination Plan:

The study is anticipated to span 12 months, which would consist of secondary analysis (6 months), and result publications (disseminations) (6 months).

To expand on the dissemination plan, Brainify.AI team will submit findings to high-impact journals in psychiatry, neurology, and AI/ML. We are considering the following journals: NeuroImage, Human Brain Mapping, Journal of Affective Disorders.

Bibliography:

Previously, our team has published several articles in high-impact journals & conferences:
Khayretdinova, M., Shovkun, A., Degtyarev, V., Kiryasov, A., Pshonkovskaya, P., & Zakharov, I. (2022). Predicting age from resting-state scalp EEG signals with deep convolutional neural networks on TD-brain dataset. Frontiers in Aging Neuroscience, 14, 1367.
Shovkun, A., Kiryasov, A., Zakharov, I., & Khayretdinova, M. (2023, June). Optimization of the Deep Neural Networks for Seizure Detection. In ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (pp. 1-2). IEEE.
Khayretdinova, M., Zakharov, I., Pshonkovskaya, P., Adamovich, T., Kiryasov, A., Zhdanov, A., & Shovkun, A. (2024). Prediction of brain sex from EEG: using large-scale heterogeneous dataset for developing a highly accurate and interpretable ML model. NeuroImage, 285, 120495.

All cited material:

Fava, M., Freeman, M. P., Flynn, M., Judge, H., Hoeppner, B. B., Cusin, C., … & Papakostas, G. I. (2020). Double-blind, placebo-controlled, dose-ranging trial of intravenous ketamine as adjunctive therapy in treatment-resistant depression (TRD). Molecular psychiatry, 25(7), 1592-1603.
Fournier, J. C., DeRubeis, R. D., Hollon, S. D., Dimidjian, S., Amsterdam, J. D., Shelton, R. C., Fawcett, J. (2010). Antidepressant drug effects and depression severity: A patient-level meta-analysis.Journal of the American Medical Association, 303, 47-53.
Glaser, J. I., Benjamin, A. S., Farhoodi, R., & Kording, K. P. (2019). The roles of supervised machine learning in systems neuroscience. Progress in neurobiology, 175, 126-137.
Hamilton, M. (1960). A rating scale for depression. Journal of neurology, neurosurgery, and psychiatry, 23(1), 56.
Kelly, K., Posternak, M., & Jonathan, E. A. (2008). Toward achieving optimal response: understanding and managing antidepressant side effects. Dialogues in clinical neuroscience, 10(4), 409-418.
Lee, C. T., Palacios, J., Richards, D., Hanlon, A. K., Lynch, K., Harty, S., … & Gillan, C. M. (2023). The Precision in Psychiatry (PIP) study: Testing an internet-based methodology for accelerating research in treatment prediction and personalisation. BMC psychiatry, 23(1), 25.
Zhdanov, A., Atluri, S., Wong, W., Vaghei, Y., Daskalakis, Z. J., Blumberger, D. M., … & Farzan, F. (2020). Use of machine learning for predicting escitalopram treatment outcome from electroencephalography recordings in adult patients with depression. JAMA network open, 3(1), e1918377-e1918377.
DOI: 10.1001/jamanetworkopen.2019.18377