An exploratory machine learning approach to identify placebo responders in pharmacological binge eating disorder trials

Rahul K. Goyal[1] 🄳 | Shamir N. Kalaria[1] 🄳 | Susan L. McElroy[2] |
Mathangi Gopalakrishnan[1]

[1]Center for Translational Medicine, School of Pharmacy, University of Maryland, Baltimore, Maryland, USA

[2]Department of Psychiatry and Behavioral Neuroscience, University of Cincinnati, College of Medicine, Mason, Ohio, USA

Correspondence
Mathangi Gopalakrishnan, Center for Translational Medicine, University of Maryland School of Pharmacy, 20 North Pine St., Baltimore, MD, 21201 USA.
Email: mgopalakrishnan@rx.umaryland.edu

Abstract

Randomized, placebo-controlled trials for binge eating disorder (BED) have revealed highly variable, and often marked, rates of short-term placebo response. Several quantitative based analyses in patients with BED have inconsistently demonstrated which patient factors attribute to an increase in placebo response. The objective of this study is to utilize machine learning (ML) algorithms to identify moderators of placebo response in patients with BED. Data were pooled from 12 randomized placebo-controlled trials evaluating different treatment options for BED. The final dataset consisted of 189 adults receiving placebo with complete information of baseline variables. Placebo responders were defined as patients experiencing ≥75% reduction in binge eating frequency (BEF) at study end point. Nine patient prerandomization variables were included as predictors. Patients were divided into training and testing subsets according to an 75%:25% distribution while preserving the proportion of placebo responders. All analysis was performed in the software Pumas 2.0. Gaussian Naïve Bayes algorithm showed the best cross-validation accuracy (~64%) and was chosen as the final algorithm. Shapley analysis suggested that patients with low baseline BEF and anxiety status were strong moderators of placebo response. Upon applying the final algorithm on the test dataset, the resulting sensitivity was 88% and prediction accuracy was 72%. This is the first application of ML to identify moderators of placebo response in BED. The results of this analysis confirm previous findings of lesser baseline disease severity and adds that patients with no anxiety are more susceptible to placebo response.

Study Highlights

WHAT IS THE CURRENT KNOWLEDGE ON THE TOPIC?

Randomized, placebo-controlled binge eating disorder (BED) trials have revealed highly variable, and often marked, rates of placebo response. Approximately 50% of all investigator-led trials previously demonstrated a lack of separation between active treatment and placebo.

WHAT QUESTION DID THIS STUDY ADDRESS?

This study provides a machine learning-based approach to enrich future BED trials by identifying potential moderators of placebo response.

WHAT DOES THIS STUDY ADD TO OUR KNOWLEDGE?

By leveraging data from twelve BED investigator-led randomized clinical trials, this study identified that patients with less disease severity and patients who do not exhibit symptoms of general anxiety disorder are more susceptible to placebo response.

HOW MIGHT THIS CHANGE CLINICAL PHARMACOLOGY OR TRANSLATIONAL SCIENCE?

A machine learning approach using pooled data from similar clinical trials can provide recommendations for increasing probability of trial success in BED.

INTRODUCTION

In 2013, the Diagnostic and Statistical Manual of Mental Disorders officially recognized binge eating disorder (BED) as a distinct eating disorder that is defined by recurrent episodes of eating an unusually large amount of food over a discrete period of time in the absence of compensatory behaviors. The global pooled lifetime prevalence of BED is ~0.9%.[1] Although previous research indicates several sociodemographic risk factors, ~75% of patients with BED also have a comorbid psychiatric condition.[1,2] Given the disparities in healthcare accessibility, psychotherapy, the cornerstone treatment option, has been significantly underutilized. Currently, lisdexamfetamine is the only drug the US Food and Drug Administration (FDA) approved for the treatment for moderate to severe BED. Because not all patients receiving treatment will experience symptomatic improvement, there is a need to evaluate the safety and efficacy of promising pharmacotherapeutic options in randomized clinical trials.[3]

A significant proportion of the trials for BED demonstrate a lack of separation between active treatment and placebo. Approximately 50% of all investigator-led trials were therefore determined to be negative.[4] Randomized, placebo-controlled BED trials have revealed highly variable, and often marked, rates of placebo response.[4,5] Depending on the primary end point, placebo response is defined as a clinically significant reduction in binge eating symptomology in patients that receive placebo. Carter et al. identified a mean placebo response rate (proportion of patients with a >50%–75% reduction in binge eating frequency) of 33% based on a literature review of randomized placebo controlled trials.[6] The placebo response rate

in BED is also within the reported range (25%–50%) for mood disorders.[7–9]

Mixed results across studies make it difficult to understand whether BED is a chronic disease, a disease that waxes and wanes, or a disease that entirely remits in patients. In a natural progression study conducted by Fairburn et al., patients with BED that were not receiving treatment were followed for 5 years and evaluated for BED symptomology.[10] Only 24% of patients with BED were still diagnosed after 15 months of study follow-up and 9% after 60 months of study follow-up. However, it is important to note that the patients with BED enrolled were younger on average compared to the average age of patients enrolled in investigator-led and randomized clinical trials. Jacobs-Pilipski et al. further analyzed a randomized, placebo controlled, placebo run-in trial that evaluated the effect of sibutramine on BED.[5] During the baseline placebo run-in period, 33% of patients (147/451) were categorized as placebo responders (≥75% reduction in binge frequency from baseline). A higher proportion of placebo responders had reported less emphasis on shape and weight upon self-evaluation and domain-specific quality of life during the baseline period. At follow-up, patients attributed their placebo response to increased awareness of their eating, increased accountability, self-monitoring, motivation to change, social support, and positive expectations regarding placebo. However, most placebo responders reported experiencing binge eating symptoms at some point during the follow-up period. In contrast, a study conducted by Crow et al. identified that 93% of patients still reported BED symptomology after a 1-year follow-up.

Machine learning (ML) is an area of artificial intelligence that involves the construction and study of systems

that can learn from data.[11] The ultimate goal is to solve a given problem by applying knowledge acquired from an automatic learning process using past experiences and data. The use of ML approaches has recently been widespread in the field of neuropsychiatry for illnesses, such as depression, Alzheimer's disease, and schizophrenia.[12–14] ML has been applied to the detection and diagnosis of conditions, prediction of long-term prognosis, timing and selection of treatment, impact assessment on public health, and mental health resource allocation.[12–17] Data from neuroimaging studies, electronic health records, healthcare insurance claims, social media, and traditional clinical trials have been commonly used to develop algorithms. Applications of ML could expedite the path to personalized medicine due to the growing interest to identify potential objective biomarkers, biological mechanisms, and genetic predictors.

Several methods have been explored to identify placebo responders a priori to limit the magnitude of placebo response and increase the probability of clinical trial success. Traditionally, placebo run-in periods have been used to identify placebo responders for exclusion before trial randomization. This can lead to longer overall trial durations and need for larger sample sizes.[17–20] ML algorithms can be used to systematically identify placebo responders for trial exclusion during the screening period. This trial enrichment method could therefore increase the efficiency of future clinical trials. The use of ML in BED is relatively unexplored. Given that many studies have demonstrated negative outcomes, different advanced quantitative methods could be utilized to improve the treatment effect in BED clinical trials. The objective of this study was to apply supervised ML algorithms to identify moderators of placebo response in patients with BED.

METHODS

Data

Subject-level, longitudinal normalized binge eating frequency (BEF) and binge day frequency (BDF) from adult patients receiving placebo in 12 investigator led trials conducted by McElroy et al. were collected to develop the main dataset ($N = 288$). Clinical trial procedures are also previously described by McElroy et al.[21–32] All studies were approved by the Institutional Review Board and patients were appropriately consented at the time of enrollment. Predictors of interest included demographics and comorbid characteristics. However, due to differences in data collection methods across clinical studies, a reduced dataset was created to ensure no missing predictors ($N = 189$). Predictors in the final reduced dataset

included: age, gender, clinical global impression of severity at baseline, weekly BEFat baseline, lifetime diagnosis of major depressive disorder, current diagnosis of major depressive disorder, history of alcohol abuse, history of substance abuse, and diagnosis of general anxiety disorder. Placebo responders were identified as patients experiencing ≥75% reduction in BED at study end point. Studies were truncated to 6 weeks in length (length of the shortest BED clinical trial incorporated into the dataset) in order to avoid confounding responder status due to trial length. Patients were randomly divided into training and testing subsets according to a 75%:25% distribution while preserving the proportion of placebo responders.

Machine learning analysis

Model building

Various supervised ML algorithms were utilized to predict placebo response status. Supervised learning algorithms are based on developing a function that maps predictors to an outcome based on a labeled learning dataset. For the purpose of this study, random forest, support vector machines, decision tree, logistic regression, k-nearest neighbors (KNNs), Gaussian naïve Bayes (GNB), Bayesian quadratic discriminant analysis, and AdaBoost and XGBoost algorithms were applied to the final reduced dataset using the MLJ package in Pumas 2.0 (www.pumas. ai).[33,34] A 10-fold cross validation with grid search was repeated five times to quantify the average accuracy of the final model.[34] As the analysis dataset was reasonably balanced among the target classes (60% nonresponders and 40% responders), accuracy was chosen as the performance metric for the choice of the best model during training of each algorithm. Furthermore, the distribution of accuracy and Cohen's kappa per fold were graphed to discriminate across algorithms.

Inference and evaluation

After the best model for every algorithm was selected, contribution of the predictors on the response was quantified using Shapley values—a model agnostic method to assess predictor contribution for a given subject/instance.[35] Briefly, Shapley value is a marginal contribution (analogous to marginal log-odds) of every covariate on the prediction of the target class, where a high Shapley value implies a higher chance of placebo response. The variable with the highest mean absolute Shapley value across all training instances was considered as the most important variable for the purpose of prediction. Sensitivity analysis

was also conducted to evaluate the impact of defining responder status using BDF and reducing the percent threshold for defining placebo responder status on feature selection. Percent reduction of 50% and 60% were used to redefine placebo response status. Finally, the algorithms were further evaluated on the test dataset based on sensitivity, specificity, and accuracy. Sensitivity (also called true-positive rate) is the probability that the algorithm will predict a responder among those who are responders and specificity (also called true-negative rate) is the probability that the algorithm will predict a nonresponder among those who are nonresponders. Because the goal for trial enrichment is to limit the number of placebo responders from being randomized, sensitivity was chosen as the metric of interest to evaluate performance of the final model. The consequence of low sensitivity (large false-negative rates) implies a higher chance of trial failure due to a greater number of placebo responders being randomized. On the other hand, the consequence of low specificity (large false-positive rates) is a longer trial enrollment period due to of the exclusion of potential nonresponders.

RESULTS

Descriptive summary

Table 1 provides a summary of demographic and clinical features in patients receiving placebo across 12 investigator-led BED clinical trials. Most patients were women with an average weekly normalized BEF and BDF of 4.7 and 4.0, respectively. The average percentage change from baseline in BEF was ~56%, indicating a large placebo response. Placebo response rate was ~40%, 52%, and 61% when placebo response was defined as a 75%, 60%, and 50% reduction in BEF from baseline, respectively. Approximately 30% of patients experienced cessation of binges at the end of 6 weeks. Figure S1 provides a scatterplot matrix of continuous variables and Figure S2 provides a bar graph for categorical variables. Average baseline BED severity (BDF and BEF at baseline) was lower in placebo responders versus placebo nonresponders (BEF: 3.90 vs. 5.14, BDF: 3.51 vs. 4.26). Furthermore, the proportion of placebo responders was higher than nonresponders in patients diagnosed with general anxiety disorder (86% vs. 14%).

Machine learning analysis

A comparison of accuracy and Cohen's kappa of all supervised ML algorithms is provided in Figure 1. Model accuracy on the validation dataset was observed to be similar across all ML algorithms. The median accuracy ranged from 54% to 64%. The kappa statistic varied across all algorithms with the median kappa statistic ranging between −0.03 and 0.31. This suggests that the model performance in predicting placebo response was slight to fair in terms of strength of agreement. Numerically, model performance was the highest with GNB algorithm and lowest with KNN.

TABLE 1 Demographic and clinical characteristics of patients with BED receiving placebo

Characteristic	Patients with BED ($N = 189$)	
	N	%
Female	161	86
Major depressive disorder – lifetime	82	43
Major depressive disorder – current	44	23
Alcohol abuse – current	13	7
Substance abuse – current	6	3
Anxiety – current	28	15
Placebo responders	75	40
	Mean	**SD**
Age, years	40.5	10.9
Binge eating frequency at baseline, per week	4.65	2.44
Binge day frequency at baseline, per week	3.97	1.35
Average percent change from baseline in weekly binge eating frequency	−56.6%	47%
	Median	**Range**
Clinical global impression severity at baseline	5	3−6

Abbreviation: BED, binge eating disorder.
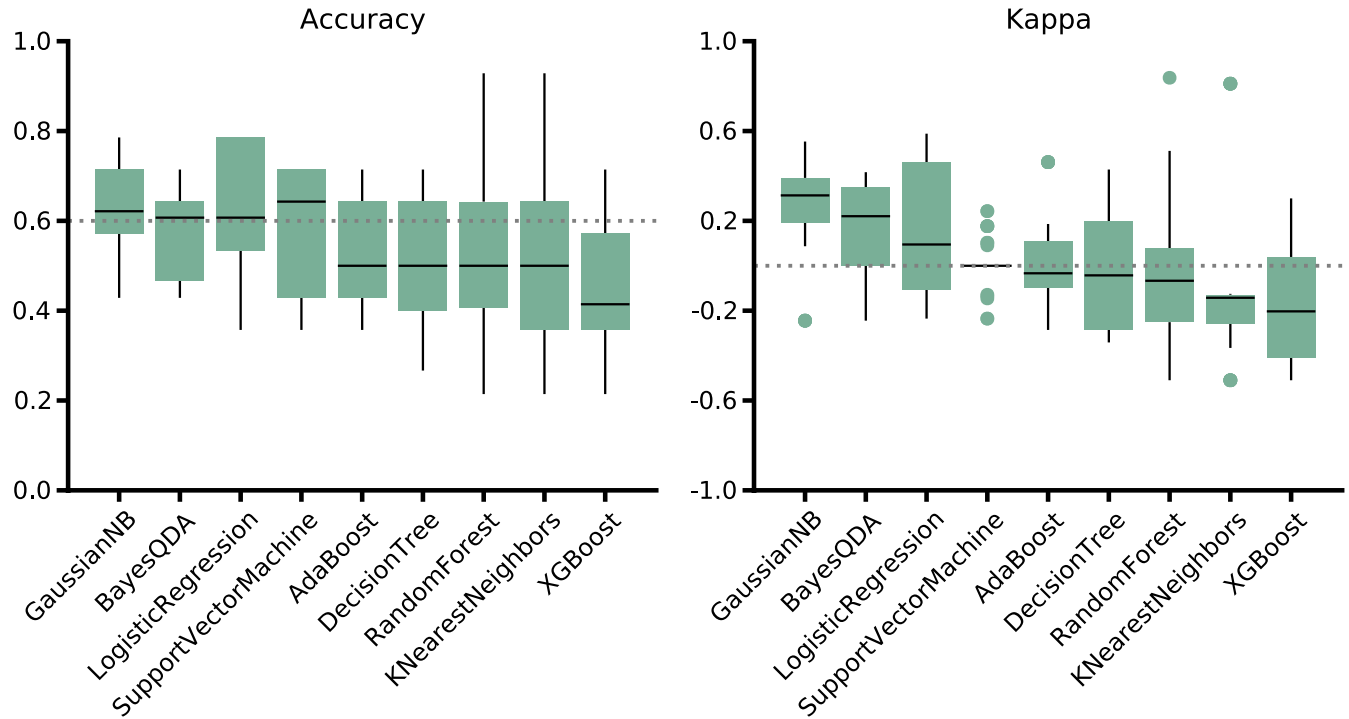
FIGURE 1 Evaluation of model performance during training and validation. Each boxplot is generated based on the distribution of 50 (5 repeats × 10 folds) cross validation performance metrics for each algorithm—accuracy or kappa. Dotted horizontal lines in each plot represent performance of a dummy classifier that would predict all subjects to be placebo nonresponders.
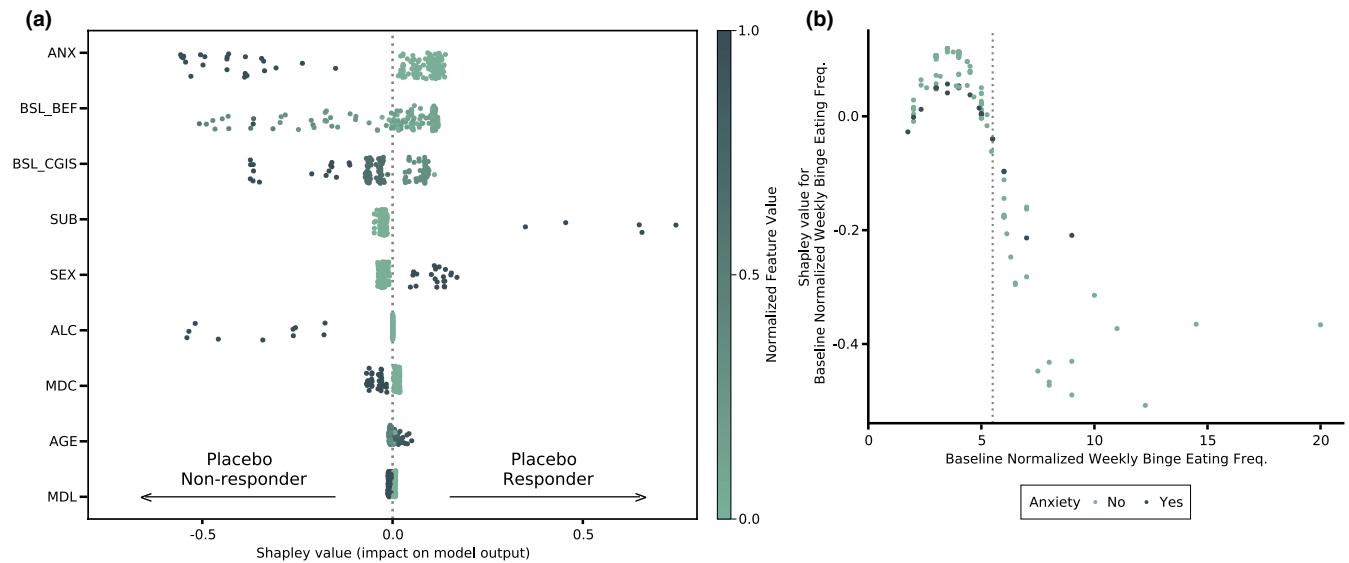
FIGURE 2 (a) Feature importance based on Gaussian Naïve Bayes Algorithm. ALC, alcohol abuse; ANX, diagnosis of generalized anxiety disorder; BSL_CGIS, baseline clinical global impression of severity; BSL_BEF, baseline normalized weekly binge eating frequency; MDC, current diagnosis of major depressive disorder; MDL, lifetime diagnosis of major depressive disorder; SUB, substance abuse. Points represents the Shapley value of the respective feature on the *y*-axis for every subject in the training dataset. Points are colored by normalized feature value. (b) Shapley dependence plot between baseline normalized weekly binge eating frequency by anxiety. Points represent the Shapley value of baseline normalized weekly binge eating frequency versus the value of the predictor for every subject in the training dataset. Points are colored by a yes/no flag for the diagnosis of anxiety.

Using the GNB algorithm, Shapley values were calculated for every feature for all training instances. A large and positive Shapley value means that odds are in favor of predicting a placebo responder. Anxiety and baseline

BED severity were the strongest predictors of placebo response (Figure 2a). In general, patients without anxiety had a higher likelihood of being a placebo responder as compared with patients with anxiety. Upon further

exploration with respect to the predictor interaction, when baseline BEF was <5.5, patients with anxiety had a lower likelihood of being a placebo responder (Figure 2b). Additionally, the correlation between anxiety and BEF was explored to test if anxiety was a potential confounder. The comparisons resulted in an insignificant *p* value at a significance level of 0.05 implying that anxiety is not correlated with BEF in the dataset. To test the robustness of the study results, the analysis was repeated using redefined placebo response criteria of at least 50% and 60% change from baseline in BEF and changing the primary outcome measure to BDF. The findings were similar to those reported with the predefined clinically significant placebo response criteria of 75% change from baseline in BEF, with anxiety and baseline BED severity identified as the strongest predictors.

A flow chart was created to summarize the incidence of placebo responders and nonresponders using the numeric cutoff for BEF baseline values obtained using (Figures 2b and 3). For patients with a BEF 5.5, the probability of experiencing a significant placebo response in the dataset is <14%. However, for a patient with a BEF <5.5 and no comorbid diagnosis of general anxiety disorder, the probability of experiencing a clinically significant placebo response is approximately 53%. On the other hand, an anxious patient with a BEF <5.5 is expected to have a probability of 17% in favor of experiencing a clinically significant placebo response.

The final models of all algorithms were applied to the testing dataset (25% of the final reduced dataset) and the sensitivity and specificity were calculated using the obtained confusion matrix (Table 2). GNB algorithm and Bayesian quadratic discriminant analysis showed >80%

sensitivity, however, the accuracy and specificity of the GNB algorithm were higher. Based on the GNB algorithm, 12% (sensitivity = 88%) of the test subjects who were responders were predicted to be placebo nonresponders (true-positive rates = 15, true-negative rates = 19, false-positive rates = 11, false-negative ratess = 2, and positive = placebo responder).

DISCUSSION

Given the continued high incidence of placebo response exhibited in randomized clinical trials, there have been relatively few successful BED drug development programs. Identification of predictors for placebo response can guide the selection of an intended population of interest and inform study design. ML approaches utilized in this study expand upon previous quantitative research in BED and considers interactions between predictors to better the understand the complexity of placebo response. This present study evaluated nine different ML algorithms to predict placebo response in patients with BED enrolled in small investigator-led trials. The overall placebo response rate observed across 12 studies was similar to other psychiatric conditions (i.e., major depressive disorder, generalized anxiety disorder, and schizophrenia).[36–38] Although the sample size of the final reduced dataset was relatively low and only a limited number of predictors were available for analysis, the GNB algorithm exhibited reasonable accuracy and demonstrated that baseline BED severity and comorbid anxiety are strong predictors for placebo response. Generally, patients that had a lower baseline severity were more likely to experience a clinically significant placebo

FIGURE 3 Flow chart of distribution of placebo responders and nonresponders in the dataset based on binge eating frequency. Baseline BEF is the baseline normalized weekly binge eating frequency. Text in bottom boxes shows majority class, percentage of patients in the majority class (number of subjects belonging to the majority class) and total number of subjects in that bucket.
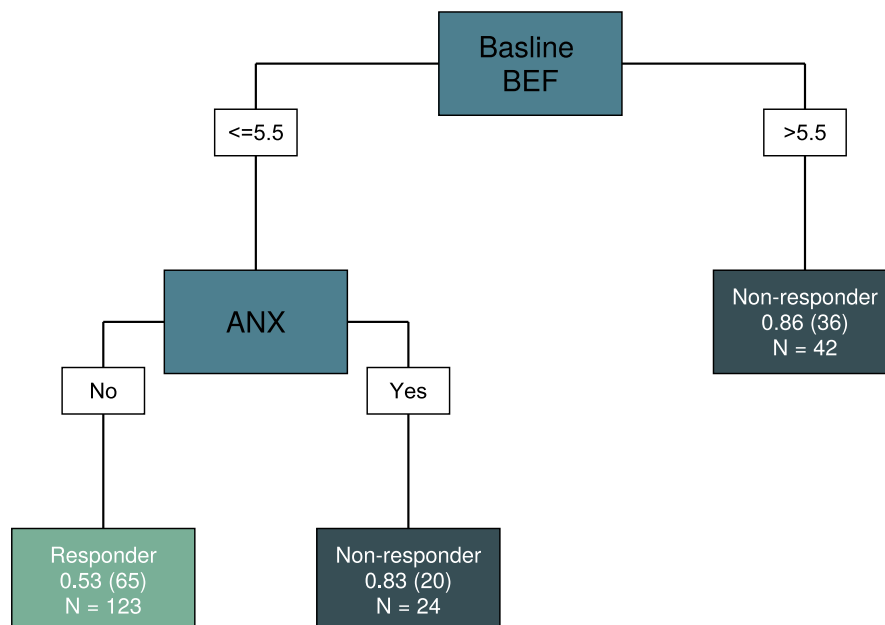
TABLE 2 Model performance on test dataset for each final supervised algorithm

Machine learning algorithm	Sensitivity	Specificity	Accuracy	Kappa
Gaussian Naïve Bayes	0.88	0.63	0.72	0.46
Bayesian QDA	0.82	0.50	0.62	0.28
AdaBoost	0.59	0.77	0.70	0.35
Logistic regression	0.53	0.73	0.66	0.26
Random forest	0.41	0.77	0.64	0.15
Decision tree	0.35	0.87	0.68	0.11
XGBoost	0.18	0.90	0.64	0.24
K-Nearest neighbors	0.12	0.87	0.60	−0.06
Support vector machine	0.00	1.00	0.64	0

response. Further, among the patients with low baseline binge eating severity, anxious patients were less likely to experience placebo response.

This finding is consistent with the results observed by Blom et al. which analyzed 10 of the 12 investigator-led trials utilized in this study.[4] Blom et al.[4] developed a mixed logistic regression model that identified lower baseline BEF and longer study participation were associated with an increased likelihood of being categorized as a placebo responder. Furthermore, the developed logistic regression model with baseline BEF as the only predictor suggested that a cutoff of 5.5 binge eating episodes per week provided maximum sensitivity and specificity. This cutoff value was similar to the decision criteria for the baseline BEF derived using the GNB algorithm (Figure 2b). Approximately 14% of patients with an observed baseline BEF of >5.5% and 48% of patients with an observed BEF <5.5 experienced a clinically significant placebo response. However, it is important to note that a baseline of 5.5 binge episodes/week represents the 78th percentile of the observed data. This implies that approximately half of the patients enrolled in a clinical trial will experience a significant placebo response. It is also expected that patients with less severe eating pathology have fewer binge eating episodes to resolve as compared with patients with a higher baseline BED severity.

Additionally, the GNB algorithm based flow-chart (Figure 3) suggests that the placebo response rate is 53% for the best-case scenario (BEF ≤ 5.5 and no anxiety) The predictor interaction plot (Figure 2b) also displayed a concordant finding that the odds of predicting placebo responders is ~1 (Shapley values near 0) when BEF ≤ 5.5. Subsequently, it can be inferred that there was a stronger signal in the data to identify placebo nonresponders as compared with placebo responders. Therefore, future analysis of placebo response in BED are recommended to include a diverse set of patient-related and disease-related features collected from various sources of data (e.g., patient-reported outcomes, BED specific questionnaires, wearable devices, neuroimaging, biomarkers, etc.) to adequately

characterize and enrich trial populations. Further, as the sample size is relatively small, all trials might not be adequately represented in the testing data. More specifically, there were 17%–35% of subjects ($N = 5$–8) per trial in the testing dataset that were filtered from 19 to 30 subjects per trial in the full dataset. Therefore, a cross-trial validation approach was tested to test for any potential bias in the final model. Two randomly chosen trials were filtered from the training data and were set aside to serve as testing data. The GNB model was fitted on the training data which contained 8 (out of 10) trials. The top three variables obtained using Shapley analysis (anxiety, BEF, and baseline clinical global impression severity) remained the same as in the original analysis (Figure S4). Likewise, the relationship among baseline disease severity, anxiety, and placebo response remained the same. Performance on the testing dataset (2 out 10 trials) was as follows: 80% sensitivity, 73% specificity, 75% accuracy, and Cohen's Kappa of 0.48. As the performance of the final model by doing cross-trial validation concurs with the original case of randomly splitting subjects into training and testing sets, it can be reasonably inferred that the results obtained from our study might be applicable to future trials and the final model is not biased due to low sample size.

Given that the presented study results are validated in future studies, the application of these findings can be highly influential to inform the design of future drug development programs. Investigators may consider modifying the inclusion/exclusion criteria of their population of interest to select patients with greater baseline BED severity or patients with mild to moderate baseline BED severity with comorbid anxiety. Study protocols should include proper diagnostic tools and clinical outcome measures to screen and prospectively monitor symptoms of comorbid conditions. However, enrolling patients with high baseline BED severity could also be difficult from a recruitment perspective and can lead to additional loss of trial efficiency due to an increased risk for dropouts. In contrast, efficacy data

obtained from only patients with low baseline BED severity may not adequately translate to benefit in patients with higher disease severity, a population with a crucial need for treatment. Several theories aimed at explaining the relationship between anxiety in BED suggest a more transactional association. Individuals who suffer with BED typically experience anxiety and distress after an episode which often leads to further binge behaviors as a coping mechanism.[39] Interestingly, treatments that target both symptoms of anxiety and BED may exhibit a higher degree of separation from placebo. Similar to other quantitative analyses comparing BEF and BDF, the proposed enrichment strategy can be applicable to studies utilizing either end point.[39]

Several limitations should be acknowledged with this study. Inclusion and exclusion criteria across studies were not identical and could have led to a different subpopulation of patients with BED. Some studies only evaluated patients with comorbid obesity, whereas others enrolled patients with just BED. Study durations varied among trials, ranging between 6 and 14 weeks. For this analysis, trials were truncated to a common week due to a plateauing effect observed after 6 weeks on the average weekly normalized BEF versus time profile. Previous analysis of placebo response for other conditions suggest that trial length is also a strong moderator for placebo response.[40,41] It is unclear whether trial length influences patient expectations prior to enrollment. Furthermore, participation in longer term studies could cause subsequent worsening of BED severity that may not be captured from short-term studies. Patients with high disease severity could experience a relapse during the trial period and falsely be categorized as a placebo nonresponder. Another limitation to this study was that placebo responder status was determined by the percent change from baseline in weekly normalized BEF at each patient's last visit. Last observation carried forward was assumed for patients that dropped out prior to week 6. In this analysis, 51 out of 189 patients (27%) did not complete at least 6 weeks of placebo therapy. Previously reported dropout analyses of patients with neuropsychiatric conditions suggest that patients with high baseline disease severity and those with limited to no change in disease symptomology were more likely to drop out before study completion.[42,43] However, other factors could have also played a role in a patient's decision to drop out that may also confound the results of this analysis. In addition, all the investigator-trials included in this analysis utilized a flexible dose titration design. When comparing study designs that include an equal number of treatment arms, flexible dose-design studies have been shown to elicit a higher placebo response as compared to fixed dose study designs.[41] It is possible that patients who are aware that their "placebo dose" is titrated based on efficacy and safety targets may have higher expectations that could result in a higher placebo response.[44–46] Finally, all investigator-led trials leveraged in this analysis was conducted by one study group. Although standard of care was not provided to patients, it is unknown whether patients participated in multiple trials or if the expertise of the research group may have influenced placebo response. Analysis of two additional large randomized, placebo controlled trials evaluating the effectiveness of lisdexamfetamine in BED demonstrated a lack of association between placebo response and less severe BED symptomology at baseline.[47] Therefore, generalization of results to multicenter and multiregional studies should be taken with caution.

Modern ML algorithms, such as the ones implemented in our analysis, can be applied for discerning more patterns that can accurately identify placebo responders. Traditional statistical algorithms, such as logistic regression, are not tractable for data with multicollinearity, high dimensional data, and assume a linear relationship between the logarithm of odds and predictors. These limitations can sometimes result in erroneous findings or inadequate model performance. In our analysis, logistic regression resulted in a 53% sensitivity, which is much lower than the 88% observed using the GNB algorithm. Given the flexibility to fit nonlinear functions, modern ML methods are recommended for drawing inferences on rich and complex data that are often collected in psychiatric clinical trials. In conclusion, we recommend that patients with baseline binge eating severity and anxiety should be taken into consideration when designing future BED trials. New data must be gathered from BED trials to validate the results of this study and to identify additional reasons for placebo response in BED.

AUTHOR CONTRIBUTIONS

R.K.G. and S.K. wrote the manuscript. S.K., S.L.M., and M.G. designed the research. R.K.G., S.K., and M.G. performed the research. R.K.G. and S.K. analyzed the data.

FUNDING INFORMATION

CONFLICT OF INTEREST

The authors decleared no competing interests for this work.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

ORCID

Rahul K. Goyal https://orcid.org/0000-0002-2648-6649
Shamir N. Kalaria https://orcid.org/0000-0001-7980-4236

REFERENCES

1. Udo T, Grilo CM. Prevalence and correlates of DSM-5–defined eating disorders in a nationally representative sample of U.S. adults. *Biol Psychiatry*. 2018;84(5):345-354. doi:10.1016/j.biopsych.2018.03.014

2. Erskine HE, Whiteford HA. Epidemiology of binge eating disorder. *Curr Opin Psychiatry*. 2018;31(6):462-470. doi:10.1097/YCO.0000000000000449

3. Appolinario JC, Nardi AE, McElroy SL. Investigational drugs for the treatment of binge eating disorder (BED): an update. *Expert Opin Investig Drugs*. 2019;28(12):1081-1094. doi:10.1080/13543784.2019.1692813

4. Blom TJ, Mingione CJ, Guerdjikova AI, Keck PE, Welge JA, McElroy SL. Placebo response in binge eating disorder: a pooled analysis of 10 clinical trials from one research group: placebo response in binge eating disorder. *Eur Eat Disord Rev*. 2014;22(2):140-146. doi:10.1002/erv.2277

5. Jacobs-Pilipski MJ, Wilfley DE, Crow SJ, et al. Placebo response in binge eating disorder. *Int J Eat Disord*. 2007;40(3):204-211. doi:10.1002/eat.20287

6. Carter WP, Hudson JI, Lalonde JK, Pindyck L, McElroy SL, Pope HG Jr. Pharmacologic treatment of binge eating disorder. *Int J Eat Disord*. 2003;34(S1):S74-S88. doi:10.1002/eat.10207

7. Charney DS. National depressive and manic-depressive association consensus statement on the use of placebo in clinical trials of mood disorders. *Arch Gen Psychiatry*. 2002;59(3):262. doi:10.1001/archpsyc.59.3.262

8. Walsh BT, Seidman SN, Sysko R, Gould M. Placebo response in studies of major depression: variable, substantial, and growing. *JAMA*. 2002;287(14):1840-1847. doi:10.1001/jama.287.14.1840

9. Schatzberg AF, Kraemer HC. Use of placebo control groups in evaluating efficacy of treatment of unipolar major depression. *Biol Psychiatry*. 2000;47(8):736-744. doi:10.1016/S0006-3223(00)00846-5

10. Fairburn CG. The natural course of bulimia nervosa and binge eating disorder in young women. *Arch Gen Psychiatry*. 2000;57(7):659-665. doi:10.1001/archpsyc.57.7.659

11. Deo RC. Machine Learning in Medicine. *Circulation*. 2015;132(20):1920-1930. doi:10.1161/CIRCULATIONAHA.115.001593

12. Rutledge RB, Chekroud AM, Huys QJ. Machine learning and big data in psychiatry: toward clinical applications. *Curr Opin Neurobiol*. 2019;55:152-159. doi:10.1016/j.conb.2019.02.006

13. Perna G, Alciati A, Daccò S, Grassi M, Caldirola D. Personalized psychiatry and depression: the role of sociodemographic and clinical variables. *Psychiatry Investig*. 2020;17(3):193-206. doi:10.30773/pi.2019.0289

14. Bzdok D, Meyer-Lindenberg A. Machine learning for precision psychiatry: opportunities and challenges. *Biol Psychiatry Cogn Neurosci Neuroimaging*. 2018;3(3):223-230. doi:10.1016/j.bpsc.2017.11.007

15. Zhdanov A, Atluri S, Wong W, et al. Use of machine learning for predicting escitalopram treatment outcome from electroencephalography recordings in adult patients with depression. *JAMA Netw Open*. 2020;3(1):e1918377. doi:10.1001/jamanetworkopen.2019.18377

16. Webb CA, Trivedi MH, Cohen ZD, et al. Personalized prediction of antidepressant v. placebo response: evidence from the EMBARC study. *Psychol Med*. 2019;49(7):1118-1127. doi:10.1017/S0033291718001708

17. Chekroud AM, Zotti RJ, Shehzad Z, et al. Cross-trial prediction of treatment outcome in depression: a machine learning approach. *Lancet Psychiatry*. 2016;3(3):243-250. doi:10.1016/S2215-0366(15)00471-X

18. Trivedi MH, Rush J. Does a placebo run-in or a placebo treatment cell affect the efficacy of antidepressant medications? *Neuropsychopharmacology*. 1994;11(1):33-43. doi:10.1038/npp.1994.63

19. Rosenkranz GK. Remarks on designs enriching for placebo non-responders. *Clin Trials*. 2016;13(3):338-343. doi:10.1177/1740774515625186

20. Pablos-Méndez A, Barr RG, Shea S. Run-in periods in randomized trials: implications for the application of results in clinical practice. *JAMA*. 1998;279(3):222. doi:10.1001/jama.279.3.222

21. McElroy SL, Guerdjikova AI, Winstanley EL, et al. Acamprosate in the treatment of binge eating disorder: a placebo-controlled trial. *Int J Eat Disord*. 2011;44(1):81-90. doi:10.1002/eat.20876

22. McElroy SL, Guerdjikova AI, Mori N, et al. Armodafinil in binge eating disorder: a randomized, placebo-controlled trial. *Int Clin Psychopharmacol*. 2015;30(4):209-215. doi:10.1097/YIC.0000000000000079

23. McElroy SL, Guerdjikova A, Kotwal R, et al. Atomoxetine in the treatment of binge-eating disorder: a randomized placebo-controlled trial. *J Clin Psychiatry*. 2007;68(3):390-398. doi:10.4088/JCP.v68n0306

24. McElroy SL, Hudson JI, Malhotra S, Welge JA, Nelson EB, Keck PE. Citalopram in the treatment of binge-eating disorder: a placebo-controlled trial. *J Clin Psychiatry*. 2003;64(7):807-813. doi:10.4088/JCP.v64n0711

25. Guerdjikova AI, McElroy SL, Winstanley EL, et al. Duloxetine in the treatment of binge eating disorder with depressive disorders: a placebo-controlled trial. *Int J Eat Disord*. 2012;45(2):281-289. doi:10.1002/eat.20946

26. Guerdjikova AI, McElroy SL, Kotwal R, et al. High-dose escitalopram in the treatment of binge-eating disorder with obesity: a placebo-controlled monotherapy trial. *Hum Psychopharmacol Clin Exp*. 2008;23(1):1-11. doi:10.1002/hup.899

27. Arnold LM, McElroy SL, Hudson JI, Welge JA, Bennett AJ, Keck PE. A placebo-controlled, randomized trial of fluoxetine in the treatment of binge-eating disorder. *J Clin Psychiatry*. 2002;63(11):1028-1033. doi:10.4088/JCP.v63n1113

28. Guerdjikova AI, McElroy SL, Welge JA, Nelson E, Keck PE, Hudson JI. Lamotrigine in the treatment of binge-eating disorder with obesity: a randomized, placebo-controlled monotherapy trial. *Int Clin Psychopharmacol*. 2009;24(3):150-158. doi:10.1097/YIC.0b013e328329c7b5

29. Guerdjikova AI, Mori N, Blom TJ, et al. Lisdexamfetamine dimesylate in binge eating disorder: a placebo controlled trial: lisdexamfetamine in BED. *Hum Psychopharmacol Clin Exp*. 2016;31(5):382-391. doi:10.1002/hup.2547

30. McElroy SL. Placebo-controlled trial of sertraline in the treatment of binge eating disorder. *Am J Psychiatry*. 2000;157(6):1004-1006. doi:10.1176/appi.ajp.157.6.1004

31. McElroy SL, Arnold LM, Shapira NA, et al. Topiramate in the treatment of binge eating disorder associated with obesity: a randomized, placebo-controlled trial. *Am J Psychiatry.* 2003;160(2):255-261. doi:10.1176/appi.ajp.160.2.255

32. McElroy SL, Kotwal R, Guerdjikova AI, et al. Zonisamide in the treatment of binge eating disorder with obesity: a randomized controlled trial. *J Clin Psychiatry.* 2006;67(12):1897-1906. doi:10.4088/JCP.v67n1209

33. Rackauckas C, Ma Y, Noack A, et al. Accelerated predictive healthcare analytics with pumas, a high performance pharmaceutical modeling and simulation platform [published online ahead of print November 2020]. *Pharmacol Toxicol.* 2020. doi:10.1101/2020.11.28.402297.

34. Blaom A, Kiraly F, Lienart T, Simillides Y, Arenas D, Vollmer S. MLJ: a Julia package for composable machine learning. *J Open Source Softw.* 2020;5(55):2704. doi:10.21105/joss.02704

35. Štrumbelj E, Kononenko I. Explaining prediction models and individual predictions with feature contributions. *Knowl Inf Syst.* 2014;41(3):647-665. doi:10.1007/s10115-013-0679-x

36. Mallinckrodt CH, Zhang L, Prucka WR, Millen BA. Signal detection and placebo response in schizophrenia: parallels with depression. *Psychopharmacol Bull.* 2010;43(1):53-72.

37. Rutherford BR, Pott E, Tandler JM, Wall MM, Roose SP, Lieberman JA. Placebo response in antipsychotic clinical trials: a meta-analysis. *JAMA Psychiat.* 2014;71(12):1409-1421. doi:10.1001/jamapsychiatry.2014.1319

38. Furukawa TA, Cipriani A, Atkinson LZ, et al. Placebo response rates in antidepressant trials: a systematic review of published and unpublished double-blind randomised controlled studies. *Lancet Psychiatry.* 2016;3(11):1059-1066. doi:10.1016/S2215-0366(16)30307-8

39. Rosenbaum DL, White KS. The role of anxiety in binge eating behavior: a critical examination of theory and empirical literature. *Health Psychol Res.* 2013;1(2):19. doi:10.4081/hpr.2013.714

40. Weimer K, Colloca L, Enck P. Placebo effects in psychiatry: mediators and moderators. *Lancet Psychiatry.* 2015;2(3):246-257. doi:10.1016/S2215-0366(14)00092-3

41. Agid O, Siu CO, Potkin SG, et al. Meta-regression analysis of placebo response in antipsychotic trials, 1970–2010. *Am J Psychiatry.* 2013;170(11):1335-1344. doi:10.1176/appi.ajp.2013.12030315

42. Kalaria SN, McElroy SL, Gobburu J, Gopalakrishnan M. An innovative disease-drug-trial framework to guide binge eating disorder drug development: a case study for topiramate. *Clin Transl Sci.* 2020;13(1):88-97. doi:10.1111/cts.12682

43. Zilcha-Mano S, Keefe JR, Chui H, Rubin A, Barrett MS, Barber JP. Reducing dropout in treatment for depression: translating dropout predictors into individualized treatment recommendations. *J Clin Psychiatry.* 2016;77(12):e1584-e1590. doi:10.4088/JCP.15m10081

44. Arakawa A, Kaneko M, Narukawa M. An investigation of factors contributing to higher levels of placebo response in clinical trials in neuropathic pain: a systematic review and meta-analysis. *Clin Drug Investig.* 2015;35(2):67-81. doi:10.1007/s40261-014-0259-1

45. Khan A, Khan SR, Walens G, Kolts R, Giller EL. Frequency of positive studies among fixed and flexible dose antidepressant clinical trials: an analysis of the food and drug administraton summary basis of approval reports. *Neuropsychopharmacology.* 2003;28(3):552-557. doi:10.1038/sj.npp.1300059

46. Khin NA, Chen YF, Yang Y, Yang P, Laughren TP. Exploratory analyses of efficacy data from major depressive disorder trials submitted to the US Food and Drug Administration in support of new drug applications. *J Clin Psychiatry.* 2011;72(4):464-472. doi:10.4088/JCP.10m06191

47. Blom TJ, Guerdjikova AI, McElroy SL. Placebo response and cessation in binge eating disorder: a pooled analysis of two randomized parallel-group clinical trials. *Eur Eat Disord Rev.* 2019;27(4):421-428. doi:10.1002/erv.2655

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.