Request for Copy of Published Material

Enclosure:

POSTER: Bell Lynum KS, Zhang Z, Atkins N et al. Presented at Psych Congress, September 17-19, 2025, San Diego, CA

Development of a Machine Learning Model Predicting Response to Aripiprazole Once-Monthly in Patients Diagnosed With Bipolar I Disorder

Scan the QR code to access a digital copy of the poster and supplementary content

Karimah S Bell Lynum, PharmD[1]; Zhen Zhang, PhD[1]; Norman Atkins, Jr., PhD[1]; Anne Walker, PhD[2]; Hong Xiao, PhD[1]; Caitlin A Stamatis, PhD[1]; Soma S Nag, PhD[1]; Mauricio Tohen, MD, DrPH[3]

[1]Otsuka Pharmaceutical Development & Commercialization Inc., Princeton, NJ, USA; [2]Lundbeck LLC, Deerfield, IL, USA; [3]Department of Psychiatry and Behavioral Sciences, University of New Mexico Health Sciences Center, Albuquerque, NM, USA

Background

- Aripiprazole once-monthly 400 mg (AOM 400) is a long-acting injectable antipsychotic that is approved in the United States for the maintenance monotherapy treatment of bipolar I disorder (BP-I) in adults.[1]
- In a double-blind, placebo-controlled, randomized withdrawal study, AOM 400 was effective and well-tolerated for the maintenance treatment of BP-I, delaying mood episode recurrence and supporting symptom and functional stability.[2-4] Longer-term data indicated that most patients treated with AOM 400 remained stable, with a minimal need for rescue medication.[5]
- Real-world data show that initiation of AOM 400 in patients diagnosed with BP-I is associated with reduced psychiatric healthcare use, including fewer and shorter hospitalizations, fewer emergency room visits, and a longer time to rehospitalization.[6,7]
- Insight into factors that predict response to AOM 400 in patients diagnosed with BP-I may assist clinicians in tailoring treatment to individual patients, potentially improving outcomes.
- Statistical methods such as regression analyses have traditionally been used to examine treatment–response relationships, guided by pre-specified hypotheses about which factors to test.[8,9] Newer methodologies involving machine learning may improve this process by detecting complex, data-driven patterns that are not limited to *a priori* assumptions.[8,9]

Here, we describe the development of a machine learning model to identify baseline factors predictive of response to AOM 400 and placebo in patients diagnosed with BP-I using data from a clinical trial.

A separate model to identify baseline factors predictive of response to AOM 400 in patients diagnosed with schizophrenia has also been developed using data from a clinical trial, with results reported in poster 65.

Methods

Source data

- Data used for the development of the model were derived from patients in a multi-phase study that evaluated the efficacy, safety, and tolerability of AOM 400 versus placebo over a 52-week period (**Supplementary Figure 1**; please scan the QR code to access supplementary content).[2]
- The primary endpoint was recurrence of any mood episode during the 52-week double-blind randomized phase of the trial (Phase D). Recurrence was reported in 35 of 132 patients randomized to AOM 400 and in 68 of 133 patients randomized to placebo.[2]

Variable screening and predictive modelling

- A two-step procedure was applied for each treatment arm.
- A univariate analysis was conducted to associate each variable, one at a time, to the outcome of recurrence of any mood episode. Tested variables are shown in **Box 1**, and included demographic characteristics, vital signs, medical/disease history, clinician- and patient-reported rating scale scores at Phase D entry, and changes in vital signs and rating scale scores from Phase B entry to Phase D entry. These variables reflected the majority of data collected at Phase D entry, and changes in data between Phase B entry and Phase D entry.
- Variables meeting predefined thresholds were carried forward to a multivariate predictive modeling step comprising a variety of machine learning algorithms.
  - The goal of variable screening was to narrow the dataset to only the most relevant factors, reducing 'noise' and improving the model's ability to identify meaningful predictors of treatment response.[10]
- Consistent with the work of others,[11] data were randomly split into 70% for training the ensemble model and 30% for out-of-sample validation.
- Model performance was assessed using standard metrics, with the importance of each variable in the final model reported using SHapley Additive exPlanations.

Box 1: Baseline variables screened in univariate analyses for each treatment arm

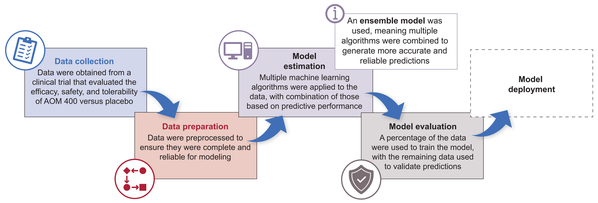
Demographics	• Age • Gender	• Race/ethnicity • Height
Vital signs	• BMI/weight/waist circumference • Heart rate	• Seated/supine diastolic/systolic BP • Temperature
Medical and disease history	• Number of lifetime manic/depressive/mixed episodes • Number of prior hospitalizations for a mood episode • Number of mood episodes in past 12 months/past 10 years • Duration of BP-I/current manic episode • Age at first diagnosis/first manic episode/first depressive episode	
Clinician-rated and patient-reported outcome measures	• YMRS • MADRS • Brief QoL.BD • FAST	• CGI-BP • AIMS • BARS • SAS

AIMS, Abnormal Involuntary Movement Scale; BARS, Barnes Akathisia Rating Scale; BMI, body mass index; BP, blood pressure; BP-I, bipolar I disorder; Brief QoL.BD, Brief Quality of Life in Bipolar Disorder; CGI-BP, Clinical Global Impression – Bipolar version; FAST, Functioning Assessment Short Test; MADRS, Montgomery-Åsberg Depression Rating Scale; SAS, Simpson-Angus Scale; YMRS, Young Mania Rating Scale.

Results

- Key steps in the predictive modeling process are shown in **Figure 1**.
- Overall, 212 baseline variables were considered (**Supplementary Table 1**; please scan the QR code to access supplementary content); of these, 34 and 38 were carried forward to the predictive modeling steps for AOM 400 and placebo, respectively.
- Two classifiers were included in the final ensemble model for AOM 400, while three were included in the final model for placebo (**Table 1**).
- The final models for AOM 400 and placebo demonstrated strong performances (**Table 2** and **Figure 2**).
- The importance ranking of the top ten variables included in the final predictive models for AOM 400 and placebo are shown in **Figure 3** and **Figure 4**, respectively.
- The importance ranking of all 34 variables included in the final predictive model for AOM 400 and of all 38 variables included in the final predictive model for placebo are shown in **Supplementary Figure 2** and **Supplementary Figure 3**, respectively (please scan the QR code to access supplementary content).

Figure 1: Key steps in the predictive modeling process



Predictive modeling/analysis was stratified by arm, to identify treatment-specific influential variables

AOM 400, aripiprazole once-monthly 400 mg

Table 1: Final ensemble models for AOM 400 and placebo

Treatment	Final ensemble model	Weight
AOM 400	eXtreme Gradient Boosting	0.502
	Elastic Net (alpha=0.1)	0.498
Placebo	Random forests	0.808
	eXtreme Gradient Boosting	0.145
	Elastic Net (alpha=0.1)	0.047

- In total, nine machine learning algorithms were evaluated for inclusion in each model; of these, **two were included in the ensemble model for AOM 400 and three were included in the ensemble model for placebo.**
- Each algorithm was assigned a weight reflecting how much it contributed to improving the overall accuracy of the combined model **(a higher weight = a greater contribution to prediction accuracy).**

All other classifiers tested had a weight of zero (indicating a trivial contribution to the ensemble)
AOM 400, aripiprazole once-monthly 400 mg

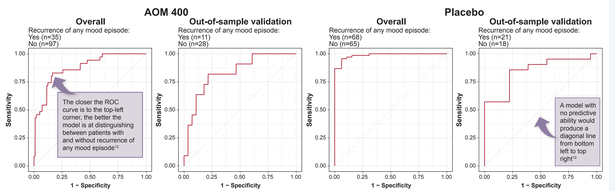
Table 2: Model performance metrics for AOM 400 and placebo

Treatment	Final ensemble classifier	Accuracy	Sensitivity	Specificity	F1-score	AUC (ROC)	AUC (PRC)
AOM 400	Overall, N=132 (35 events)	0.83	0.83	0.84	0.73	0.88	0.74
	Out-of-sample validation, random 30% split, n=39 (11 events)	0.8	0.82	0.79	0.69	0.83	0.65
Placebo	Overall, N=133 (68 events)	0.95	0.96	0.94	0.95	0.99	0.99
	Out-of-sample validation, random 30% split, n=39 (21 events)	0.82	0.86	0.78	0.84	0.85	0.89

The models demonstrated a **strong performance** for predicting recurrence of any mood episodes.

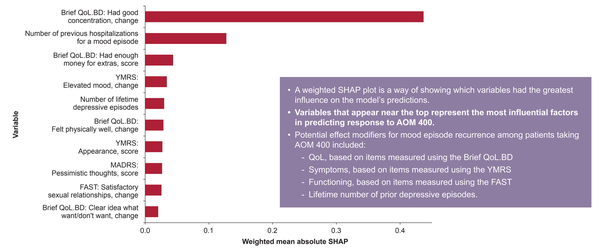
AOM 400, aripiprazole once-monthly 400 mg; AUC, area under the curve; PRC, precision-recall curve; ROC, receiver-operating characteristic

Figure 2: Model performance plots for AOM 400 and placebo



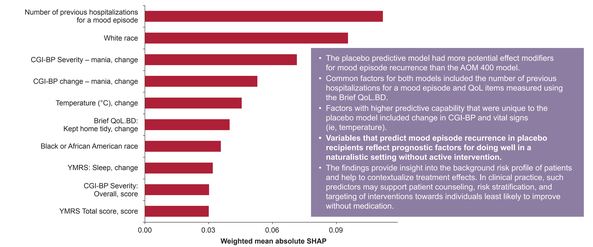
AOM 400, aripiprazole once-monthly 400 mg; ROC, receiver-operating characteristic

Figure 3: Variable importance ranking for AOM 400 (top 10 variables)



AOM 400, aripiprazole once-monthly 400 mg; Brief QoL.BD, Brief Quality of Life in Bipolar Disorder; FAST, Functioning Assessment Short Test; MADRS, Montgomery-Åsberg Depression Rating Scale; QoL, quality of life; SHAP, SHapley Additive exPlanations; YMRS, Young Mania Rating Scale

Figure 4: Variable importance ranking for placebo (top 10 variables)



AOM 400, aripiprazole once-monthly 400 mg; Brief QoL.BD, Brief Quality of Life in Bipolar Disorder; CGI-BP, Clinical Global Impression – Bipolar version; QoL, quality of life; SHAP, SHapley Additive exPlanations; YMRS, Young Mania Rating Scale

Limitations

- The input data were derived from a trial that was not designed with predictive modeling in mind. As such, variable selection was limited to what was collected during the trial rather than being guided by a theory-driven framework. This may impact model predictive power and real-world applicability.

Conclusions

- Robust machine learning models have been developed using clinical trial data to identify baseline factors predictive of response to AOM 400 or placebo in patients diagnosed with BP-I.
- Initial findings support the potential utility of disease chronicity for predicting response to AOM 400, notably the number of previous hospitalizations for a mood episode and the number of lifetime depressive episodes.
- The identification of patient-reported quality of life and functioning as potential effect modifiers in AOM 400 recipients highlights the importance of the subjective patient experience, indicating that environmental context might be more than just a secondary concern.
- Results from the AOM 400 model may inform clinician-led monitoring and supportive interventions (eg, psychoeducation, therapy, medication adjustment) in patients diagnosed with BP-I who are being considered for AOM 400 treatment, with a view to optimizing treatment outcomes.
- Further validation of the model using real-world data is planned.

References

1. Otsuka America Pharmaceutical, Inc. Abilify Maintena (aripiprazole). Prescribing information. March 2025.
2. Calabrese et al. J Clin Psychiatry 2017; 78 (3): 324–331.
3. Calabrese et al. J Affect Disord 2018; 227: 649–656.
4. Calabrese et al. J Affect Disord 2018; 241: 425–432.
5. Calabrese et al. Int J Bipolar Disord 2018; 6 (1): 14.
6. Waters et al. Curr Med Res Opin 2023; 39 (2): 299–306.
7. Goto et al. Neuropsychopharmacol Rep 2023; 43 (3): 425–433.
8. Chekroud et al. World Psychiatry 2021; 20 (2): 154–170.
9. Goldstein et al. Eur Heart J 2017; 38 (23): 1805–1814.
10. Michel et al. BMC Med Inform Decis Mak 2021; 21 (Suppl 4): 130.
11. Wu et al. JAMA Netw Open 2020; 3 (2): e1921660.
12. Jiao & Du. Quant Biol 2016; 4: 320–330.

Disclosures

KSBL, ZZ, HX, CAS, SN: employees of Otsuka Pharmaceutical Development & Commercialization Inc.
NA: consultant for Otsuka Pharmaceutical Development & Commercialization Inc.
AW: employee of Lundbeck LLC.
MT: received honoraria or consultation fees from Abbott, AbbVie, Alkermes, AstraZeneca, Elan, Gedeon Richter, Intracellular Therapies, Johnson & Johnson, Lilly, Lundbeck, Merck, Minerva, Neurocrine Biosciences, Otsuka, Pfizer, Roche, Sunovion, and Teva. MT was an employee at Lilly (1997–2008); his spouse was an employee at Lilly (1998–2013).

Acknowledgments

Medical writing and editorial assistance were provided by Lyndal Staples, BSc, and colleagues of Cambridge (a division of Prime, Cambridge, UK), with funding from Otsuka Pharmaceutical Development & Commercialization Inc. (Princeton, NJ, USA) and H. Lundbeck A/S (Valby, Denmark).
The sponsors thank the patients and their families who participated in this study.

Sponsorship

This work was supported by Otsuka Pharmaceutical Development & Commercialization Inc. (Princeton, NJ, USA) and H. Lundbeck A/S (Valby, Denmark).