OPEN ACCESS

ORIGINAL RESEARCH

# Comparison of prediction methods for treatment continuation of antipsychotics in children and adolescents with schizophrenia

Soo Min Jeon ,[1] Jaehyeong Cho,[2] Dong Yun Lee ,[3] Jin-Won Kwon [1]

► Additional supplemental material is published online only. To view, please visit the journal online (http://dx.doi.org/10.1136/ebmental-2021-300404).

[1]BK21 FOUR Community–Based Intelligent Novel Drug Discovery Education Unit, College of Pharmacy and Research Institute of Pharmaceutical Sciences, Kyungpook National University, Daegu, Republic of Korea
[2]Department of Biomedical Systems Informatics, Yonsei University College of Medicine, Seodaemun-gu, Seoul, Republic of Korea
[3]Department of Biomedical informatics, Ajou University School of Medicine, Ajou University, Suwon, Gyeonggi-do, Republic of Korea

**Correspondence to**
Professor Jin-Won Kwon, BK21 FOUR Community–Based Intelligent Novel Drug Discovery Education Unit, College of Pharmacy and Research Institute of Pharmaceutical Sciences, Kyungpook National University, Daegu 41566, Korea (the Republic of); jwkwon@knu.ac.kr

Received 13 December 2021
Accepted 25 March 2022
Published Online First
13 April 2022

Check for updates



**To cite:** Jeon SM, Cho J, Lee DY, *et al*. *Evid Based Ment Health* 2022;**25**:e26–e33.

## ABSTRACT

**Objective** There is little evidence for finding optimal antipsychotic treatment for schizophrenia, especially in paediatrics. To evaluate the performance and clinical benefit of several prediction methods for 1-year treatment continuation of antipsychotics.

**Design and Settings** Population-based prognostic study conducting using the nationwide claims database in Korea.

**Participants** 5109 patients aged 2–18 years who initiated antipsychotic treatment with risperidone/aripiprazole for schizophrenia between 2010 and 2017 were identified.

**Main outcome measures** We used the conventional logistic regression (LR) and common six machine-learning methods (least absolute shrinkage and selection operator, ridge, elstic net, randomforest, gradient boosting machine, and superlearner) to derive predictive models for treatment continuation of antipsychotics. The performance of models was assessed using the Brier score (BS), area under the receiver operating characteristic curve (AUROC) and area under the precision-recall curve (AUPRC). The clinical benefit of applying these models was also evaluated by comparing the treatment continuation rate between patients who received the recommended medication by models and patients who did not.

**Results** The gradient boosting machine showed the best performance in predicting treatment continuation for risperidone (BS, 0.121; AUROC, 0.686; AUPRC, 0.269). Among aripiprazole models, GBM for BS (0.114), SuperLearner for AUROC (0.688) and random forest for AUPRC (0.317) showed the best performance. Although LR showed lower performance than machine learnings, the difference was negligible. Patients who received recommended medication by these models showed a 1.2–1.5 times higher treatment continuation rate than those who did not.

**Conclusions** All prediction models showed similar performance in predicting the treatment continuation of antipsychotics. Application of prediction models might be helpful for evidence-based decision-making in antipsychotic treatment.

### Summary box

**What is already known about this subject?**
⇒ Antipsychotics are a mainstay of early-onset schizophrenia treatment. However, there is no clear evidence on how to determine the most appropriate antipsychotic medication for each patient, especially for paediatric patients.

**What are the new findings?**
⇒ Most machine learning algorithms did not outperform the conventional logistic regression in predicting the treatment continuation of antipsychotics in children and adolescents with schizophrenia.
⇒ Regardless of the modelling algorithms, the recommendation of antipsychotics for each patient by prediction models had a clinical benefit in terms of treatment continuation.

**How might it impact clinical practice in the foreseeable future?**
⇒ The use of prediction models might help identify appropriate antipsychotic treatments in children and adolescents with schizophrenia.

## BACKGROUND

Schizophrenia is a challenging mental health disorder that is characterised by a wide range of psychological symptoms such as hallucinations, delusions and extremely disordered thinking and behaviour that impair daily functioning.[1] Early-onset schizophrenia (EOS) is generally defined as schizophrenia onset before the age of 18 years. In childhood, the incidence of schizophrenia is extremely rare, but the incidence increases steeply in adolescence.[2] Compared with adulthood-onset schizophrenia, EOS has more severe negative symptoms, premorbid adjustment, impaired cognitive function and worse treatment efficacy.[3 4] As schizophrenia is a lifelong mental health disorders, the symptom severity of EOS could deteriorate over time and persist into adulthood,[3 4] which causes significant and long-lasting health, social and financial burdens. Therefore, the clinical management of EOS deserves separate attention.

Antipsychotics are a keystone of pharmacological treatment for schizophrenia. To date, most clinical guidelines recommended atypical antipsychotics as the first-line pharmacological treatment for patients with schizophrenia at any age.[5] However, due to the different pharmacological mechanisms of each antipsychotic and the clinical heterogeneity of schizophrenia (disease severity, symptoms and comorbidities),[4 6] it is difficult to make a decision

about which medication could maximise the treatment outcome for each individual patient. Consequently, it is necessary to attempt multiple antipsychotics as a trial-and-error approach for most patients before identifying an effective antipsychotic,[7] which causes delayed treatment of early phase and exposes patients to unwarranted risk of side effects. Given the challenges in antipsychotic treatment for schizophrenia, a prediction model for identifying the treatment response before initiating treatment would be a beneficial approach to increase treatment effectiveness and decrease the risk of side effects.

Although some previous studies have developed prediction models to determine the treatment response of antipsychotics,[8–12] there were some concerns regarding those studies. First, most only applied the conventional regression method without considering other approaches, such as machine learning methods, which have the potential to improve prediction.[8 9] Second, few studies included children or adolescents. The difference between adult-onset schizophrenia and EOS has been reported in several previous studies, and the treatment response of children and adolescents in the gradual process of development in the dopamine and neurotransmitter system may differ from that of adults with a stable system.[13] Third, most studies used the complicated predictors that clinicians cannot easily access, such as genetics type[10] and imaging data.[11] Furthermore, to our best knowledge, only one study evaluated the clinical benefit (improvement of treatment continuation for 1 year) of this kind of prediction model in selecting antipsychotic treatment for schizophrenia.[12]

## OBJECTIVE

This study was conducted to compare several prediction modelling algorithms for predicting the treatment continuation of antipsychotics using the nationwide population-based claims data of children and adolescents with schizophrenia. The focus of this study was on risperidone and aripiprazole, the two most frequently prescribed medications in this population.[14] In addition, we preliminary evaluated whether these prediction models could improve the treatment continuation of antipsychotics when applied to personalised treatment recommendation.

## METHODS

### Data source

We used the database provided by the Health Insurance Review and Assessment (HIRA) Service, a government-affiliated agency that reviews the accuracy of national health insurance claims, which are mandatory for all Korean populations. The HIRA database includes variables that identify individual characteristics, such as age, sex and economic vulnerability (ie, approximately 2.8% of the total population who qualify for medical insurance advantage beneficiaries). Clinical data include information billed by healthcare providers such as diagnoses, prescriptions, procedures and devices. The diagnosis was identified by the international classification of disease, 10th revision (ICD-10) codes. There were no missing values in this database.

This study was approved by the Institutional Review Board of Kyungpook National University (IRB number: KNU 2018–0141), and the requirement for informed consent was waived. The Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis guideline was used to develop and validate the prediction model in this study.[15]

### Study participants

First, data of all patients who were prescribed antipsychotics and diagnosed with schizophrenia (ICD-10: F20–F29) between 2008 and 2017 were collected. Each patient was required to have a look-back period of at least 2 years to identify the new users of antipsychotics for the first episode of schizophrenia. Patients who had any history of schizophrenia diagnosis or antipsychotic prescription in the first 2 years (2008/2009) were excluded because their look-back period was insufficient. The first date of antipsychotic prescription was defined as the index date for each patient. Moreover, patients who were not included in the age range of 2–18 years at the index date were excluded. To reduce the heterogeneity of disease severity within populations,[16] patients who were under psychiatric hospitalisation at the index date or prescribed multiple/long-acting injection-type antipsychotics at the index date were excluded. Finally, patients who used risperidone or aripiprazole as the first-line antipsychotic treatment were selected. Most patients (approximately 75%) in our data set used one of these two medications, and these two medications are the only Korea Food & Drug Administration-approved atypical antipsychotics for treating schizophrenia in children and adolescents[17] (online supplemental figure S1).

### Outcome

Treatment continuation is an important proxy measure of antipsychotic effectiveness, reflecting drug efficacy, safety and tolerability for both patients and clinicians.[18] Hence, the treatment continuation was measured as the outcome of interest in this study. The treatment continuation was defined as the continuous treatment of their first-line antipsychotic medications for at least 1 year without stopping over 60 days of missed prescription or any event of all-cause treatment discontinuation. All-cause treatment discontinuation included the occurrence of any of the following events during the follow-up period: suicide (X60–X84, Y87), psychiatry hospitalisation, emergency department visit with any mental disorder and switch to or addition of other antipsychotics. Patients who met all these criteria of treatment continuation were assigned as 1; all others were assigned as 0.

### Predictors

Approximately 60 predictors, including patient's demographic factors (sex, age and type of insurance) and clinical characteristics (inpatient or outpatient visit, diagnosis and prescription records), were measured for 1 year before the index date. For risperidone and aripiprazole, respectively, each variable was constructed as a binary or continuous predictor (online supplemental tables S1 and S2). Continuous predictors were normalised using min–max scaling, which is a simple method to rescale the values of the predictor range of 0 to 1. Data normalisation is the transformation of predictors to a fixed range, which is essential for preparing data for prediction modelling.[19] The predictors observed less than 10 times for either medication were not included in the model development.[20] There were no missing values for any predictors because we used claims data.

### Development and validation of prediction models

We developed two separate prediction models for the 1-year treatment continuation of patients with risperidone and of those with aripiprazole. The respective clinical benefit of the recommended antipsychotic treatment agent (ie, risperidone or aripiprazole) was investigated based on these individual prediction models. The following seven different prediction algorithms were investigated in this process: (1) multivariable logistic regression (LR),

**Table 1** Baseline characteristics of patients treated with risperidone and aripiprazole

| | Risperidone (n=5109) | Aripiprazole (n=3393) |
|---|---|---|
| Treatment continuation for 1 year (n, %) | 724 (14.2) | 471 (13.9) |
| Demographics* (n, %) | | |
| Sex | | |
| Male | 3195 (62.5) | 1944 (57.3) |
| Female | 1914 (37.5) | 1449 (42.7) |
| Age (years, mean±SD) | 13.8±3.4 | 14.1±3.4 |
| Insurance type | | |
| Health insurance | 4343 (85.0) | 3182 (93.8) |
| Medical aid | 766 (15.0) | 211 (6.2) |
| Comorbidities of mental health disorders† (n, %) | | |
| First diagnosis of schizophrenia at index date | 1969 (38.5) | 1648 (48.6) |
| Anxiety disorder | 776 (15.2) | 741 (21.8) |
| Depression | 1257 (24.6) | 1184 (34.9) |
| ADHD | 1161 (22.7) | 694 (20.5) |
| Somatoform disorder | 105 (2.1) | 91 (2.7) |
| Reaction to severe stress and adjustment disorder | 438 (8.6) | 325 (9.6) |
| Other neurotic disorders | 56 (1.1) | 50 (1.5) |
| Mental retardation | 532 (10.4) | 184 (5.4) |
| Other behaviour and emotional disorders with onset generally occurring in childhood and adolescence | 124 (2.4) | 66 (1.9) |
| Emotional disorders with onset specific to childhood | 285 (5.6) | 135 (4.0) |
| Nonorganic disorder | 127 (2.5) | 114 (3.4) |
| Tic disorder | 192 (3.8) | 220 (6.5) |
| Bipolar disorder | 120 (2.3) | 117 (3.4) |
| Autism spectrum disorder | 277 (5.4) | 145 (4.3) |
| Other mental health disorders | 1150 (22.5) | 819 (24.1) |
| Prescriptions of other psychotropics† (n, %) | | |
| Anticholinergic drugs | 34 (0.7) | 18 (0.5) |
| Antianxiety drugs | 1489 (29.1) | 1097 (32.3) |
| SSRI/SNRI | 1369 (26.8) | 1169 (34.5) |
| TCA | 107 (2.1) | 76 (2.2) |
| MAO inhibitor | 196 (3.8) | 187 (5.5) |
| Methylphenidate | 892 (17.5) | 445 (13.1) |
| Nonstimulants for ADHD | 178 (3.5) | 178 (5.2) |
| Antiepileptic drugs | 612 (12.0) | 358 (10.6) |
| Lithium | 58 (1.1) | 41 (1.2) |
| Comorbidities of paediatric chronic conditions†‡ (n, %) | | |
| Neurologic and neuromuscular | 128 (2.5) | 57 (1.7) |
| Cardiovascular | 71 (1.4) | 50 (1.5) |
| Respiratory | 1 (0.0) | 5 (0.1) |
| Renal and urologic | 59 (1.2) | 31 (0.9) |
| Gastrointestinal | 81 (1.6) | 63 (1.9) |
| Haematologic or immunologic | 12 (0.2) | 13 (0.4) |
| Metabolic | 220 (4.3) | 160 (4.7) |
| Other congenital or genetic defects | 122 (2.4) | 65 (1.9) |
| Malignancy | 26 (0.5) | 7 (0.2) |
| Neonatal | 3 (0.1) | 1 (0.0) |
| Asthma | 798 (15.6) | 550 (16.2) |
| Diabetes mellitus | 54 (1.1) | 30 (0.9) |
| Seizure | 397 (7.8) | 242 (7.1) |
| Prescriptions of other general medications† (n, %) | | |
| Antipyretics, analgesics | 2960 (57.9) | 2113 (62.3) |
| Antihistamines | 3423 (67.0) | 2341 (69.0) |
| Other antiallergic agents | 7 (0.1) | 11 (0.3) |
| Cardiovascular system drugs | 363 (7.1) | 311 (9.2) |
| Respiratory system drug | 3629 (71.0) | 2420 (71.3) |
| Drugs for diabetes mellitus | 7 (0.1) | 7 (0.2) |

Continued

**Table 1** Continued

| | Risperidone (n=5109) | Aripiprazole (n=3393) |
|---|---|---|
| Antibiotic preparations | 3600 (70.5) | 2460 (72.5) |
| Narcotic analgesics | 67 (1.3) | 52 (1.5) |
| Utilisation of healthcare system† (n, %) | | |
| Number of any route visits for mental health disorders | 1.70±3.4 | 1.95±3.70 |
| Number of inpatient visits for mental health disorders | 0.07±2.23 | 0.13±2.89 |
| Number of outpatient visits for mental health disorders | 1.62±2.16 | 1.81±2.21 |
| Number of emergency room visits for mental health disorders | 0.01±0.09 | 0.01±0.11 |
| Number of any route visits for non-mental health disorders | 12.08±13.41 | 12.43±12.67 |
| Number of inpatient visits for non-mental health disorders | 0.19±0.97 | 0.17±0.62 |
| Number of outpatient visits for non-mental health disorders | 11.92±13.21 | 12.28±12.53 |
| Number of emergency room visits for non-mental health disorders | 0.04±0.29 | 0.04±0.25 |

*Assessed on the cohort entry.
†Assessed for 1 year before the index date.
‡The paediatric complex chronic conditions were classified according to the modified criteria of Feudtner *et al*.
ADHD, attention deficit hyperactivity disorder; MAO, monoamine oxidase inhibitor; SNRI, selective norepinephrine reuptake inhibitor; SSRI, selective serotonin reuptake inhibitor.

(2) least absolute shrinkage and selection operator (LASSO), (3) ridge, (4) elastic net, (5) random forest, (6) gradient boosting machine (GBM) and (7) SuperLearner. Although LR could be categorised into machine learning algorithms, it was considered as a conventional prediction model in this study because it does not require a complex process of tuning hyperparameters compared with other algorithms.

The data set for each medication was randomly divided into training (75%) and test (25%) data sets. The training data set was used to tune the final prediction model, and the remaining test data set was only used to evaluate the performance of the finalised model. The multivariable LR was fitted using the backward stepwise algorithm based on the Akaike information criterion. In machine learning methods, the grid search method was used to identify the regularisation parameter of LASSO/ridge/elastic net and some hyperparameters of random forest/GBM. In this process, we selected the final hyperparameters that maximised the area under the receiver operating characteristic curve (AUROC) in the fivefold CV of the training data set. SuperLearner, an ensemble method, was generated by combining six tuned models (backward stepwise LR, LASSO, ridge, elastic net, random forest, and GBM) into a single prediction model; the optimal weight for combining these models was determined using the fivefold CV process. The detailed methods used to develop each model are described in online supplemental method S1.

### Statistical analysis

#### Demographics of study population

The data of patients using risperidone and aripiprazole (demographics, medical history and medications) were presented as number (%) for categorical predictors and mean (SD) for continuous predictors. The differences between the training set and test data set were compared using independent two-sample t tests (for continuous variables) and $\chi^2$ tests (for categorical variables).

#### Performance evaluation of candidate algorithms

The performance of the finalised prediction models was evaluated in the test data set of each medication. To evaluate the performance of predicting algorithms using these two approaches, the following eight metrics were used: Brier score (BS), AUROC area under the precision-recall curve (AUPRC), sensitivity, specificity, accuracy, f1-score and the Matthews correlation coefficient (MCC) (online supplemental figure S2). Among these metrics, BS, AUROC and AUPRC were selected as the main metrics to evaluate the algorithms' performance. Moreover, the importance of predictors was calculated as the relative value to the highest value (1.00) of each model based on their beta coefficient (backward stepwise LR, LASSO, ridge and elastic net) or the decrease in accuracy/Gini score (random forest and GBM).

The finalised prediction models (for each medication) were used to estimate their predicted treatment continuation for the specific medication, and then the medication with a relatively high probability of treatment continuation was labelled as the optimal medication. If the patient received the identical medication with the labelled optimal medication, they were classified as 'recommended,' and others were classified as 'non-recommended'. The Kaplan-Meier survival analysis and log-rank test were used to analyse the difference in treatment continuation in the recommended group versus the non-recommended group (online supplemental figure 3).

All p values $<0.05$ were considered to be statistically significant, and all analyses were conducted using SAS Enterprise Guide V.6.1 (SAS Institute, Cary, North Carolina) and R software V.3.6.1 (R foundation, Vienna, Austria).

## FINDINGS

Table 1 shows the characteristics of study participants and the proportion of 1-year treatment continuations for risperidone and aripiprazole. There were 5109 patients in the risperidone cohort and 3393 patients in the aripiprazole cohort, and approximately 14.0% of them were continuously prescribed their initial medication for 1 year. The mean age of patients in the risperidone and aripiprazole cohorts was 13.8 (SD 3.4) and 14.1 (SD 3.4) years, respectively. Male patients were more prevalent (risperidone 57.3%; aripiprazole 62.5 %), and most patients were covered by the national health insurance (risperidone 85.0%; aripiprazole 93.8%). For 1 year before the index date, selective serotonin reuptake inhibitors (risperidone 26.8%; aripiprazole 34.5%) and antianxiety medications (risperidone 29.1%; aripiprazole 32.3%) were the top two psychotropic medications prescribed to patients. Depression was the most frequent diagnosis among the comorbidities of mental health disorder (risperidone 24.6%; aripiprazole 34.9%). The prescription of other medications and the utilisation of the healthcare system showed similar distribution in both cohorts. Overall, the distribution of characteristics

**Table 2** Performance of machine learning models in predicting treatment response

| | Risperidone | | | Aripiprazole | | |
|---|---|---|---|---|---|---|
| Model | BS | AUROC | AUPRC | BS | AUROC | AUPRC |
| Backward stepwise LR | 0.124 | 0.659 | 0.244 | 0.117 | 0.672 | 0.300 |
| LASSO | 0.122 | 0.678 | 0.253 | 0.115 | 0.681 | 0.304 |
| Ridge | 0.123 | 0.680 | 0.255 | 0.117 | 0.683 | 0.307 |
| e-net | 0.122 | 0.676 | 0.252 | 0.115 | 0.683 | 0.305 |
| Random forest | 0.124 | 0.658 | 0.250 | 0.115 | 0.678 | 0.317 |
| GBM | 0.121 | 0.686 | 0.269 | 0.114 | 0.682 | 0.314 |
| SuperLearner | 0.121 | 0.685 | 0.264 | 0.115 | 0.688 | 0.314 |

AUPRC, area under the precision-recall curve; AUROC, area under the receiver operating characteristic curve; BS, Brier score; e-net, elastic net; GBM, gradient boosting machine; LASSO, least absolute shrinkage and selection operator; LR, logistic regression.

between the training data set and test data set showed no statistical difference (online supplemental tables S3 and S4).

Table 2 shows the performance of the machine learning and backward stepwise LR models for predicting the treatment continuation of both risperidone and aripiprazole in the test data set. Machine learning algorithms demonstrated better performance than the conventional stepwise LR in predicting the treatment continuation of both risperidone and aripiprazole; however, the performance of other machine learning algorithms also showed a close value to these metrics (online supplemental figure S4 and table S5). GBM was the best modelling algorithm in predicting the treatment continuation of risperidone, with the lowest BS value (0.121) and the highest AUROC (0.686) and AUPRC (0.269) values, closely followed by SuperLearner (BS 0.121; AUROC 0.685; AUPRC 0.262). In the prediction of aripiprazole, GBM, SuperLearner and random forest exhibited the best performance in terms of BS (GBM 0.114), AUROC (SuperLearner 0.688) and AUPRC (random forest 0.317). Regarding sensitivity, specificity, precision, accuracy, f1-score and MCC, all models demonstrated comparable performance to each other (online supplemental figure S5). Furthermore, although the conventional stepwise LR demonstrated worse performance than the machine learning algorithms, the difference in performance was small. For instance, the range of absolute difference in AUROC was just −0.027 to 0.001 and −0.016 to 0.006 for risperidone and aripiprazole, respectively.

Figure 1 shows the clinical benefit of using the prediction models for improving treatment continuation by supporting treatment selection. Patients taking the recommended medication had a favourable outcome in treatment continuation compared with that of patients taking the non-recommended medication. In the overall cohort of the test data set, the proportion of patients with a treatment continuation in the recommended group was approximately 1.2 to 1.5-fold greater than that in the non-recommended group. Furthermore, during the follow-up period, patients in the recommended medication group had a significantly higher survival probability of treatment continuation than those in the non-recommended medication group. This tendency was consistently observed in the subgroup analysis stratified for the observed medications, although the statistical significance was not obtained in the aripiprazole group (online supplemental figure S6).

Table 3 shows the top 10 important predictors in the prediction models (all predictors for each model were presented in the online supplemental tables S6 and S7). Age and utilisation of healthcare system were important predictors in prediction models of both risperidone and aripiprazole. In the prediction models of risperidone, four models consistently considered anticholinergic use and history of attention deficit hyperactivity disorder as important predictors. Lithium use and bipolar disorder diagnosis were considered as important predictors in four of the six prediction models of aripiprazole.

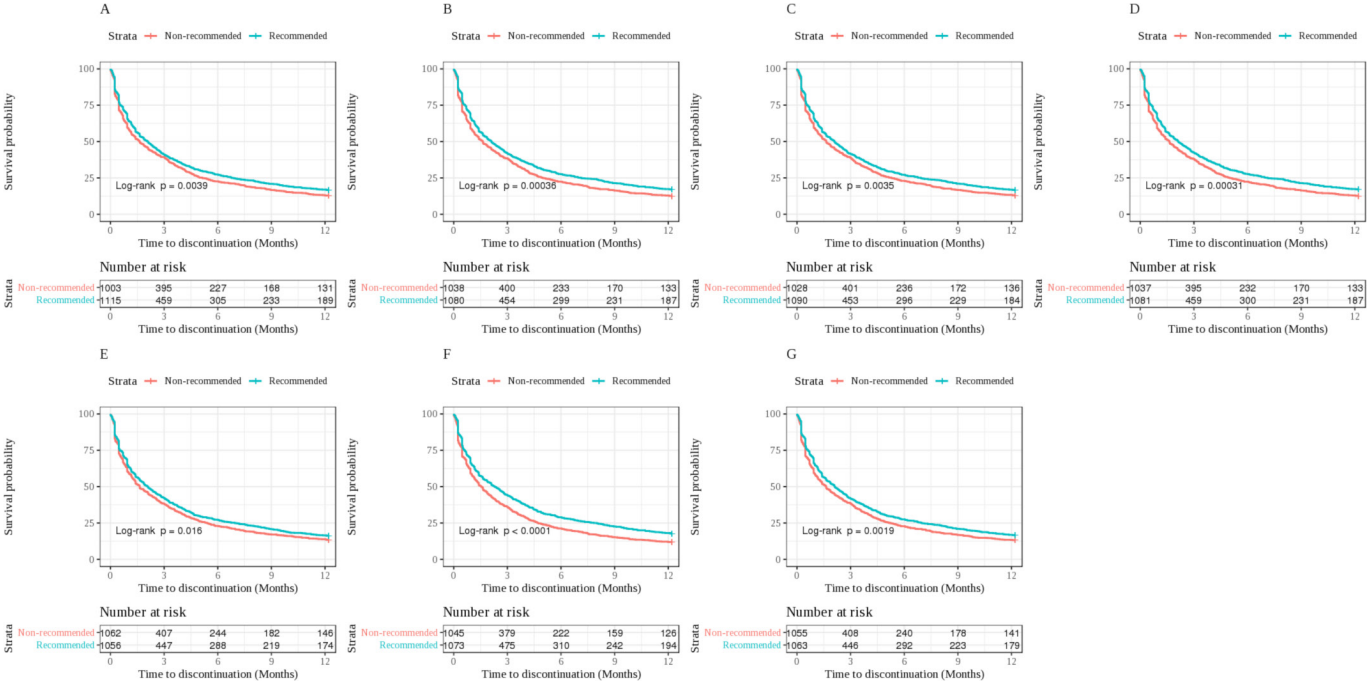


**Figure 1** Treatment continuation for overall populations in the test data set. (A) backward stepwise LR, (B) LASSO, (C) ridge, (D) elastic net, (E) random forest, (F) GBM and (G) SuperLearner. GBM, gradient boosting machine; LASSO, least absolute shrinkage and selection operator; LR, logistic regression.

**Table 3** Top 10 important predictors of the prediction models for risperidone and aripiprazole*

| Risperidone | | | | | | |
|---|---|---|---|---|---|---|
| **Rank** | **Backward stepwise LR** | **LASSO** | **Ridge** | **Elastic net** | **Random forest** | **GBM** |
| 1 | Number of outpatient visits, non-MHD | Number of inpatient visits, non-MHD | Number of EM visits, MHD | Number of inpatient visits, MHD | Age | Age |
| 2 | Number of any route visits, non-MHD | Age | Number of inpatient visits, MHD | Age | Number of any route visits, non-MHD | Number of visits, non-MHDs |
| 3 | Number of inpatient visits, MHD | Number of EM visits, non-MHDs | Number of EM visits, non-MHDs | Number of EM visits, non-MHDs | Number of outpatient visits, non-MHD | Number of outpatient visits, non-MHD |
| 4 | Number of EM visits, non-MHDs | Number of EM visits, MHD | Age | Number of EM visits, MHD | Number of outpatient visits, MHD | ASD |
| 5 | Age | Anticholinergic drugs | Anticholinergic drugs | Anticholinergic drugs | Number of any route visits, MHD | Number of any route visits, MHD |
| 6 | Anticholinergic drugs | EBD | ASD | EBD | Number of outpatient visits, non-MHD | Number of outpatient visits, MHD |
| 7 | EBD | ASD | EBD | ASD | First diagnosis of Schizophrenia | Mental retardation |
| 8 | Bipolar disorder | Insurance type: Medical aid | Number of inpatient visits, non-MHD | Insurance type: Medical aid | Sex: female | Insurance type: Medical aid |
| 9 | ASD | Mental retardation | Other congenital or genetic defect | Mental retardation | Insurance type: Medical aid | Seizure |
| 10 | Insurance type: Medical aid | Sex: female | Insurance type: Medical aid | Reaction to severe stress and adjustment disorder | Other MHDs | Antipyretics, analgesics |
| **Aripiprazole** | | | | | | |
| **Rank** | **Backward stepwise LR** | **LASSO** | **Ridge** | **Elastic NET** | **Random forest** | **GBM** |
| 1 | Number of outpatient visits, non-MHD | Age | Number of EM visits, MHD | Age | Age | Age |
| 2 | Number of any route visits, non-MHD | Number of inpatient visits, non-MHD | Number of inpatient visits, non-MHD | Number of inpatient visits, non-MHD | Number of outpatient visits, non-MHD | Number of any route visits, non-MHDs |
| 3 | Number of EM visits, MHD | Number of EM visits, MHD | Number of EM visits, non-MHDs | Number of EM visits, MHD | Number of any route visits, non-MHD | Number of outpatient visits, non-MHD |
| 4 | Number of inpatient visits, non-MHD | Lithium | Age | Lithium | Number of any route visits, MHD | ASD |
| 5 | Age | Number of EM visits, non-MHDs | Lithium | Number of EM visits, non-MHDs | Number of outpatient visits, MHD | Number of any route visits, MHD |
| 6 | Lithium | Bipolar disorder | Anticholinergic drugs | Bipolar disorder | First diagnosis of Schizophrenia | Number of outpatient visits, MHD |
| 7 | Tic disorder | Sex: female | Bipolar disorder | Sex: female | Antipyretics, analgesics | Mental retardation |
| 8 | MAO inhibitor | Number of any route visits, MHD | ASD | Number of any route visits, MHD | Sex: female | Insurance type: Medical aid |
| 9 | Bipolar disorder | Anticholinergic drugs | Diabetes mellitus | Anticholinergic drugs | Antihistamines | Seizure |
| 10 | ED | MAO inhibitor | MAO inhibitor | MAO inhibitor | Other MHDs | Antipyretics, analgesics |

*The relative importance of predictors in SuperLearner was not evaluated because this algorithm ensembles the predicted probabilities of each model used in construction.
ASD, autism spectrum disorder; BP, bipolar disorder; EBD, other behaviour and emotional disorders with onset generally occurring in childhood and adolescence; ED, emotional disorders with onset specific to childhood; EM, emergency room; GBM, gradient boosting machine; LASSO, least absolute shrinkage and selection operator; MAO, monoamine oxidase; MHD, mental health disorder.

## DISCUSSION

To our knowledge, this is the first study to develop and validate prediction models for 1-year treatment continuation of antipsychotics in children and adolescents with schizophrenia. Although some studies have reported prediction models for diverse treatment responses of antipsychotics, there is a scarcity of studies particularly conducted in this population. Moreover, the majority of previous studies applied only one prediction algorithm, whereas the present study has applied and compared several modelling algorithms (including both conventional LR and modern machine learning). The clinical benefit of prediction models for antipsychotic treatment was also evaluated.

In the present study, machine learning algorithms provided better performance than the conventional stepwise LR in predicting treatment continuation of antipsychotics. However, the increase in performance was too limited to confirm their outperformance over the conventional LR algorithm. This finding supported previous studies, such as Christodoulou *et al*, that compared the performance of LR and machine learning algorithms for clinical prediction modelling. They demonstrated no evidence that machine learning has a performance benefit over LR models in predicting clinical risk.[21] Machine learning has the advantage of improving performance by controlling both linear and non-linear interactions between outcomes and predictors; however, all these algorithms have disadvantages that require the tuning of complex hyperparameters, a high burden of data and cost and difficulty with interpretation.[22] Meanwhile, the conventional regression method has the advantage of being easy to interpret and consuming less time to develop.[22] Hence, conventional LR could be an attractive approach for predicting the treatment continuation of antipsychotics because of its advantages if it shows similar performance to machine-learning.

Machine learning is generally considered a better approach for prediction models than conventional LR, but machine learning failed to demonstrate its superiority to conventional LR in this study. There are some possible explanations for this finding. First, the number of sample cases and predictors might be insufficient for machine learning algorithms. Machine learning methods are data-driven algorithms,[22] which show better performance than conventional regression models when using massive data consisting of numerous predictors.[23] Second, most of the predictors used in our study were clinically meaningful predictors (socioeconomic factors, concomitant use of other psychotropics and comorbidities of mental health disorders), which might limit the ability of machine learning to discover the hidden interaction between predictors and outcomes. Previous studies that used structured data with clinically meaningful predictors, like our study, also reported little difference in performance between algorithms.[24 25] Third, as this study was conducted using claims data, most predictors were formed as categorical variables. One previous study suggested that machine learnings are more suitable for continuous variables than for categorical variables because they are non-parametric models and cannot assume linearity for predicting outcome association.[26] Furthermore, in the sensitivity analysis considering all predictors as categorical variables, the performance of models was worse than that of the primary analysis (online supplemental table S8).

Furthermore, the patients in the recommended group were more likely to continue their first-line treatment than those in the non-recommended group. The ratio of treatment continuation rate at the last follow-up between the recommended and nonrecommended groups ranged from 1.20 to 1.50. In a recent study conducted by Wu *et al*, mostly on adult patients (mean age 36.7 years, SD 14.3) in Taiwan, they developed a prediction model for each antipsychotic medication using the machine learning method.[12] They reported a similar range of increasing treatment success rate (treatment continuation for 1 year) from 1.16-fold to 1.48-fold when the individualised treatment rule was applied according to prediction models compared with the non-applied treatment.[12] Due to the heterogeneity of data source, study definition and patient demographics, it is difficult to directly compare the result of the present study with their study. However, considering the consistent result with that study, the present result provided additional empirical evidence that the prediction model has the potential to be a useful tool to improve treatment response in clinical practice.

This study has several limitations. First, as schizophrenia was determined using the ICD-10 code in our study, the possibility of misdiagnosis might not be excluded, especially for EOS. Vernal *et al* have shown that the diagnosis of EOS has generally good validity for the paediatric populations in claims data but found a slightly higher number of false positives for diagnosis in the outpatient setting.[27] Second, patients might discontinue or change the antipsychotic treatment regimen because of worsening of symptoms or the onset of other mental health disorders, not because they showed a poor response to their first-line antipsychotic agent. To overcome this limitation, several events indicating the worsening of psychiatric symptoms were included to identify poor treatment results.[12 28] Third, it was impossible to include all predictors that might improve the accuracy of the prediction model, such as the type of schizophrenia, disease severity and residential area, due to the lack of data. Fourth, it is possible there are other machine learning algorithms that might outperform conventional regression. Finally, there might be potential bias because of the lack of external validation in populations with different demographic profiles.[29] However, the purpose of our study was preliminary research to develop and compare the performance of prediction models for antipsychotic treatment response. This limitation regarding the generalisability may be addressed in future studies through external validation in various clinical settings.[30]

## CLINICAL IMPLICATIONS

Machine learning algorithms provided significantly limited performance improvement over conventional LR for predicting the 1-year treatment continuation of antipsychotics in children and adolescents with schizophrenia. Given the difficulties of hyperparameter tuning and understanding the result of machine learning, conventional LR might still be an attractive approach for constructing prediction models. The application of prediction models might be a helpful decision tool for antipsychotic treatment at the individual patient level.

**Contributors** SMJ, JC and J-WK had full access to all study data and take responsibility for the integrity of the data and the accuracy of the data analysis. SMJ and J-WK were responsible for the study concept and design. All authors were involved in the acquisition and interpretation of the data. SMJ and JC drafted the manuscripts. All authors critically revised the manuscripts. All authors provided their final approval of the version to be published and agreed to be accountable for all aspects of the work. SMJ and JC contributed equally. J-WK is responsible for the overall content as the gurantor.

**Funding** This study was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (NRF-2018R1D1A1B07043416). This research was supported by the BK21 FOUR Community-Based Intelligent Novel Drug Discovery Education Unit, College of Pharmacy and Research Institute of Pharmaceutical Sciences, Kyungpook National University. The supporting agencies had no role in the design and conduct of the study; collection, management, analysis and interpretation of the data; preparation, review or approval of the manuscript; and decision to submit the manuscript for publication.

**Competing interests** None declared.

**Patient consent for publication** Not applicable.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Data availability statement** Data are available upon reasonable request. This study used Health Insurance Review and Assessment Service database (M20190117533). Requests to access these datasets should be directed to HIRA; Official website of HIRA: https://opendata.hira.or.kr; Contact information of data access committee: +82-33-739-1083.



**ORCID iDs**
Soo Min Jeon http://orcid.org/0000-0002-5100-5739
Dong Yun Lee http://orcid.org/0000-0002-3678-9862
Jin-Won Kwon http://orcid.org/0000-0003-3467-7805

## REFERENCES

1 Kahn RS, Sommer IE, Murray RM, *et al*. Schizophrenia. *Nat Rev Dis Primers* 2015;1:15067.
2 Kendhari J, Shankar R, Young-Walker L. A review of childhood-onset schizophrenia. *Focus* 2016;14:328–32.
3 Clemmensen L, Vernal DL, Steinhausen H-C. A systematic review of the long-term outcome of early onset schizophrenia. *BMC Psychiatry* 2012;12:150.

4 Schimmelmann BG, Conus P, Cotton S, *et al*. Pre-Treatment, baseline, and outcome differences between early-onset and adult-onset psychosis in an epidemiological cohort of 636 first-episode patients. *Schizophr Res* 2007;95:1–8.
5 Keepers GA, Fochtmann LJ, Anzia JM, *et al*. The American psychiatric association practice guideline for the treatment of patients with schizophrenia. *Am J Psychiatry* 2020;177:868–72.
6 Case M, Stauffer VL, Ascher-Svanum H, *et al*. The heterogeneity of antipsychotic response in the treatment of schizophrenia. *Psychol Med* 2011;41:1291–300.
7 Schulz P, Berney P. Clinicians' predictions of patient response to psychotropic medications. *Dialogues Clin Neurosci* 2004;6:105–11.
8 Kinon BJ, Chen L, Ascher-Svanum H, *et al*. Predicting response to atypical antipsychotics based on early response in the treatment of schizophrenia. *Schizophr Res* 2008;102:230–40.
9 Beex-Oosterhuis MM, Heerdink ERR, Van Gool AR, *et al*. Predicting unsuccessful clozapine treatment after first use in adult patients with psychotic disorders. *J Clin Psychopharmacol* 2018;38:604–8.
10 Lee BS, McIntyre RS, Gentle JE, *et al*. A computational algorithm for personalized medicine in schizophrenia. *Schizophr Res* 2018;192:131–6.
11 Cao B, Cho RY, Chen D, *et al*. Treatment response prediction and individualized identification of first-episode drug-naïve schizophrenia using brain functional connectivity. *Mol Psychiatry* 2020;25:906–13.
12 Wu C-S, Luedtke AR, Sadikova E, *et al*. Development and validation of a machine learning individualized treatment rule in First-Episode schizophrenia. *JAMA Netw Open* 2020;3:e1921660.
13 Bylund DB, Reed AL. Childhood and adolescent depression: why do children and adults respond differently to antidepressant drugs? *Neurochem Int* 2007;51:246–53.
14 Patten SB, Waheed W, Bresee L. A review of pharmacoepidemiologic studies of antipsychotic use in children and adolescents. *Can J Psychiatry* 2012;57:717–21.
15 Collins GS, Reitsma JB, Altman DG, *et al*. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. *Br J Surg* 2015;102:148–58.
16 Altamura AC, Sassella F, Santini A, *et al*. Intramuscular preparations of antipsychotics: uses and relevance in clinical practice. *Drugs* 2003;63:493–512.
17 Korea MoFaDSi. Welcome to the Ministry of food and drug safety: Minisry of food and drug safety, 2022. Available: https://www.mfds.go.kr/eng/index.do [Accessed 26 Jan 2022].
18 Stroup TS, McEvoy JP, Swartz MS, *et al*. The National Institute of mental health clinical antipsychotic trials of intervention effectiveness (CATIE) project: schizophrenia trial design and protocol development. *Schizophr Bull* 2003;29:15–31.
19 Singh D, Singh B. Investigating the impact of data normalization on classification performance. *Appl Soft Comput* 2020;97:105524.
20 Shipe ME, Deppen SA, Farjah F, *et al*. Developing prediction models for clinical use using logistic regression: an overview. *J Thorac Dis* 2019;11:S574–84.
21 Christodoulou E, Ma J, Collins GS, *et al*. A systematic review shows no performance benefit of machine learning over logistic regression for clinical prediction models. *J Clin Epidemiol* 2019;110:12–22.
22 Bzdok D, Altman N, Krzywinski M. Statistics versus machine learning. *Nat Methods* 2018;15:233–4.
23 Beam AL, Kohane IS. Big data and machine learning in health care. *JAMA* 2018;319:1317–8.
24 Ye Y, Xiong Y, Zhou Q, *et al*. Comparison of machine learning methods and conventional logistic regressions for predicting gestational diabetes using routine clinical data: a retrospective cohort study. *J Diabetes Res* 2020;2020:1–10.
25 Khera R, Haimovich J, Hurley NC, *et al*. Use of machine learning models to predict death after acute myocardial infarction. *JAMA Cardiol* 2021;6:633–41.
26 Desai RJ, Wang SV, Vaduganathan M, *et al*. Comparison of machine learning methods with traditional models for use of administrative claims with electronic medical records to predict heart failure outcomes. *JAMA Netw Open* 2020;3:e1918962.
27 Vernal DL, Stenstrøm AD, Staal N, *et al*. Validation study of the early onset schizophrenia diagnosis in the Danish psychiatric central research register. *Eur Child Adolesc Psychiatry* 2018;27:965–75.
28 Oh S, Lee TY, Kim M, *et al*. Effectiveness of antipsychotic drugs in schizophrenia: a 10-year retrospective study in a Korean tertiary hospital. *NPJ Schizophr* 2020;6:32.
29 Obermeyer Z, Powers B, Vogeli C, *et al*. Dissecting racial bias in an algorithm used to manage the health of populations. *Science* 2019;366:447–53.
30 Suchard MA, Schuemie MJ, Krumholz HM, *et al*. Comprehensive comparative effectiveness and safety of first-line antihypertensive drug classes: a systematic, multinational, large-scale analysis. *Lancet* 2019;394:1816–26.