Supplemental material

BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s)

Evid Based Ment Health

Table of contents

Figure S1. Flow chart for the selection of study populations

25,781 patients with the record of antipsychotics prescriptions and schizophrenia diagnosis between Jan 1, 2008 and Dec 31, 2017

Exclued
- Patients who had any history of schizophrenia diagnosis or antipsychotics prescriptions at least two years before the index date (n=9,295)
- Patients who aged <2 or >18 at the index date (n=3,753)
- Patients who were inpatients of psychiatry at the index date (n=227)
- Patients who used long-acting injection antipyschotics or polypharmacy of antipyschotics at the index date (n=1,276)
- Patients who used other antipsychotics except risperidone or aripirpazole (n=2,728)

8,502 patients who were finally included in anlaysis
- **5,109** received Risperidone
- **3,393** received Aripiprazole

6384 Included in training dataset

2,118 Included in test dataset

Table S1. The list of code for the medications and diagnosis used for developing the prediction models for treatment continuations of antipsychotic treatment

Comorbidities of other mental health disorders

Classification	ICD-10 code
Anxiety disorder	F40, F41
Depression	F32, F33
ADHD	F90
Somatoform disorder	F45
Reaction to severe stress and adjustment disorder	F43
Other neurotic disorders	F48
Mental retardation	F70-F79
Other behavior and emotional disorders with onset generally occurring in childhood and adolescence	F98
Emotional disorders with onset specific to childhood	F93
Nonorganic disorder	F51
Tic disorder	F95
Bipolar disorder	F31
Autism spectrum disorder	F84
Other mental health disorders	Other diagnosis with F code

Prescription of other psychotropic medications

Classification	ATC code	Agents in that classification
Antiepileptic drugs	N03A	Phenobarbital, Primidone, Phenytoin, Fosphenytoin, Ethosuximide, Clonazepam, Carbamazepine, Oxcarbazepine, Rufinamide, Valproic acid, Vigabatrin, Lamotrigine, Topiramate, Gabapentin, Levetiracetam, Zonisamide, Pregabalin, Stiripentol, Lacosamide
Anticholinergic drugs	N04A	Trihexyphenidyl, Biperiden, Procyclidine, Benztropine
Antianxiety drugs	N05B	Diazepam, Chlordiazepoxide, Lorazepam, Bromazepam, Clobazam, Alprazolam, Ethyl loflazepate, Etizolam, Clotiazepam, Tofisopam, Hydroxyzine, Buspirione
Antidepressants drugs	N06A	Fluoxetine, Citalopram, Paroxetine, Sertraline, Fluvoxamine, Escitalopram, Venlafaxine, Milnacipran, Duloxetine, Desvenlafaxine, Clomipramine, Amitriptyline, Doxepin, Amoxapine, Moclobemide, Trazodone, Mirtazapine, Bupropion, Tianeptine
Stimulants for ADHD	N06B	Methylphenidate
Nonstimulants for ADHD	N06B	Modafinil, Atomoxetine, Clonidine

Comorbidities of general clinical conditions

Classification	ICD-10 code
Neurologic and neuromuscular	Q00, Q01, Q02, Q03, Q04, Q05, Q06, Q07, G901, E750, E751, E752, E754, G111, G112, G113, G114, G118, G119, G120, G121, G122, G128, G129, G310, G318, ,G328, G938, G939, G94, G911, G319, G253, G951, G958, G909, G851, G80, G371, G372, G378, G81, G82, G83, G931, G935, R402, I63, I635, G71, G72, G10, G20, G210, G211, G218, G230, G231, G232, G238, G240, G248, G253, G254, G255, G258, G259, G803, T850, T851, Z982
Cardiovascular	Q20, Q21, Q22, Q23, Q24, Q25, Q26, Q282, Q283, Q289, I340, I348, I360, I368, I370, I378, I42, I43, I515, I44, I45, I47, I48, I490, I491, I492, I493, I494, I495, I498, I499, R001, I270, I271, I272, 278, I279, I509, I517, I518, I631, I632, Z951, T825, T821, T820, T822, T826, T827, Z450, Z950, Z952, Z953, Z958, Z959, T862, Z941

3

Respiratory	Q30, Q31, Q32, Q33, Q34, P280, G473, I278, I43, J841, J961, Z902, E84, J950, J958, Z430, Z930, Z990, Z991, Z942
Renal and urologic	Q60, Q61, Q62, Q63, Q64, N18, Z905, Z906, G834, N312, N319, Z935, Z936, Z992, Z435, Z436, Z466, T861, Z940
Gastrointestinal	Q390, Q391, Q392, Q393, Q394, Q41, Q42, Q43, Q44, Q45, K73, K74, K754, K760, K761, K762, K763, K765, K768, K50, K51, I820, K551, K562, K593, Z980, Z903, Z904, Z931, Z932, Z933, Z934, Z431, Z432, Z433, Z434, Z465, T864, Z944

Classification	**ICD-10 code**
Hematologic or immunologic	D55, D56, D57, D58, D60, D61, D71, D80, D81, D82, D83, D84, D85, D86, D87, D88, D89, D720, M303, M359, D66, D682, D694, D70, D761, D762, D763, D86, B20, B21, B22, B23, B24, M300, M310, M311, M313, M314, M316, M321, M339, M340, M341, M349
Metabolic	E700, E702, E703, E708, E709, E710, E711, E712, E713, E714, E715, E720, E721, E722, E723, E724, E728, E729, E740, E741, E742, E743, E744, E748, E749, E75, E770, E771, E780, E781, E782, E783, E784, E785, E786, E787, E788, E789, E881, E888, E760, E761, E762, E763, E85, E791, E798, E804, E805, E806, E807, E830, E831, E833, E834, D841, E88, H498, E009, E230, E232, E222, E233, E237, E240, E242, E243, E248, E249, E268, E250, E258, E259, Z4681, Z9641
Other congenital or genetic defects	Q90, Q91, Q92, Q93, Q94, Q95, Q96, Q97, Q98, Q99, E343, M410, M412, M4130, M418, M419, M4330, M965, Q722, Q750, Q752, Q759, Q760, Q761, Q762, Q764, Q765, Q766, Q767, Q77, Q78, K449, Q790, Q791, Q792, Q793, Q794, Q795, Q799, Q81, Q871, Q872, Q873, Q874, Q878, Q897, Q899, Q992
Malignancy	C00, C01, C02, C03, C04, C05, C06, C07, C08, C09, C10, C11, C12, C13, C14, C15, C16, C17, C18, C19, C20, C21, C22, C23, C24, C25, C26, C27, C28, C29, C30, C31, C32, C33, C34, C35, C36, C37, C38, C39, C40, C41, C42, C43, C44, C45, C46, C47, C48, C49, C50, C51, C52, C53, C54, C55, C56, C57, C58, C59, C60, C61, C62, C63, C64, C65, C66, C67, C68, C69, C70, C71, C72, C73, C74, C75, C76, C77, C78, C79, C80, C81, C82, C83, C84, C85, C86, C87, C88, C89, C90, C91, C92, C93, C94, C95, C96, D00, D01, D02, D03, D04, D05, D06, D07, D08, D09, D37, D38, D39, D40, D41, D42, D43, D44, D45, D46, D47, D48, Q850, T860, Z948
Premature and neonatal	P050, P051, P052, P059, P070, P072, P100, P101, P104, P524, P528, P115, P210, P219, P250, P251, P252, P253, P258, P270, P271, P278, P916, P950, p350, P351, P252, P560, P570, P578, P613, P614, P77, P832, P912,
Miscellaneous, not elsewhere classified	T840, T841, T844, T845, T846, T847, T868, T869, T870, T871, T872, Y831, Y833, Z998, T868, T869
Asthma	J45, J46
Diabetes mellitus	E10, E11, E12, E13, E14
Seizure	R56, G40
Prescriptions of other general medications [a]	
Classification	**Drug class (no.)**
Antipyretics, analgesics	114
Antihistamines	141
Other antiallergic agents	142
Cardiovascular system drugs	210
Respiratory system drug	220
Drugs for diabetes Mellitus	396
Antibiotic preparations	610
Narcotic analgesics	811,821

4

Table S2. Definition of predictors included for the development of prediction models [a]

Classification	Definition
Sex	Sex recorded in the data source
Age	Age measured at the index date
Insurance type	Health insurance type recorded in the data source at the index date (Health insurance and medical aid)
1st diagnosis of schizophrenia at the index date	The first diagnosis date of schizophrenia was the index date (yes or no)
Comorbidities of mental health disorders a	At least one diagnosis record of each mental health disorder for one year before the index date (yes or no)
Prescriptions of other psychotropic medications	At least one prescription record of each other psychotropic medication (yes or no)
Comorbidities of other chronic conditions	At least one diagnosis record of each chronic condition (yes or no)
Prescriptions of other general medications	At least one prescription record of each general medication (yes or no)
Utilization of health care system for mental health disorders, any route	Number of any route visits for mental health disorders
Utilization of health care system for mental health disorders, inpatient	Number of inpatient visits for mental health disorders
Utilization of health care system for mental health disorders, outpatient	Number of outpatient visits for mental health disorders
Utilization of health care system for mental health disorders, emergency room	Number of emergency room visits for mental health disorders
Utilization of health care system for nonmental health disorders, any route	Number of any route visits for non-mental health disorders
Utilization of health care system for nonmental health disorders, inpatient	Number of inpatient visits for non-mental health disorders
Utilization of health care system for nonmental health disorders, outpatient	Number of outpatient visits for non-mental health disorders
Utilization of health care system for nonmental health disorders, emergency room	Number of emergency room visits for non-mental health disorders

a. All potential predictors were measured for one year before the index date.

b. The definition of mental health disorders was considered as F00-F99 according to ICD-10 codes.

Method S1. Machine Learning Model Development

The LASSO, ridge, and elastic net are the penalized regression algorithms. Their penalized regulation was tuned by performing 5-fold cross-validation (CV) with the grid search between 0.0001 and 100000(for elastic net, the penalized value was set to 0.5 by grid search, which equally balanced the model between LASSO and ridge penalties equally[1]).

Random forest and GBM are the tree-based ensemble methods, which are constructed by combining single decision tree models to strengthen the performance of the prediction models. Random forest was fitted using the number of trees as 500, which was best according to the 5-fold CV. In this process, we used the number of variables randomly sampled as candidates at each split as the default value (the root square of the number of predictors) because this value was insensitive to the performance of models.[2] For GBM, an interaction depth of 4 and a learning rate as 0.01 were used, and the 5-fold CV was conducted to tune the number of trees to build. The GBMs showed a consistent performance when these parameters were used in various conditions.[3]

SuperLearner is an ensemble method generated from a set of single prediction models. We used the SuperLearner package[45] to build this ensemble model using these tuned machine-learning models (LASSO, ridge, elastic net, randomforest, and GBM) and backward stepwise LR. The 5-fold CV process was performed to identify the optimal weight for combining these models.

Table. Corresponding hyperparameters for machine-learning algorithms

Algorithm	R package	Hyperparameter	Hyperparameter value	Best parameter setting [a]
LASSO	glmnet [6]	Lambda	Automatic grid search with 500 different values ranging from 0.001 to 100000	Risperidone: 0.004542706/ Aripiprazole: 0.006332908
Ridge	glmnet [6]	Lambda	Automatic grid search with 500 different values ranging from 0.001 to 100000	Risperidone: 0.1630891/ Aripiprazole: 0.2539859
Elastic net	glmnet [6]	alpha	0.25, 0.5, 0.75	0.5
		Lambda	Automatic grid search with 500 different values ranging from 0.001 to 100000	Risperidone: 0.007074559 Aripiprazole: 0.012770633
Random forest	randomforest [2]	Number of trees to build	100, 200, 300, 500, 1000	500
		Max depth	2, 4, 5	4
		Predictors	7	7
GBM	gbm [7]	Number of trees to build	500, 1000, 1500, 3000, 5000, 7000, 10000	10000
		Boosting learn rate	0.01	0.01
		Maximum nodes per tree	4	4

LASSO, least absolute shrinkage and selection operator; GBM, gradient boosting machine
a. The best parameter setting was determined to be the value that maximized the AUROC in the training dataset.
b. Other unmentioned hyperparameters were set to their default values.

Supplemental material

BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s)

Evid Based Ment Health

Figure S2. Classification matrix and equation of prediction metrics

Classification matrix

	Predicted category	
Actual category	Treatment continuation	Treatment discontinuation
Treatment continuation	True positive (TP)	False negative (FN)
Treatment discontinuation	False positive (FP)	True negative (TN)

$$\text{Sensitivity (recall)} = \frac{TP}{TP+FN}$$

$$\text{Specificity} = \frac{TN}{TN+FP}$$

$$\text{Precision} = \frac{TP}{TP+FP}$$

$$\text{Accuracy} = \frac{TP+TN}{TP+FN+FP+TN}$$

$$\text{F1} - \text{score} = 2 \times \frac{PR \times RE}{PR+RE}$$

$$\text{MCC} = \frac{TP \times TN - FP \times FN}{\sqrt{(TP+FN)(TP+FP)(TN+FP)(TN+FN)}}$$

Figure S3. Scheme of assignment of optimal medication for individual patients using the prediction models
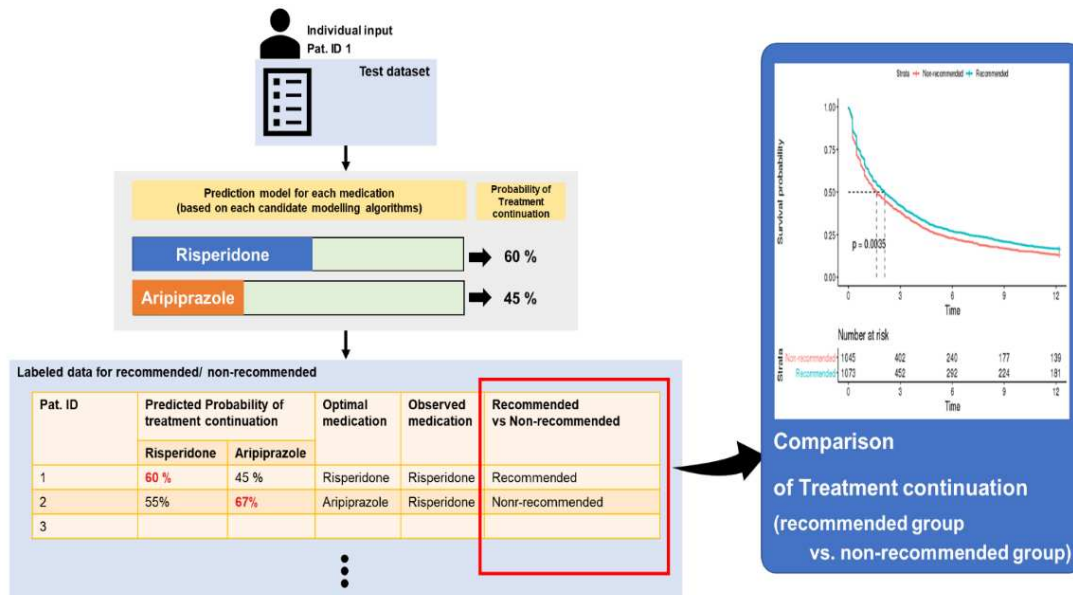
Proportion of patients with treatment continuation

$$= \frac{Number\ of\ patients\ with\ treatment\ continuation}{Number\ of\ patients\ in\ the\ recommended\ (or\ nonrecommended)\ group}$$

8

Table S3. Baseline characteristics of patients with risperidone

	Train (n=3826)	Test (n=1283)	p-value
Treatment continuation for 1 year (n, %)	531 (13.9)	193 (15.0)	0.323
Demographics [a] (n, %)			
Sex			
Male	2395 (62.6)	800 (62.4)	0.902
Female	1431 (37.4)	483 (37.6)	
Age (years, Mean ± SD)	13.8 ± 3.5	13.8 ± 3.4	0.839
Insurance type			
Health insurance	3249 (84.9)	1094 (85.3)	0.796
Medical aid	577 (15.1)	189 (14.7)	
Comorbidities of mental health disorders [b] (n, %)			
First diagnosis of schizophrenia at index date	1445 (37.8)	524 (40.8)	0.054
Anxiety disorder	581 (15.2)	195 (15.2)	1.000
Depression	929 (24.3)	328 (25.6)	0.375
ADHD	892 (23.3)	269 (21.0)	0.089
Somatoform disorder	78 (2.0)	27 (2.1)	0.976
Reaction to severe stress, and adjustment disorder	329 (8.6%)	109 (8.5)	0.955
Other neurotic disorders	37 (1.0)	19 (1.5)	0.169
Mental retardation	408 (10.7)	124 (9.7)	0.337
Other behavior and emotional disorders with onset generally occurring in childhood and adolescence	94 (2.5)	30 (2.3)	0.893
Emotional disorders with onset specific to childhood	229 (6.0)	56 (4.4)	0.034
Nonorganic disorder	86 (2.2)	41 (3.2)	0.075
Tic disorder	145 (3.8)	47 (3.7)	0.903
Bipolar disorder	90 (2.4)	30 (2.3)	1.000
Autism spectrum disorder	200 (5.2)	77 (6.0)	0.323
Other mental health disorders	873 (22.8)	277 (21.6)	0.383
Prescriptions of other psychotropics [b] (n, %)			
Anticholinergic drugs	21 (0.5)	13 (1.0)	0.116
Antianxiety drugs	1103 (28.8)	386 (30.1)	0.411
SSRI/SNRI	1035 (27.1)	334 (26.0)	0.499
TCA	75 (2.0)	32 (2.5)	0.297
MAO inhibitor	150 (3.9)	46 (3.6)	0.648
Stimulants for ADHD	692 (18.1)	200 (15.6)	0.046
Non-stimulants for ADHD	138 (3.6)	40 (3.1)	0.460
Antiepileptic drugs	475 (12.4)	137 (10.7)	0.108
Lithium	45 (1.2)	13 (1.0)	0.746
Comorbidities of pediatrics chronic conditions [b, c] (n, %)			
Neurologic and neuromuscular	94 (2.5)	34 (2.7)	0.780
Cardiovascular	52 (1.4)	19 (1.5)	0.853
Respiratory	1 (0.0)	0 (0.0)	1.000
Renal and urologic	51 (1.3)	8 (0.6)	0.056
Gastrointestinal	65 (1.7)	16 (1.2)	0.321
Hematologic or immunologic	7 (0.2%)	5 (0.4)	0.322
Metabolic	167 (4.4)	53 (4.1)	0.781
Other congenital or genetic defects	92 (2.4)	30 (2.3)	0.977

9

Supplemental material

BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s)

Evid Based Ment Health

Malignancy	23 (0.6)	3 (0.2)	0.170
Neonatal	2 (0.1)	1 (0.1)	1.000
Asthma	580 (15.2)	218 (17.0)	0.129
Diabetes mellitus	36 (0.9)	18 (1.4)	0.214
Seizure	305 (8.0)	92 (7.2)	0.386
Prescriptions of other general medications [b] (n, %)			
Antipyretics, analgesics	2195 (57.4)	765 (59.6)	0.167
Antihistamines	2555 (66.8)	868 (67.7)	0.588
Other antiallergic agents	5 (0.1)	2 (0.2)	1.000
Cardiovascular system drugs	273 (7.1)	90 (7.0)	0.934
Respiratory system drug	2698 (70.5)	931 (72.6)	0.173
Drugs for diabetes mellitus	2 (0.1)	5 (0.4)	0.017
Antibiotic preparations	2676 (69.9)	924 (72.0)	0.169
Narcotic analgesics	48 (1.3)	19 (1.5)	0.635
Utilization of health care system [b] (Mean ± SD)			
Number of any route visits for mental health disorders	1.7 ± 2.2	1.8 ± 4.9	0.372
Number of inpatient visits for mental health disorders	0.0 ± 0.5	0.2 ± 4.4	0.295
Number of outpatient visits for mental health disorders	1.6 ± 2.2	1.6 ± 2.2	0.972
Number of emergency room visit for mental health disorders	0.0±0.1	0.0±0.1	0.466
Number of any route visits for nonmental health disorders	12.0 ± 13.3	12.4 ± 13.7	0.296
Number of inpatient visits for nonmental health disorders	0.2 ± 1.0	0.2 ± 1.0	0.232
Number of outpatient visits for nonmental health disorders	11.8 ± 13.1	12.2 ± 13.5	0.307
Number of emergency room visits for nonmental health disorders	0.0 ± 0.3	0.0 ± 0.3	0.402

ADHD, attention deficit hyperactivity disorder; SSRI, selective serotonin reuptake inhibitor; SNRI, selective norepinephrine reuptake inhibitor; MAO inhibitor, monoamine oxidase inhibitor

a. Assessed on the cohort entry

b. Assessed for one-year before the index date

c. The pediatric complex chronic conditions were classified according to the modified the criteria of Feudtner et al (2014).

10

Table S4. Baseline characteristics of patients with aripiprazole

	Train dataset (n=2,558)	Test dataset (n=835)	p-value
Treatment continuation for 1 year (n, %)	350 (13.7)	121 (14.5)	0.597
Demographics [a] (n, %)			
Sex			0.416
Male	1,455 (56.9)	489 (58.6)	
Female	1,103 (43.1)	346 (41.4)	
Age (years, Mean ± SD)	14.0 ± 3.4	14.2 ± 3.3	0.195
Insurance type			0.572
Health insurance	2,395 (93.6)	787 (94.3)	
Medical aid	163 (6.4)	48 (5.7)	
Comorbidities of mental health disorders [b] (n, %)			
First diagnosis of schizophrenia at index date	1,227 (48.0)	421 (50.4)	0.234
Anxiety disorder	552 (21.6)	189 (22.6)	0.553
Depression	883 (34.5)	301 (36.0)	0.446
ADHD	513 (20.1)	181 (21.7)	0.337
Somatoform disorder	76 (3.0)	15 (1.8)	0.089
Reaction to severe stress, and adjustment disorder	248 (9.7)	77 (9.2)	0.737
Other neurotic disorders	35 (1.4)	15 (1.8)	0.468
Mental retardation	140 (5.5)	44 (5.3)	0.891
Other behavior and emotional disorders with onset generally occurring in childhood and adolescence	52 (2.0)	14 (1.7)	0.615
Emotional disorders with onset specific to childhood	102 (4.0)	33 (4.0)	1.000
Nonorganic disorder	95 (3.7)	19 (2.3)	0.058
Tic disorder	169 (6.6)	51 (6.1)	0.669
Bipolar disorder	84 (3.3)	33 (4.0)	0.418
Autism spectrum disorder	117 (4.6)	28 (3.4)	0.157
Other mental health disorders	615 (24.0)	204 (24.4)	0.856
Prescriptions of other psychotropics [b] (n, %)			
Anticholinergic drugs	12 (0.5)	6 (0.7)	0.557
Antianxiety drugs	822 (32.1)	275 (32.9)	0.699
SSRI/SNRI	877 (34.3)	292 (35.0)	0.749
TCA	60 (2.3)	16 (1.9)	0.553
MAO inhibitor	139 (5.4)	48 (5.7)	0.796
Stimulants for ADHD	327 (12.8)	118 (14.1)	0.346
Non-stimulants for ADHD	125 (4.9)	53 (6.3)	0.120
Antiepileptic drugs	256 (10.0)	102 (12.2)	0.082
Lithium	32 (1.3)	9 (1.1)	0.830
Comorbidities of pediatrics chronic conditions [b, c] (n, %)			
Neurologic and neuromuscular	49 (1.9)	8 (1.0)	0.086
Cardiovascular	39 (1.5)	11 (1.3)	0.790
Respiratory	3 (0.1)	2 (0.2)	0.779
Renal and urologic	21 (0.8)	10 (1.2)	0.433
Gastrointestinal	53 (2.1)	10 (1.2)	0.140
Hematologic or immunologic	10 (0.4)	3 (0.4)	1.000
Metabolic	125 (4.9)	35 (4.2)	0.466

Supplemental material

BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s)

Evid Based Ment Health

|			
Other congenital or genetic defects	48 (1.9)	17 (2.0)	0.884
Malignancy	5 (0.2)	2 (0.2)	1.000
Neonatal	1 (0.0)	0 (0.0)	1.000
Asthma	422 (16.5)	128 (15.3)	0.459
Diabetes mellitus	21 (0.8)	9 (1.1)	0.634
Seizure	177 (6.9)	65 (7.8)	0.444
Prescriptions of other general medications [b] (n, %)			
Antipyretics, analgesics	1,579 (61.7)	534 (64.0)	0.267
Antihistamines	1,766 (69.0)	575 (68.9)	0.958
Other antiallergic agents	9 (0.4)	2 (0.2)	0.885
Cardiovascular system drugs	248 (9.7)	63 (7.5)	0.072
Respiratory system drug	1,829 (71.5)	591 (70.8)	0.721
Drugs for diabetes mellitus	5 (0.2)	2 (0.2)	1.000
Antibiotic preparations	1,846 (72.2)	614 (73.5)	0.469
Narcotic analgesics	41 (1.6)	11 (1.3)	0.674
Utilization of health care system [b] (Mean ± SD)			
Number of any route visits for mental health disorders	1.9 ± 2.9	2.0 ± 5.4	0.606
Number of inpatient visits for mental health disorders	0.1 ± 1.9	0.3 ± 4.8	0.366
Number of outpatient visits for mental health disorders	1.8 ± 2.2	1.8 ± 2.2	0.523
Number of emergency room visits for mental health disorders	0.0 ± 0.1	0.0 ± 0.1	0.137
Number of any route visits for nonmental health disorders	12.5 ± 13.0	12.1 ± 11.7	0.380
Number of inpatient visits for nonmental health disorders	0.2 ± 0.5	0.2 ± 0.8	0.205
Number of outpatient visits for nonmental health disorders	12.4 ± 12.8	11.9 ± 11.5	0.334
Number of emergency room visits for nonmental health disorders	0.0 ± 0.2	0.1 ± 0.3	0.092

ADHD, attention deficit hyperactivity disorder; SSRI, selective serotonin reuptake inhibitor; SNRI, selective norepinephrine reuptake inhibitor; MAO inhibitor, monoamine oxidase inhibitor

a. Assessed on the cohort entry

b. Assessed for one-year before the index date

c. The pediatric complex chronic conditions were classified according to the modified the criteria of Feudtner et al (2014).

Figure S4. Receiver operating curves and precision-recall curves for prediction models treatment continuation of risperidone and aripiprazole
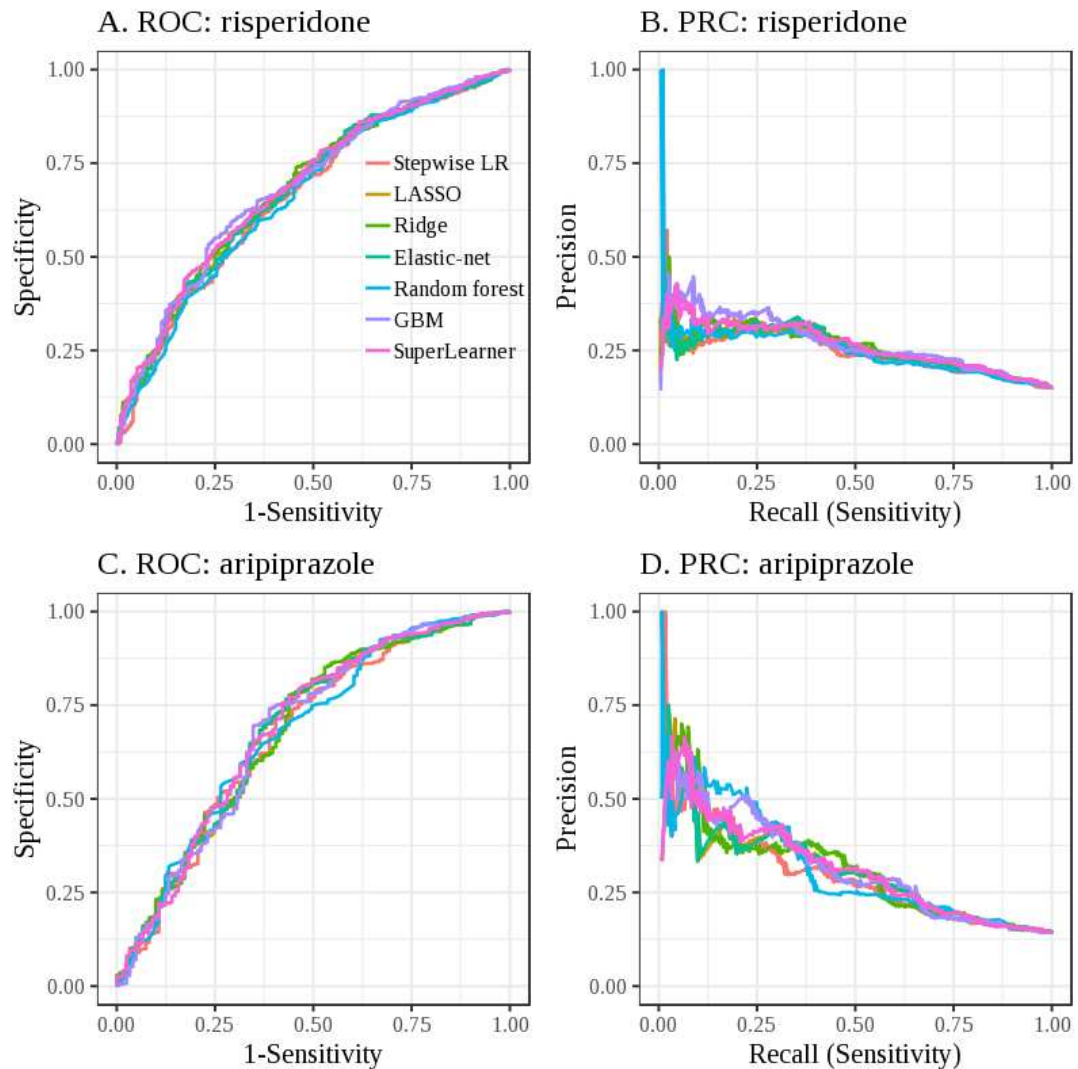
ROC, receiver operating curves; PRC, precision-recall curves; LR, logistic regression; LASSO, least absolute shrinkage and selection operator; GBM, gradient boosting machine

13

Table S5. AUROC with confidence intervals in the test dataset

	AUROC (95% CI)	
	Risperidone	**Aripiprazole**
Backward stepwise LR	0.659 (0.617–0.701)	0.672 (0.616–0.728)
LASSO	0.678 (0.638–0.719)	0.681 (0.625–0.738)
Ridge	0.680 (0.639–0.720)	0.683 (0.627–0.738)
e-net	0.676 (0.635–0.716)	0.683 (0.627–0.739)
Random forest	0.658 (0.616–0.700)	0.678 (0.623–0.733)
GBM	0.686 (0.646–0.726)	0.682 (0.626–0.739)
SuperLearner	0.685 (0.645–0.725)	0.688 (0.632–0.744)

AUROC, area under the receiver operating characteristic curve; CI, confidence intervals; LR, logistic regression, LASSO, least absolute shrinkage and selection operator; e-net, elastic net; GBM, gradient boosting machine

Supplemental material

BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s)

Evid Based Ment Health

Figure S5. Prediction performance of specificity, sensitivity, precision, accuracy, F1 score, and MCC in predicting the treatment continuation of risperidone and aripiprazole

(a) Risperidone

LR, logistic regression; LASSO, least absolute shrinkage and selection operator; GBM, gradient boosting machine; MCC, matthews correlation coefficients
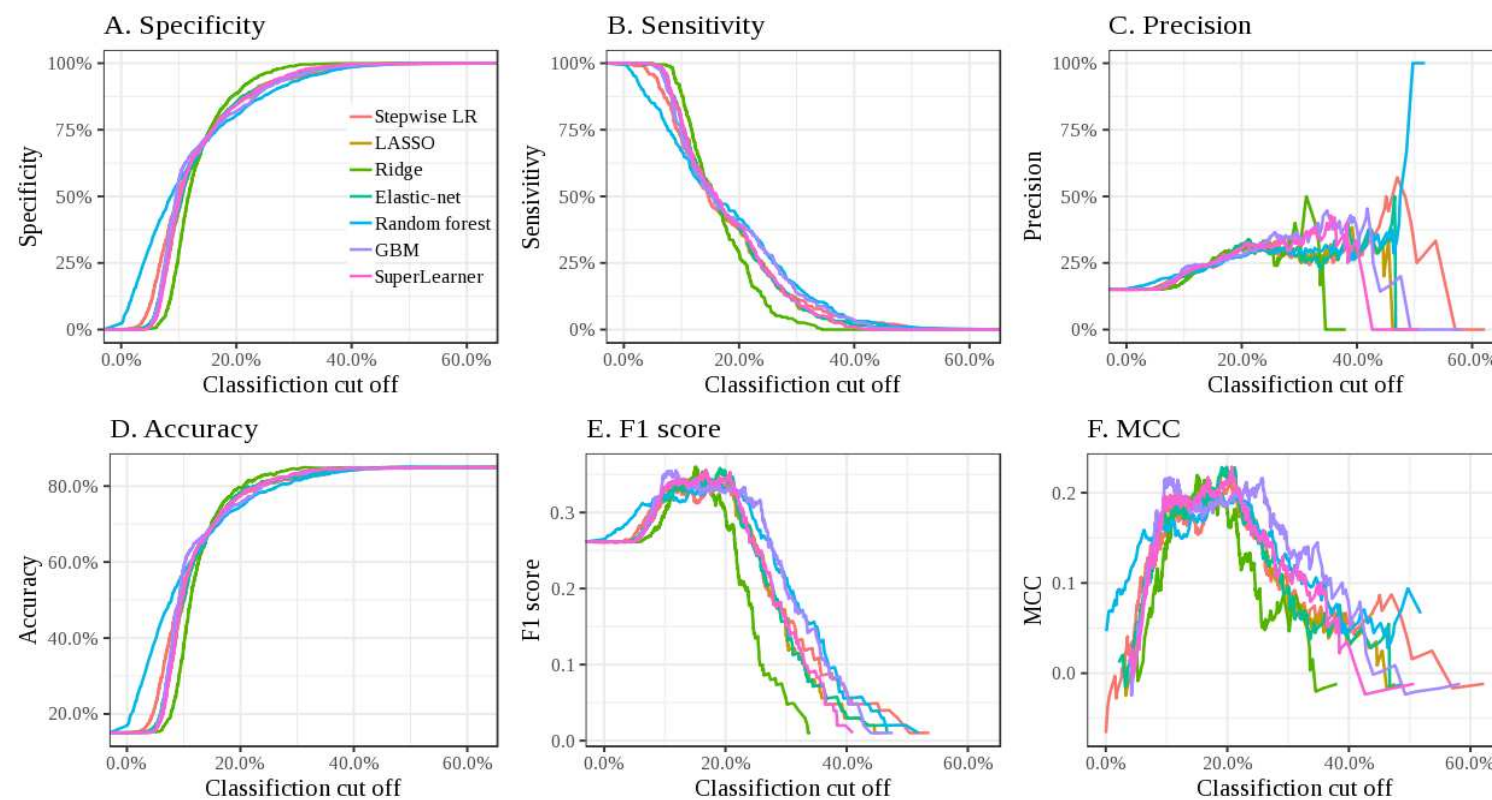
15

Figure S5. (Continued)

(b) Aripiprazole

LR, logistic regression; LASSO, least absolute shrinkage and selection operator; GBM, gradient boosting machine; MCC, matthews correlation coefficients

16

Figure S6. Treatment continuation for overall populations and each cohort in the test dataset

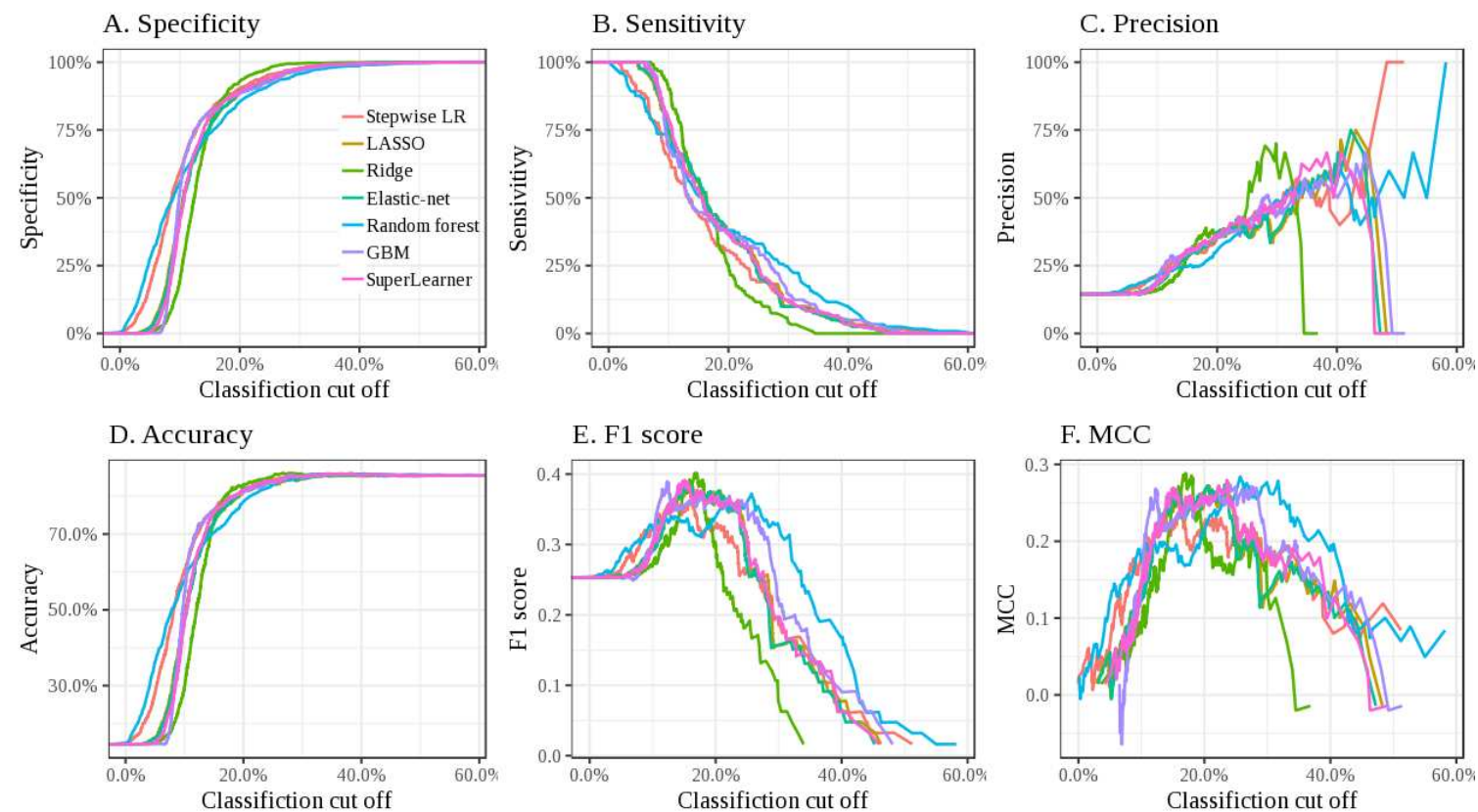(a) Risperidone

LR, logistic regression; LASSO, least absolute shrinkage and selection operator; GBM, gradient boosting machine (A) backward stepwise LR, (B) LASSO, (C) ridge, (D) elastic net, (E) random forest, (F) GBM, and (G) SuperLearner

17

Figure S6. (Continued)
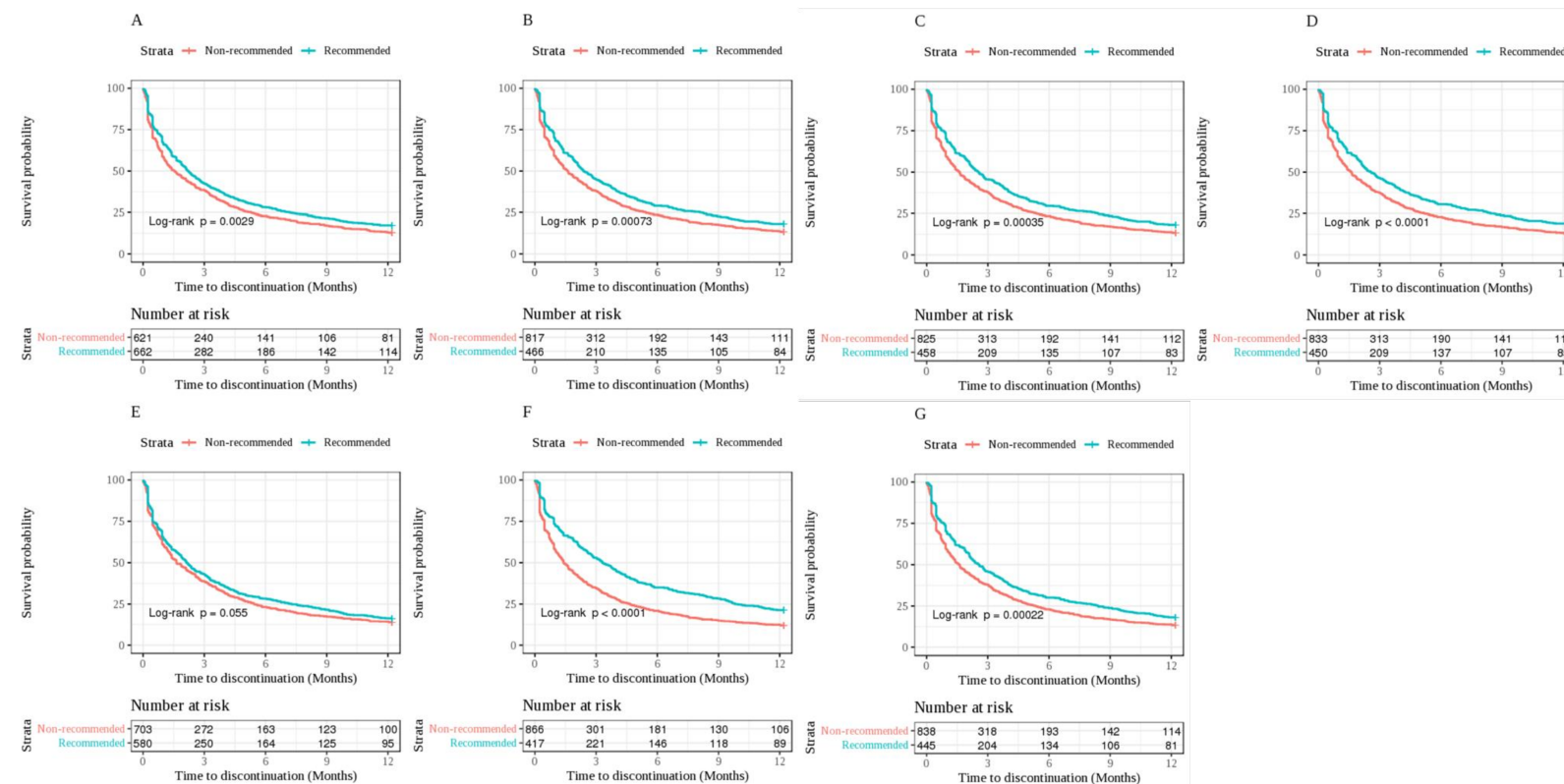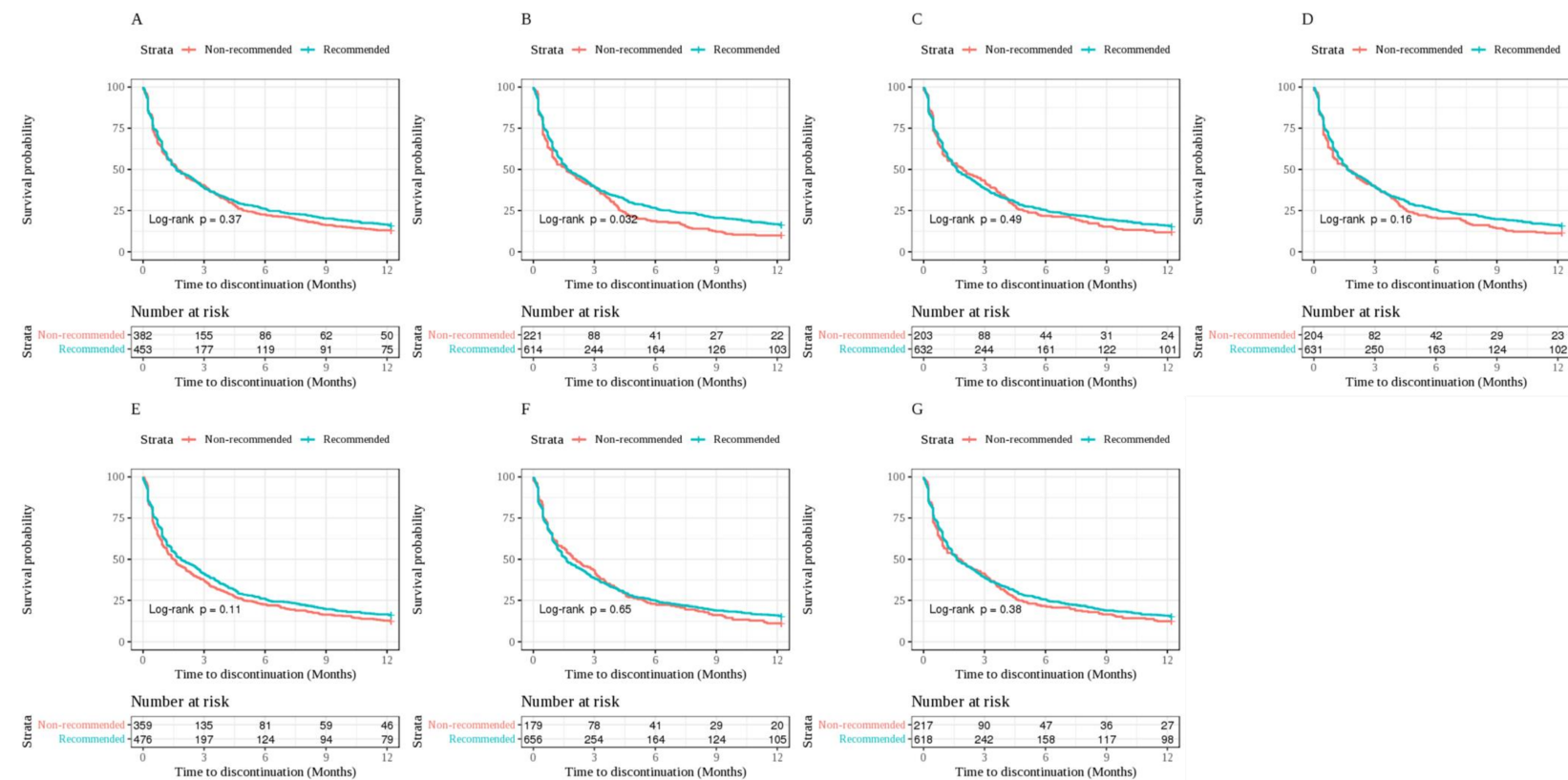
(b) Aripiprazole

LR, logistic regression; LASSO, least absolute shrinkage and selection operator; GBM, gradient boosting machine (A) backward stepwise LR, (B) LASSO, (C) ridge, (D) elastic net, (E) random forest, (F) GBM, and (G) SuperLearner

18

Table S6. Relative predictor importance of the different prediction models for treatment continuation of risperidone

	Backward Stepwise LR	LASSO	Ridge	Elastic net	Random forest	GBM
Demographics						
Female sex [a]	0.032	0.120	0.115	0.108	0.176	0.082
Age (years) [a]	0.160	0.749	0.573	0.634	1.000	1.000
Insurance type: medical aid	0.045	0.182	0.172	0.161	0.166	0.130
Comorbidities of mental health disorders [a]						
First diagnosis of schizophrenia at index date	0.018	0.055	0.082	0.054	0.184	0.071
Anxiety disorder	0.028	0.080	0.102	0.078	0.107	0.044
Depression			0.053	0.005	0.141	0.060
ADHD	0.016	0.059	0.104	0.052	0.134	0.081
Somatoform disorder		0.054	0.162	0.078	0.019	0.005
Reaction to severe stress, and adjustment disorder	0.039	0.117	0.136	0.110	0.076	0.025
Other neurotic disorders			0.105		0.013	0.001
Mental retardation	0.034	0.140	0.165	0.122	0.151	0.167
Other behavior and emotional disorders with onset generally occurring in childhood and adolescence	0.062	0.223	0.271	0.206	0.083	0.094
Emotional disorders with onset specific to childhood			0.071	0.010	0.099	0.086
Nonorganic disorder			0.095		0.023	0.005
Tic disorder	0.041	0.065	0.074	0.076	0.076	0.061
Bipolar disorder	0.050	0.048	0.113	0.064	0.038	0.017
Autism spectrum disorder	0.046	0.207	0.275	0.186	0.150	0.211
Other mental health disorders			0.037		0.163	0.059

LR, logistic regression; LASSO, least absolute shrinkage and selection operator; GBM, Gradient boosting machine; ADHD, attention deficit hyperactivity disorder; SSRI, selective serotonin reuptake inhibitor; SNRI, selective norepinephrine reuptake inhibitor; MAO inhibitor, monoamine oxidase inhibitor; MHD, mental health disorders
Continued on the next page

Table S6. (Continued)

	Stepwise LR	LASSO	Ridge	Elastic net	Random forest	GBM
Prescriptions of other psychotropics [a]						
Anticholinergic drugs	0.091	0.240	0.367	0.231	0.029	0.043
Antianxiety drugs	0.022	0.073	0.072	0.065	0.161	0.076
SSRI/SNRI			0.027		0.140	0.032
Tricyclic antidepressants			0.041		0.027	0.013
MAO inhibitor			0.064		0.045	0.011
Methylphenidate		0.023	0.099	0.025	0.121	0.087
Nonstimulants for ADHD	0.036	0.081	0.139	0.086	0.073	0.093
Antiepileptic drugs	0.035	0.061	0.083	0.060	0.118	0.072
Lithium		0.002	0.145	0.032	0.013	0.002
Comorbidities of general clinical conditions [a]						
Neurologic and neuromuscular		0.086	0.177	0.092	0.064	0.069
Cardiovascular			0.017		0.031	0.018
Renal and urologic			0.021		0.038	0.043
Gastrointestinal			0.093		0.025	0.003
Hematologic or immunologic			0.010		0.005	0.000
Metabolic	0.040	0.099	0.132	0.101	0.076	0.048
Other congenital or genetic defects		0.024	0.087	0.042	0.065	0.037
Asthma			0.028		0.146	0.071
Diabetes Mellitus			0.000		0.018	0.008
Seizure		0.078	0.117	0.074	0.106	0.102
Prescriptions of other general medications [a]						
Antipyretics, analgesics	0.016	0.037	0.063	0.038	0.198	0.099
Antihistamines	0.017	0.021	0.058	0.030	0.159	0.038
Cardiovascular system drugs			0.004		0.084	0.026
Respiratory system drug			0.010		0.162	0.070
Antibiotic preparations			0.041	0.005	0.155	0.056
Narcotic analgesics			0.100		0.021	0.001
Utilization of health care system [a]						
No. any route visits, MHD			0.009		0.418	0.193
No. inpatient visits, MHD	0.831	1.000	0.973	1.000	0.021	0.009
No. outpatient visits, MHD			0.023		0.421	0.182
No. EM visits, MHDs		0.446	1.000	0.559	0.004	0.000
No. any route visits, non-MHD	0.920		0.022		0.842	0.486
No. inpatient visits, non-MHD			0.228		0.155	0.080
No. outpatient visits, non-MHD	1.000	0.005	0.009	0.053	0.842	0.377
No. EM visits, non-MHDs	0.284	0.599	0.632	0.594	0.043	0.015

LR, logistic regression; LASSO, least absolute shrinkage and selection operator; GBM, Gradient boosting machine; ADHD, attention deficit hyperactivity disorder; SSRI, selective serotonin reuptake inhibitor; SNRI, selective norepinephrine reuptake inhibitor; MAO inhibitor, monoamine oxidase inhibitor; MHD, mental health disorders

Supplemental material

BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s)

Evid Based Ment Health

Table S7. Relative predictor importance of the different prediction models for treatment continuation of aripiprazole

	Backward Stepwise LR	LASSO	Ridge	Elastic net	Random forest	GBM
Demographics						
Female sex [a]	0.005	0.187	0.174	0.189	0.184	0.099
Age (years) [a]	0.024	1.000	0.694	1.000	1.000	1.000
Insurance type: medical aid	0.006	0.082	0.108	0.078	0.085	0.034
Comorbidities of mental health disorders [a]						
First diagnosis of schizophrenia at index date		0.000	0.011	0.000	0.220	0.102
Anxiety disorder	0.004	0.063	0.093	0.065	0.133	0.057
Depression	0.003	0.085	0.107	0.091	0.142	0.054
ADHD	0.005	0.129	0.186	0.138	0.159	0.119
Somatoform disorder			0.011		0.053	0.022
Reaction to severe stress, and adjustment disorder			0.094		0.077	0.014
Other neurotic disorders			0.135		0.010	
Mental retardation			0.012		0.098	0.057
Other behavior and emotional disorders with onset generally occurring in childhood and adolescence			0.070		0.039	0.025
Emotional disorders with onset specific to childhood	0.008	0.085	0.134	0.081	0.056	0.034
Nonorganic disorder	0.010	0.086	0.185	0.090	0.026	0.004
Tic disorder	0.008	0.300	0.352	0.313	0.154	0.181
Bipolar disorder			0.040		0.056	0.040
Autism spectrum disorder		0.071	0.215	0.081	0.108	0.123
Other mental health disorders			0.054		0.175	0.097

LR, logistic regression; LASSO, least absolute shrinkage and selection operator; GBM, Gradient boosting machine; ADHD, attention deficit hyperactivity disorder; SSRI, selective serotonin reuptake inhibitor; SNRI, selective norepinephrine reuptake inhibitor; MAO inhibitor, monoamine oxidase inhibitor; MHD, mental health disorders
Continued on the next page

21

Supplemental material

BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s)

Evid Based Ment Health

Table S7. (Continued)

	Stepwise LR	LASSO	Ridge	Elastic net	Random forest	GBM
Prescriptions of other psychotropics [a]						
Anticholinergic drugs		0.156	0.395	0.159	0.029	0.001
Antianxiety drugs		0.065	0.078	0.065	0.171	0.072
SSRI/SNRI		0.000	0.080		0.130	0.068
TCA		0.000	0.114		0.027	0.004
MAO inhibitor	0.009	0.142	0.192	0.143	0.037	0.015
Methylphenidate		0.000	0.058		0.101	0.098
Nonstimulants for ADHD		0.000	0.142		0.075	0.070
Antiepileptic drugs	0.005	0.099	0.113	0.096	0.111	0.068
Lithium	0.018	0.549	0.475	0.547	0.050	0.154
Comorbidities of general clinical conditions [a]						
Neurologic and neuromuscular			0.046		0.035	0.030
Cardiovascular			0.172		0.018	
Renal and urologic			0.069		0.016	0.003
Gastrointestinal			0.160		0.019	0.001
Hematologic or immunologic			0.056		0.013	
Metabolic		0.020	0.127	0.026	0.048	0.013
Other congenital or genetic defects			0.111		0.047	0.032
Asthma	0.003	0.069	0.148	0.076	0.172	0.135
Diabetes mellitus			0.203		0.008	0.000
Seizure	0.005	0.115	0.175	0.120	0.103	0.104
Prescriptions of other general medications [a]						
Antipyretics, analgesics	0.003	0.047	0.083	0.047	0.199	0.086
Antihistamines			0.004		0.176	0.076
Cardiovascular system drugs		0.043	0.119	0.048	0.071	0.015
Respiratory system drug			0.069		0.158	0.060
Antibiotic preparations			0.022		0.171	0.077
Narcotic analgesics			0.011		0.022	0.015
Utilization of health care system [a]						
No. any route visits, MHD		0.184	0.143	0.174	0.440	0.229
No. inpatient visits, MHD			0.183		0.031	0.005
No. outpatient visits, MHD	0.005		0.056		0.425	0.164
No. EM visits, MHDs	0.734	0.611	1.000	0.594	0.003	0.001
No. any route visits, non-MHD	1.000		0.104		0.839	0.522
No. inpatient visits, non-MHD	0.161	0.701	0.753	0.668	0.130	0.049
No. outpatient visits, non-MHD	1.047		0.114		0.852	0.464
No. EM visits, non-MHDs		0.338	0.742	0.354	0.043	0.019

LR, logistic regression; LASSO, least absolute shrinkage and selection operator; GBM, Gradient boosting machine; ADHD, attention deficit hyperactivity disorder; SSRI, selective serotonin reuptake inhibitor; SNRI, selective norepinephrine reuptake inhibitor; MAO inhibitor, monoamine oxidase inhibitor; MHD, mental health disorders

Table S8. Performance of machine-learning models in predicting treatment continuation in the test dataset (using age and utilization of health care system as categorical predictors)

Model	Risperidone		Aripiprazole	
	BS	AUROC (95% CI)	BS	AUROC (95% CI)
Backward stepwise LR	0.122	0.675 (0.634–0.717)	0.116	0.675 (0.620–0.731)
LASSO	0.121	0.685 (0.644–0.725)	0.115	0.688 (0.633–0.743)
Ridge	0.122	0.686 (0.645–0.727)	0.115	0.687 (0.632–0.743)
e-net	0.122	0.686 (0.645–0.727)	0.115	0.687 (0.632–0.742)
Random forest	0.139	0.627 (0.583–0.671)	0.133	0.639 (0.583–0.696)
GBM	0.127	0.670 (0.629–0.711)	0.123	0.675 (0.618–0.732)
SuperLearner	0.121	0.686 (0.645–0.727)	0.115	0.687 (0.632–0.743)

BS, brier score; AUROC, area under the receiver operating characteristic curve; CI, confidence intervals; LR, logistic regression; LASSO, least absolute shrinkage and selection operator; e-net, elastic net; GBM, gradient boosting machine

Supplemental material

BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance
placed on this supplemental material which has been supplied by the author(s)

Evid Based Ment Health

Reference

1. Zou H, Hastie TJJotrsssB. Regularization and variable selection via the elastic net. 2005;67(2):301-20.

2. Breiman L. Random Forests. *Machine Learning* 2001;45(1):5-32. doi: 10.1023/A:1010933404324

3. Trevor Hastie RT, Jerome Friedman. Boosting and additive trees. In: The Elements of Statistical Learning. 2 ed: Springer, New York, NY 2009.

4. van der Laan MJ, Polley EC, Hubbard AE. Super Learner. *Statistical Applications in Genetics and Molecular Biology* 2007;6(1) doi: doi:10.2202/1544-6115.1309

5. Kennedy C. Guide to SuperLearner 2017 [Available from: https://cran.r-project.org/web/packages/SuperLearner/vignettes/Guide-to-SuperLearner.html accessed August 12 2021.

6. Friedman J, Hastie T, Tibshirani R. Glmnet: Lasso and elastic-net regularized generalized linear models. *R Package Version* 2009;1

7. Generalized Boosted Models: A guide to the gbm package; 2006.

8. Steyerberg EW, Vickers AJ, Cook NR, et al. Assessing the performance of prediction models: a framework for traditional and novel measures. *Epidemiology* 2010;21(1):128-38. doi: 10.1097/EDE.0b013e3181c30fb2 [published Online First: 2009/12/17]