Original Article
https://doi.org/10.9758/cpn.2022.20.3.450 pISSN 1738-1088 / eISSN 2093-4327
Clinical Psychopharmacology and Neuroscience 2022;20(3):450-461 Copyrightⓒ 2022, Korean College of Neuropsychopharmacology

Machine-Learning for Prescription Patterns: Random Forest in the Prediction of Dose and Number of Antipsychotics Prescribed to People with Schizophrenia

Mattia Marchi[1], Giacomo Galli[1], Gianluca Fiore[1], Andrew Mackinnon[2], Giorgio Mattei[1], Fabrizio Starace[3], Gian M. Galeazzi[1,3,4]

[1]Department of Biomedical, Metabolic and Neural Sciences, University of Modena and Reggio Emilia, Modena, Italy, [2]Black Dog Institute, University of New South Wales, Sydney, NSW, Australia, [3]Department of Mental Health and Drug Abuse, Azienda Unità Sanitaria Locale (AUSL) Modena, Modena, [4]Dipartimento di Salute Mentale e Dipendenze Patologiche, AUSL-IRCSS di Reggio Emilia, Reggio Emilia, Italy

Objective: We aimed to predict antipsychotic prescription patterns for people with schizophrenia using machine learning (ML) algorithms.
Methods: In a cross-sectional design, a sample of community mental health service users (SUs; n = 368) with a primary diagnosis of schizophrenia was randomly selected. Socio-demographic and clinical features, including the number, total dose, and route of administration of the antipsychotic treatment were recorded. Information about the number and the length of psychiatric hospitalization was retrieved. Ordinary Least Square (OLS) regression and ML algorithms (i.e., random forest [RF], supported vector machine, K-nearest neighborhood, and Naïve Bayes) were used to estimate the predictors of total antipsychotic dosage and prescription of antipsychotic polytherapy (APP).
Results: The strongest predictor of the total dose was APP. The number of Community Mental Health Centers (CMHC) contacts was the most important predictor of APP and, with APP omitted, of dosage. Treatment with anticholinergics predicted APP, emphasizing the strong correlation between APP and higher antipsychotic dose. RF performed better than OLS regression and the other ML algorithms in predicting both antipsychotic dose (root square mean error = 0.70, R^2 = 0.31) and APP (area under the receiving operator curve = 0.66, true positive rate = 0.41, and true negative rate = 0.78).
Conclusion: APP is associated with the prescription of higher total doses of antipsychotics. Frequent attenders at CMHCs, and SUs recently hospitalized are often treated with APP and higher doses of antipsychotics. Future prospective studies incorporating standardized clinical assessments for both psychopathological severity and treatment efficacy are needed to confirm these findings.

KEY WORDS: Schizophrenia; Antipsychotic drugs; Drug prescription; Drug polytherapy; Machine learning.

INTRODUCTION

Schizophrenia is a complex phenotype, understood as a neurodevelopmental, polygenic, and multifactorial disorder [1,2]. It is characterized by polymorphic symptomatic manifestations, including reality distortion, cognitive disturbance, and negative symptoms, typically presenting phases of remission and relapses. Genetic, environmental and social factors have been established as playing important roles in its origin [3,4].

Although the prevalence of schizophrenia spectrum disorders is relatively low—approximately 0.47% for schizophrenia and 3.0% for other clinical diagnoses of psychotic disorders [5,6]—they are responsible for tremendous personal, economic and societal burden, with 218 disability adjusted life years (DALYs) per 100,000, making schizophrenia the fifth leading cause of DALYs in the 15—44 age group of [7].

Regarding the treatment of schizophrenia, the UK

Received: April 27, 2021 / **Revised:** August 30, 2021
Accepted: September 23, 2021
Address for correspondence: Mattia Marchi
Department of Biomedical, Metabolic and Neural Sciences, University of Modena and Reggio Emilia, Via del Pozzo 71, 41124 Modena, Italy
E-mail: mattiamarchimd@gmail.com
ORCID: https://orcid.org/0000-0003-2970-1276

ⓒ This is an Open-Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/4.0) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

National Institute for Health and Care Excellence (NICE) guidelines recommend the use of a single form of antipsychotic medication (monotherapy) as the first-line treatment [8]. NICE guidelines recommendations regarding the use of antipsychotics for the treatment of schizophrenia have not changed substantially since 2002 despite review and an update in 2014. Critically, there have been no changes to the recommendation to avoid use of combined antipsychotics and to keep the total antipsychotic dose within the limits indicated by the summary of product characteristics. Maintaining a record the reasons of any use outside the endorsed dosage ranges is also recommended [8,9].

In the 2014 update, the NICE guidelines advised the regular monitoring of indicators of overall physical health (i.e., weight, waist circumference, pulse and blood pressure, fasting blood glucose, lipid blood profile, nutritional status and physical activity), response to treatment and presence of any side effects potentially attributable to antipsychotic treatment [8].

The more recently promulgated Maudsley Prescriber guidelines [10] also recommend use the lowest possible dose of antipsychotics in the treatment of schizophrenia and psychosis. The guidelines state that dose should be titrated to the lowest effective level for each patient, and that the treating psychiatrist should consider an increase in dosage only if the patient does not respond after 2 weeks of treatment.

High-dose antipsychotic medication regimes may result from the prescription of either a single antipsychotic agent or from the combination of two or more antipsychotics as part of the clinical practice called antipsychotic polytherapy (APP) [11]. There is concern regarding the use of APP, mainly due to a lack of evidence of its safety and efficacy [12] as well as due to the potentially high total dose of antipsychotics resulting from APP. Many side effects of the antipsychotic treatment are either dose-related or drug specific. These include the wide array of antipsychotic-induced movement disorders (i.e., drug-induced parkinsonism [the term now preferred to extrapyramidal symptoms], akathisia, and tardive dyskinesia), sedation, anticholinergic effects (i.e., tachycardia, constipation, urinary retention, xerostomia, and cognitive impairment), sexual dysfunction, metabolic impairment (up to a frank metabolic syndrome), QTc prolongation, and coronary heart disease [13,14]. High number and dose of antipsychotics are commonly associated with greater burden of adverse reactions. Pharmacological strategies for reducing the number of antipsychotics delivered, favoring monotherapy, have been proposed [15]. Nevertheless, the practice of prescribing APP is common worldwide [11,16-18]. Observational studies report various reasons for APP. These include the need of prolonged cross-titration when switching between antipsychotics, combination of different mechanisms of action, minimizing side effects of a single antipsychotic agent, treating comorbidities, and ineffective antipsychotic monotherapy [12,19,20].

In recent years, the use of machine learning (ML) has grown to include applications in the social sciences and medical research. Classical Ordinary Least Square (OLS) regression models are limited in their ability to manage nonlinearity and collinearity issues. To overcoming these limitations, many ML algorithms have been developed. These include (but are not limited to) random forest (RF), supported vector machine (SVM), K-nearest neighborhood (KNN), Naïve Bayes (NB). In psychiatry, emerging ML studies have significantly predicted the probability of response to antipsychotic treatment in patients diagnosed with schizophrenia using SVM and RF for continuous outcomes and KNN and NB for classification problems [21,22]; and psychosocial outcomes in people treated with antipsychotic monotherapy, through RF [23].

ML algorithms have different predictive capabilities when applied to different tasks and different types of data [24]. A recent systematic review concluded that RF generally yields superior accuracy for disease prediction [25], but no study compared the performance of these methods in predicting drugs prescription. However, ML may be appropriate in the study of prescription patterns as it is able to unravel patterns in complex, multivariate, and non-linear datasets [26].

Identifying predictors of prescription patterns that diverge from the current, established guidelines could provide a profile of the types of SUs who are more likely to receive higher doses of antipsychotic and APP in the future, guiding clinical practice and interventions in real-world settings.

The aim of this study was to compare the performance of OLS regression and different ML algorithms in predicting the dose and the number of antipsychotics prescribed to people with schizophrenia, using real-world data.

Specifically, based on the available literature [25], we chose four of the most widely used ML algorithms: RF, SVM, KNN, NB.

METHODS

Study Design and Sample Size

The population studied consisted of a cross-sectional random sample of Community Mental Health Service Users (SUs) from the catchment area of the Department of Mental Health and Drug Abuse of Modena (Italy) with a primary diagnosis of Schizophrenia Disorders, corresponding to ICD-10 code F20, and subclasses [27]. A target sample size of 400 was selected based on a desire to maintain 95% confidence intervals (95% CIs) within ±5% (i.e., 0.05), conservatively assuming a prevalence of 0.5 [28,29]. The sample was selected using a computer-generated random sequence to include an equal number of SUs taking oral or long-acting injectable (LAI) antipsychotics.

Data were collected from SUs' electronic medical records and clinical paper records held by Community Mental Health Centers (CMHC) for outpatients. In Italy, CMHC are the main providers of care for people with mental health conditions. Records were screened for the required information by the investigators (G.G. and G.F.); when the information was not clear or the screener was in doubt, the specific issue was discussed between the members of the research team and with the referring psychiatrist's SU.

The following data were collected:

- Presence/absence of the following indicators:
  - Medication side effects;
  - Electrocardiogram (ECG) and QTc measure;
  - Body mass index (BMI);
  - Blood exams;
- Mode of administration of antipsychotic medication, namely oral or LAI;
- Prescription of APP;
- The total dose of the prescribed antipsychotic medication in chlorpromazine equivalents milligrams [10,30];
- Prescription of medications with anticholinergic effect;
- Stability of the prescription in the previous year, in terms of in dose and type of antipsychotic medication prescribed;
- Number, length (days) and type (voluntary or compulsory admission) of psychiatric hospitalization in the last year;
- Number and type (urgent or routine) of CMHC contacts both with the psychiatrist and with other mental health professionals;
- Socio-demographic attributes:
  - Age;
  - Sex;
  - Marital status;
  - Employment;
  - Schooling;
  - Citizenship.

A total of 400 clinical records were randomly selected. There was substantial missing information in 32 records, which were excluded from analyses (i.e., the final n = 368).

Figure 1 displays the flow chart of the enrollment and selection process, compliant with the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guidelines.

Statistical Analyses

Frequencies and percentages were used for describing categorical and dichotomous variables. Continuous variables were summarized by means, standard deviations (SD), medians, and ranges.

Analyses were performed to investigate the role of possible predictors of the total dose (continuous) and the APP prescribed (dichotomous), implementing ML algorithms for regression and classification problems, respectively. The analyses were performed with R [31], using the "caret" package [32]. We implemented a train/test split ratio of 0.75/0.25 and employed 5 × 5 cross-validation, to avoid bias in prediction, for all the analyses. Performance of the models was compared on the basis of each model's root mean square error (RMSE) or R^2 for the regression models. Classification models were compared using differences in the area under the receiving operator curve (AUROC), true positive rate, true negative rate, and overall accuracy.

Prediction of Total Dose

Total antipsychotic dose was log transformed to reduce righthand skew of the distribution. There were n = 276 ob-

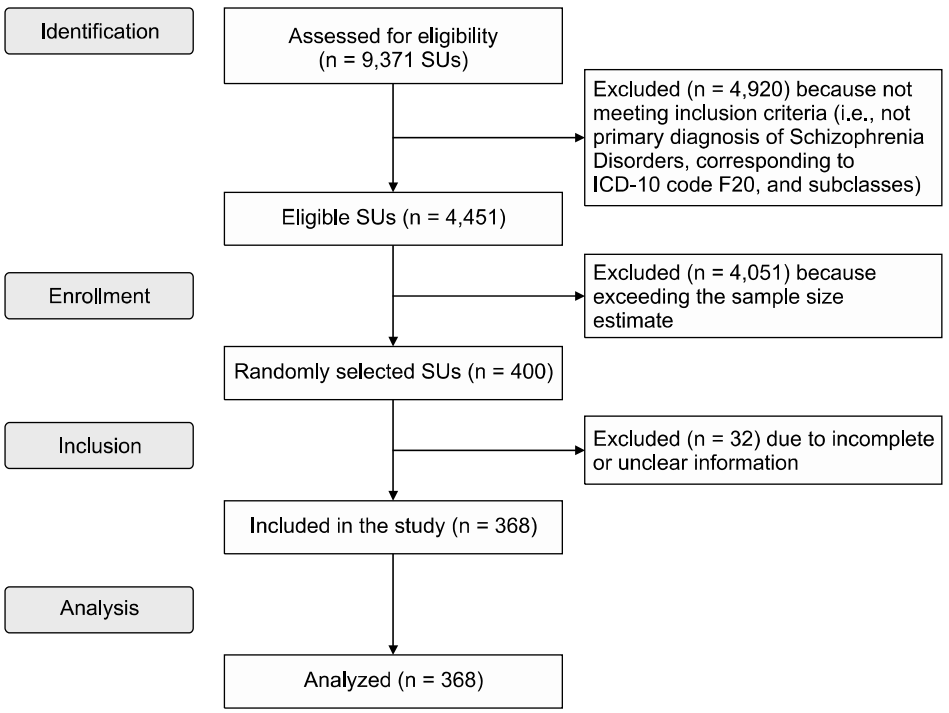

Fig. 1. STROBE flow chart. STROBE, Strengthening the Reporting of Observational Studies in Epidemiology; SUs, Community Mental Health Service Users (SUs).

servations in the training set, and n = 92 observations in the test set. Twenty possible predictors of the total antipsychotic dose were assessed. The model's hyperparameters were tuned to find the model with the highest prediction accuracy across ten (number of predictors/2) possible models. We used the "*train()*" function from the "*caret*" package to automate this process, employing "*tuneLenght=10*", and the function "*trainControl()*" to set repeated 5 × 5 folds cross-validation.

Prediction of APP

There were n = 277 observations in the train set, and n = 91 observations in the test set. The number of predictors of APP was 19. The model's hyperparameters were automatically tuned across ten possible models, employing the same procedure used for the prediction of dose.

Statement of Ethics

The study was approved by the local ethics committee on July 3, 2019 (Comitato Etico AVEN, reference 549/2019, protocol number 19133) and was conducted according to the Declaration of Helsinki, Good Clinical Practice principles for medical research, and current regulations relating to the protection and processing of personal and sensitive data (European Regulation n. 679/2016).

RESULTS

The final sample consisted of 368 clinical records, corresponding to 368 different SUs. Mean age was 52.9 years (SD = 12.8; median = 53 years); 56.5% (n = 208) were males. The majority of the sample were Italian citizens (n = 334, 92.8%) and married (n = 274, 79.4%); only 35.8% (n = 120) had at least 8 years of schooling (attended high school), and just 33.0% (n = 115) were employed. N = 137 (37.2%) SUs were taking APP, 185 (50.3%) received LAI medication, and 67 (18.2%) anticholinergics; the average dose of antipsychotics (expressed as chlorpromazine equivalent milligrams) was 238.1 mg (SD = 204.8; median = 181.5 mg). The most common follow-up recorded in the clinical records was blood exam monitoring (n = 143, 38.9%). Average days of admission in the previous 3 year was 6.4 (SD = 27.9 days), with the majority of the sample that were not admitted to psychiatric ward in the last year (n = 333, 90.5%). Mean number of community mental health outpatient contacts in the last year was 39.7 (SD = 59.6). Table 1 displays full description of the sample.

Prediction of Dose

As can be seen in Table 2, the best prediction perform-

ance was provided by RF (mean absolute error = 0.56; RMSE = 0.70; R^2 = 0.31). We presented the results of performance comparison as the difference in the parameters (i.e., the first-the second), along with the two-tailed test p value, with Bonferroni adjustment, to avoid Type I error due to multiple testing.

In addition to model performance indices, we regressed the predicted dose on the actual dose to understand which model yielded the strongest association. The results are displayed in Supplementary Table 1 (available online), highlighting best linear correlation yielded by RF predictions (unstandardized regression coefficient [B] = 1.01; 95% CI = 0.627−1.39; $p < 0.001$).

Accordingly, the RF model was fitted on the whole set of data (i.e., n = 368; 20 predictors). That RF model employed 500 iterations (i.e., number of trees), and randomly selected six variables at each split. To gain insight into the model, the 'importance' of each variable (essentially, its relative strength of predictive power) was calculated by dividing its prediction score over the maximum score (note that, due to this normalization, the importance score obtained for the most important variable is always 1, i.e., 100%). The relative importance plot is displayed in Figure 2. The most important predictor of the total antipsychotic dose was the APP, followed by the number of CMHC contacts, SU age, and by variables describing length of psychiatric hospitalizations, and the type of CMHC contacts in the last year. When APP was omitted from the set of predictors the relative order and size of the predictors is maintained, and the most important predictors became the total number of CMHC contacts in the last year (see Supplementary Fig. 1; available online).

The results of linear regression on the full set are displayed in Table 3. For increasing the readability of the coefficients, the outcome (dose, log-transformed) was multiplied by 10.

Table 1. Descriptive analysis of the sample

Variable	Value
Sex, male	208 (56.5)
Married	71 (20.6)
Employed	115 (33.0)
Schooling (> 8 yr)	120 (35.8)
Italian Citizenship	334 (92.8)
Antipsychotic medication (LAI)	185 (50.3)
Anticholinergic medication	67 (18.2)
APP	137 (37.2)
Therapy unchanged in the last year	168 (45.8)
Compulsory psychiatric hospitalizations in the last year	8 (2.2)
Number of psychiatric hospitalizations in the last year	
0	333 (90.5)
1	20 (5.4)
2	12 (3.3)
3	3 (0.8)
Recorded BMI (in the last year)	38 (10.3)
Recorded blood exams (in the last year)	143 (38.9)
Recorded ECG (in the last year)	124 (33.7)
Recorded side effects (in the last year)	110 (29.9)
Age (yr)	52.9 ± 12.8 (24−82)
Antipsychotic dose (mg chlorpromazine equivalence)	238.1 ± 204.8 (17−1,050)
Length of psychiatric hospitalization in the last year (day)	1.6 ± 6.2 (0−36)
Length of psychiatric hospitalization in the last 3 years (day)	6.4 ± 27.9 (0−65)
Overall community mental health center services in the last year	39.7 ± 59.6 (2−335)
Medical community mental health center services in the last year	8.0 ± 6.5 (0−32)
Urgent community mental health center services in the last year	0.7 ± 2.4 (0−9)

Values are presented as number (%) or mean ± standard deviation (range).
LAI, long-acting injectable antipsychotic; APP, antipsychotic polytherapy; BMI, body mass index; ECG, electrocardiogram.

Prediction of APP

As can be seen in Table 4, RF performed better than the

Table 2. Model's performance comparisons in the prediction of the total antipsychotic dose

Comparison	ΔMAE (p value)	ΔRMSE (p value)	Δ R^2 (p value)
LM vs. RF	0.032 (0.001)	0.073 (0.009)	−0.034 (0.014)
SVM vs. RF	0.060 (< 0.001)	0.125 (0.001)	−0.079 (< 0.001)
NB vs. RF	0.031 (0.001)	0.073 (0.010)	−0.034 (0.014)
KNN vs. RF	0.082 (< 0.001)	0.086 (< 0.001)	−0.181 (< 0.001)

MAE, mean absolute error; RMSE, root mean square error; LM, linear regression model; RF, random forest; SVM, supported vector machine; NB, Naïve Bayes; KNN, K-nearest neighborhood.
NB: A positive difference represents better performance by RF, on the contrary for the R^2.

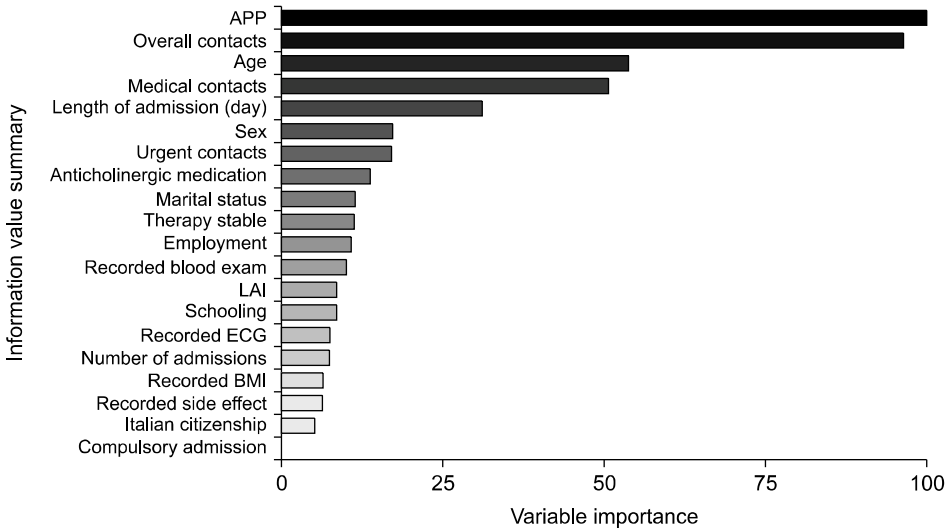

Fig. 2. Plot of variable importance of predictors of antipsychotic dose. ECG, electrocardiogram; LAI, long-acting injectable antipsychotics; BMI, body mass index; APP, antipsychotic polytherapy.

Table 3. Predictors of antipsychotic dose using multivariate linear regression model

Variable	Coefficient	95% CI	*p* value	Standardized beta coefficient
APP[a]	6.61	4.63 to 8.60	< 0.01**	0.417
Sex[a]	−2.36	−4.38 to −0.33	0.02*	−0.235
Length of admission (day)	0.02	−0.26 to 0.30	0.86	0.119
Overall community mental health contacts	0.01	−0.01 to 0.03	0.15	0.030
Compulsory admission	2.72	−4.08 to 9.52	0.43	0.026
Number of admissions	0.42	−3.26 to 4.09	0.82	0.054
Urgent community mental health contacts	0.23	−0.19 to 0.65	0.28	0.069
Marital status	0.58	−1.96 to 3.13	0.65	−0.011
Medical community mental health contacts	0.03	−0.14 to 0.20	0.70	0.080
Age	−0.05	−0.14 to 0.03	0.19	−0.059
Italian citizenship[a]	−6.83	−11.63 to −2.03	< 0.01**	−0.088
Anticholinergic medication[a]	2.48	0.10 to 4.87	0.04*	0.101
Recorded BMI	2.47	−0.60 to 5.55	0.11	0.013
Employment[a]	−3.45	−5.40 to −1.49	< 0.01**	−0.077
Therapy stable	1.35	−0.53 to 3.15	0.16	0.056
Recorded blood exam	0.48	−2.10 to 3.07	0.71	0.032
Recorded side effect	−1.06	−3.25 to 1.13	0.34	−0.001
LAI	−0.32	−2.18 to 1.54	0.73	−0.002
Schooling	0.59	−1.36 to 2.54	0.55	0.047
Recorded ECG	−0.48	−3.13 to 2.17	0.72	0.018

95% CI, 95% confidence interval; APP, antipsychotic polytherapy; BMI, body mass index; LAI, long-acting injectable antipsychotic; ECG, electrocardiogram.
Significant associations are highlighted using [a].
p values: *< 0.05; **< 0.01.

other models in terms of true positive rate and AUROC in predicting APP. Yet, the true negative rate resulted not significantly different from that of logistic regression, but lower than that of SMV, NB, and KNN. RF performance indices for the prediction of APP are: AUROC = 0.66, true positive rate = 0.41, and true negative rate = 0.78. We present the results of performance comparison as the difference in the parameters (i.e., the first-the second), along with the two-tailed test *p* value, with Bonferroni adjustment, to avoid Type I error due to multiple testing.

However, when testing if the overall accuracy is better than the "no information rate", which is taken to be the largest class percentage in the data (i.e., 63% antipsychotic monotherapy), the overall accuracy resulted

Table 4. Model's performance comparison in the prediction of APP

Comparison	ΔAUROC (p value)	ΔTrue positive rate (p value)	ΔTrue negative rate (p value)
LM vs. RF	−0.034 (0.044)	−0.078 (0.001)	0.037 (0.106)
SVM vs. RF	−0.048 (0.003)	−0.311 (< 0.001)	0.194 (< 0.001)
NB vs. RF	−0.035 (0.042)	−0.096 (< 0.001)	0.047 (0.045)
KNN vs. RF	−0.022 (0.022)	−0.125 (< 0.001)	0.114 (< 0.001)

APP, antipsychotic polytherapy; AUROC, area under the receiving operator curve; LM, logistic regression model; RF, random forest; SVM, supported vector machine; NB, Naïve Bayes; KNN, K-nearest neighborhood.
NB: A positive difference represents better performance by RF.

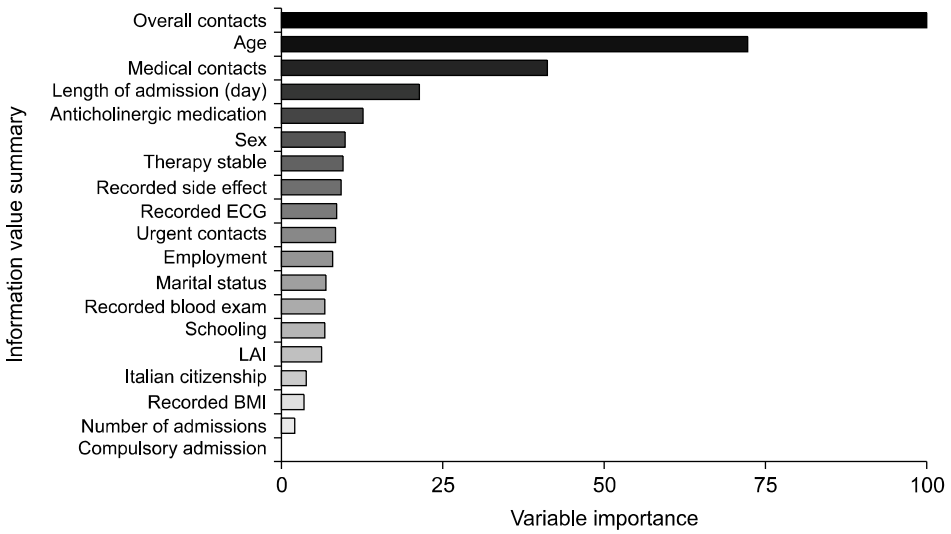

Fig. 3. Variable importance plot of the predictors of APP.
APP, antipsychotic polytherapy; BMI, body mass index; ECG, electrocardiogram; LAI, long-acting injectable antipsychotics.

not above of that baseline accuracy for all the models. The confusion matrix is displayed in Supplementary Table 2 (available online).

The RF model was then fitted to the complete dataset (n = 368; 19 predictors) to determine the relative importance of each variable in the prediction of APP. The importance plot is displayed in Figure 3. The final RF model employed 500 iterations (i.e., number of trees) and randomly selected 17 variables at each split.

The most important predictors of APP identified were: the number of CMHC contacts, the length of psychiatric hospitalization in the last year, receiving anticholinergic medication, the stability of the prescription in the last year, the age, and sex.

The results of logistic regression on the full set are displayed in Table 5, the significant associations are bolded.

DISCUSSION

Random Forest Performance

Our classification and regression application of RF confirmed that RF models can have higher prediction accuracy than corresponding parametric models such as logistic regression and linear regression [33]. As previously suggested, the main advantage of RF over the respective parametric model has been observed in managing non-linearity and collinearity issues [34]. This is an important strength of RF over a typical regression-type approach because it allows identifying which predictors of the outcome under study are most strongly related to it in a non-parametric way. Moreover, this avoids the risk of misspecifications and arbitrary categorizations of continuous variables with non-normal distribution, such as the overall antipsychotic dose or the length of hospitalization. Most of these features are common to other ML algorithms but RF yielded more accurate predictions also than the other ML models explored in this application.

Table 5. Odds ratios predictors of APP derived from logistic regression model

Variable	Odds ratio	95% CI	p value
LAI	1.16	0.59−2.27	0.67
Total community mental health contacts[a]	1.02	1.01−1.03	< 0.01**
Therapy stable	0.73	0.38−1.41	0.35
Recorded side effect	0.95	0.43−2.13	0.91
Medical community mental health contacts	1.03	0.96−1.09	0.42
Age[a]	1.06	1.03−1.09	< 0.01**
Recorded blood exam	1.81	0.70−4.68	0.22
Recorded ECG	0.91	0.35−2.39	0.85
Recorded BMI	0.68	0.21−2.23	0.53
Employment	0.76	0.37−1.57	0.47
Sex	0.59	0.28−1.24	0.16
Anticholinergic medication[a]	1.68	1.09−2.61	0.019*
Urgent community mental health contacts	1.05	0.91−1.21	0.55
Schooling	1.37	0.67−2.83	0.39
Length of admission (day)	1.06	0.96−1.17	0.28
Marital status	0.47	0.18−1.21	0.12
Number of admissions	0.31	0.07−1.28	0.11
Italian citizenship	3.94	0.69−22.5	0.12
Compulsory admission	2.92	0.001−32.3	0.38

APP, antipsychotic polytherapy; 95% CI, 95% confidence interval; LAI, long-acting injectable antipsychotic; ECG, electrocardiogram; BMI, body mass index.
Significant associations are highlighted using [a].
p values: *< 0.05; **< 0.01.

Clinical Outcomes

The aim of this study was to assess antipsychotics prescription patterns for people with schizophrenia, using ML techniques applied to real-world data to investigate predictors of antipsychotic dose and number of antipsychotics prescribed.

In our sample a relatively high proportion of patients were receiving APP (37.2%). The average chlorpromazine equivalent dose in the whole sample was relatively low, at 238.1 mg/day (the approved range is 200 to 1,000 mg/day [10]). The high rate of APP is not surprising, despite guidelines recommendations, as it reflects common practice widespread throughout the world [17,18,35,36]. This pattern reflects the fact that the clinical improvement brought by antipsychotics is often only partial. The presence of residual symptoms may drive the combined prescription of medication [10].

APP was the most important predictor of total antipsychotic dose in the RF and was also strongly associated with dose in the linear regression. Our findings suggest that APP is adjunctive rather than compensatory: effectively 'more' drug is added, not just different drugs. Other studies have highlighted a correlation between total dose and the number of antipsychotics delivered, since the addition of medications is very likely to increase the final dose of medication prescribed [11,36,37]. However, this correlation may also reflect the long-standing challenge in clinical settings of implementing an effective tapering strategy according to individual patient response and needs, leading to the stalling of the cross-titration process when switching between antipsychotics and resulting in *de facto* APP. In clinical practice, slow cross-titration is frequently observed because of the widely held notion that this method diminishes risk of symptom exacerbation and side effects, making this a safer approach. However, the evidence collected to date in this respect is not compelling [38] and studies addressing this issue are in progress [39,40].

Concerning total antipsychotic dose, other than APP, sex was found to be a strong predictor both in RF and in linear regression whereby males received higher doses than females. This result may be intuitively understood as deriving from the higher average body weight of males than females, leading the former to need more medication to reach a similar dose per kg ratio [36]. However, the current literature suggests a worse course of schizophrenia in males [41], consequently inducing the clinician to adopt a more intense pharmacological approach. Furthermore,

it has been also suggested that women might be more prone to develop side effects [42] and that they show a better response to lower doses due to the different hormonal milieu and, possibly, due to the relatively greater blood flow in the brain which delivers more drug to target receptors [43].

Length of hospital admissions in the previous year and other indicators of CMHC utilization (i.e., the number of CMHC contacts, both medical and urgent) were important predictors of total dose of antipsychotic in the RF but not in the linear model. The source of the inconsistency between the two methods may be attributable to the low number of cases (i.e., SUs admitted in the previous year) that makes difficult for regression model to detect a linear relation with the outcome. That could also reflect a difference in metrics. RF importance is normalized while OLS 'importance' is judged against a fixed standard of significance. However, a more severe hospital admission and CMHC utilization status is a likely indicator of a worse control of the psychotic symptoms, warranting higher antipsychotic doses for these patients. Furthermore, the pressure on the inpatient psychiatric unit and increased demand for hospital beds might induce clinicians to adopt prescription patterns resulting in higher dosages at discharge [36]. Being unemployed and receiving anticholinergic medication was significantly associated with the total antipsychotic dose. These results may be indicating a suboptimal antipsychotic prescription, leaving patients with residual symptoms or debilitating side effects, which need the prescription of anti-side effects medication. In this perspective, psychiatric rehabilitation (such as vocational rehabilitation programs) may represent a useful approach to facilitate reduction of psychotropic medication and foster recovery [44]. Alternatively, unemployment and the need for anticholinergics may be considered proxies of more severe psychopathology, requiring higher doses of antipsychotics to achieve a psychosocial recovery. Unfortunately, the cross-sectional design of the present study does not allow resolution of the causal direction of these associations.

Treatment with anticholinergics was found to be important predictor also of APP in the RF model and in the logistic regression, emphasizing the strong correlation between APP and higher antipsychotic dose. The severity of the symptoms and insufficient control by monotherapy could explain why a higher number of CMHC contacts was a predictor of APP both in the RF and in the linear regression models. According to the results of the present study, for each additional community mental health center contact per year, the odds of being prescribed APP rather than monotherapy are 2% higher, suggesting that frequent attenders at mental health services require more intensive psychiatric treatment due to their experiencing more severe psychopathology. However, some studies have highlighted the challenge for mental health professionals of dealing with frequent attenders. Polytherapy may result from attempts to address the problem by adding prescription to prescription, even though this practice diverges dramatically from guideline recommendations [45,46].

Finally, maintaining records of side effects, blood exams, ECG, and BMI in the clinical charts proved to be predictors of APP. These results are consistent with previous evidence suggesting that patients taking more complex pharmacotherapy usually suffer from other physical health conditions, being more vulnerable to more severe side effects, and requiring closer monitoring of physical parameters [45]. Overall, the increased monitoring to which SUs receiving APP are subject is in in accord with relevant guideline recommendations [8,10]. However, issues of patient compliance and satisfaction needs to be addressed, balancing prescribing for efficacy with side effect management, since both lack of efficacy and poor tolerability have been shown to be determinants of treatment discontinuation by patients [13,47].

Prospectively, ML techniques applied to the amount of data routinely produced by CMHCs (that may be conceived as "big data") may help to implement new models of community psychiatry in Italy. The Mental Health Department service model developed since 1978 is currently only able to partially respond to the increase need for mental health care, due to budget cuts, reduced personnel, and increased psychiatric disorders throughout Europe. This may particularly occur in times of economic crises or severe epidemics [48,49]. Also, the circumstances may be a cause of increased high-dose antipsychotic prescription, i.e. a "medicalizing response" to suffering. The adoption of a new approach based on big data analysis, ML techniques, digitalization and informatization of mental health care may represent an avenue of innovation for Italian Mental Health Departments, specifically for CMHCs [50,51].

Limitations

Some limitations of this study must be acknowledged. The use of cross-sectional design does not account whether some associations revealed are causes or consequences of higher antipsychotic dose, or APP. Nevertheless, we aimed to apply an independent, atheoretical approach, using real-world data to take a snapshot of what commonly occurs in the everyday clinical practice. RF models have been predominantly used in cross-sectional samples [34,52,53], and we performed direct comparisons with the classic parametric models to assess whether the outcomes from the two approaches offer different insights.

We relied on the evaluation of the most used performance indices for comparing RF with other ML algorithms and the parametric models, however, the prediction accuracy in our classification problem (i.e., APP prediction) resulted not different from random guessing for all the models. That is likely due to the imbalanced prevalence of APP within the sample.

Finally, only indirect indicators of the clinical state have been considered in the absence of a standardized assessment for the overall psychopathology or the effectiveness of APP and current antipsychotic dose. Future studies should particularly attempt to overcome this last issue, ideally implementing prospective evaluations, for a more comprehensive understanding of the prescription patterns for SUs living with schizophrenia.

■ Funding

None.

■ Acknowledgments

Authors wish to thank Flavia Baccari, Massimo Marcon, Samantha Perseguiti, Giuseppe Tibaldi for support in planning and conducting this study.

■ Conflicts of Interest

No potential conflict of interest relevant to this article was reported.

■ Author Contributions

Conceptualization: Mattia Marchi, Giacomo Galli, Gianluca Fiore, Gian M. Galeazzi. Data acquisition: Giacomo Galli, Gianluca Fiore. Formal analysis: Mattia Marchi, Andrew Mackinnon. Supervision: Gian M. Galeazzi, Fabrizio Starace. Writing original draft: Mattia Marchi, Giacomo Galli, Gianluca Fiore. Reviewing and editing: Giorgio Mattei, Gian M. Galeazzi, Andrew Mackinnon, Fabrizio Starace. All authors have contributed to and approved the final manuscript.

■ ORCID

Mattia Marchi https://orcid.org/0000-0003-2970-1276
Giacomo Galli https://orcid.org/0000-0002-9238-697X
Gianluca Fiore https://orcid.org/0000-0003-4647-1759
Andrew Mackinnon https://orcid.org/0000-0003-0831-9801
Giorgio Mattei https://orcid.org/0000-0001-5326-303X
Fabrizio Starace https://orcid.org/0000-0002-1123-8917
Gian M. Galeazzi https://orcid.org/0000-0003-2706-3362

REFERENCES

1. Seidman LJ, Mirsky AF. *Evolving notions of schizophrenia as a developmental neurocognitive disorder. J Int Neuropsychol Soc 2017;23:881-892.*
2. Guloksuz S, Pries LK, Delespaul P, Kenis G, Luykx JJ, Lin BD, *et al. Examining the independent and joint effects of molecular genetic liability and environmental exposures in schizophrenia: results from the EUGEI study. World Psychiatry 2019;18:173-182.*
3. Owen MJ, Sawa A, Mortensen PB. *Schizophrenia. Lancet 2016;388:86-97.*
4. van Os J, Kenis G, Rutten BP. *The environment and schizophrenia. Nature 2010;468:203-212.*
5. Perälä J, Suvisaari J, Saarni SI, Kuoppasalmi K, Isometsä E, Pirkola S, *et al. Lifetime prevalence of psychotic and bipolar I disorders in a general population. Arch Gen Psychiatry 2007;64:19-28.*
6. Tandon R, Keshavan MS, Nasrallah HA. *Schizophrenia, "just the facts" what we know in 2008. 2. Epidemiology and etiology. Schizophr Res 2008;102:1-18.*
7. Murray CJ, Vos T, Lozano R, Naghavi M, Flaxman AD, Michaud C, *et al. Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet 2012;380:2197-2223.*
8. National Institute for Health and Care Excellence. *Psychosis and schizophrenia in adults: prevention and management [Internet]. National Institute for Health and Care Excellence; 2014 Feb 12 [cited at 2020 Apr 6]. Available from: https://www.nice.org.uk/guidance/cg178/chapter/1-Recommendations#subsequent-acute-episodes-of-psychosis-or-schizophrenia-and-referral-in-crisis-2.*
9. National Collaborating Centre for Mental Health; National Institute for Health and Clinical Excellence. *Schizophrenia: core interventions in the treatment and management of schizophrenia in adults in primary and secondary care. London:Royal College of Psychiatrists;2010.*

10. Taylor DM, Barnes TRE, Young AH. *The Maudsley prescribing guidelines in psychiatry. Hoboken:Wiley Blackwell;2018.*
11. Roh D, Chang JG, Kim CH, Cho HS, An SK, Jung YC. *Antipsychotic polypharmacy and high-dose prescription in schizophrenia: a 5-year comparison. Aust N Z J Psychiatry 2014;48:52-60.*
12. Bruijnzeel DM, Tandon R. *Antipsychotic polypharmacy: State of the Science and Guidelines for Practice. It's difficult to stop once you start. Asian J Psychiatr 2018;33:A1-A2.*
13. Stahl SM, Sy S, Maguire GA. *How and when to treat the most common adverse effects of antipsychotics: expert review from research to clinical practice. Acta Psychiatr Scand 2021;143:172-180.*
14. Ray WA, Chung CP, Murray KT, Hall K, Stein CM. *Atypical antipsychotic drugs and the risk of sudden cardiac death. N Engl J Med 2009;360:225-235.*
15. Kamei H, Yamada H, Hatano M, Hanya M, Yamada S, Iwata N. *Effectiveness in switching from antipsychotic polypharmacy to monotherapy in patients with schizophrenia: a case series. Clin Psychopharmacol Neurosci 2020;18:159-163.*
16. Correll CU, Rummel-Kluge C, Corves C, Kane JM, Leucht S. *Antipsychotic combinations vs monotherapy in schizophrenia: a meta-analysis of randomized controlled trials. Schizophr Bull 2009;35:443-457.*
17. Li Q, Xiang YT, Su YA, Shu L, Yu X, Chiu HF, *et al. Antipsychotic polypharmacy in schizophrenia patients in China and its association with treatment satisfaction and quality of life: findings of the third national survey on use of psychotropic medications in China. Aust N Z J Psychiatry 2015;49:129-136.*
18. Qiu H, He Y, Zhang Y, He M, Liu J, Chi R, *et al. Antipsychotic polypharmacy in the treatment of schizophrenia in China and Japan. Aust N Z J Psychiatry 2018;52:1202-1212.*
19. Kang SG, Cho SE, Na KS, Pae CU, Cho SJ. *Clinical usefulness of amisulpride add-on therapy in schizophrenia patients without treatment response to second-generation Antipsychotics. Clin Psychopharmacol Neurosci 2021;19:117-124.*
20. Lee JS, Yun JY, Kang SH, Lee SJ, Choi JH, Nam B, *et al. Korean Medication Algorithm for Schizophrenia 2019, second revision: treatment of psychotic symptoms. Clin Psychopharmacol Neurosci 2020;18:386-394.*
21. Cao B, Cho RY, Chen D, Xiu M, Wang L, Soares JC, *et al. Treatment response prediction and individualized identification of first-episode drug-naïve schizophrenia using brain functional connectivity. Mol Psychiatry 2020;25:906-913.*
22. Ambrosen KS, Skjerbæk MW, Foldager J, Axelsen MC, Bak N, Arvastson L, *et al. A machine-learning framework for robust and reliable prediction of short- and long-term treatment response in initially antipsychotic-naïve schizophrenia patients based on multimodal neuropsychiatric data. Transl Psychiatry 2020;10:276.*
23. Li Y, Zhang L, Zhang Y, Wen H, Huang J, Shen Y, *et al. A random forest model for predicting social functional improvement in Chinese patients with schizophrenia after 3 months of atypical antipsychotic monopharmacy: a cohort study. Neuropsychiatr Dis Treat 2021;17:847-857.*
24. Kelleher JD, Namee BM, D'Arcy A. *Fundamentals of machine learning for predictive data analytics: algorithms, worked examples, and case studies. Cambridge:MIT Press;2015. 624 p.*
25. Uddin S, Khan A, Hossain ME, Moni MA. *Comparing different supervised machine learning algorithms for disease prediction. BMC Med Inform Decis Mak 2019;19:281.*
26. Tai AMY, Albuquerque A, Carmona NE, Subramanieapillai M, Cha DS, Sheko M, *et al. Machine learning and big data: implications for disease modeling and therapeutic discovery in psychiatry. Artif Intell Med 2019;99:101704.*
27. World Health Organization. *ICD-10: international statistical classification of diseases and related health problems: tenth revision [Internet]. Geneva: World Health Organization; 2004 [cited at 2020 Nov 2]. Available from: https://apps.who.int/iris/handle/10665/42980.*
28. Arya R, Antonisamy B, Kumar S. *Sample size estimation in prevalence studies. Indian J Pediatr 2012;79:1482-1488.*
29. Lwanga SK, Lemeshow S, World Health Organization. *Sample size determination in health studies: a practical manual [Internet]. Geneva: World Health Organization; 1991 [cited at 2021 Aug 25]. Available from: https://apps.who.int/iris/handle/10665/40062.*
30. Leucht S, Samara M, Heres S, Davis JM. *Dose equivalents for antipsychotic drugs: the DDD method. Schizophr Bull 2016;42 Suppl 1:S90-S94.*
31. RStudio Team. *RStudio: Integrated Development Environment (IDE) for R. [Internet]. Boston: RStudio Team; 2021 [cited at 2021 Aug 10]. Available from: https://www.rstudio.com/products/rstudio/.*
32. Kuhn M. *Building predictive models in R using the caret package. J Stat Softw 2008;28:1-26.*
33. Breiman L. *Random forests. Mach Learn 2001;45:5-32.*
34. Schonlau M, Zou RY. *The random forest algorithm for statistical learning. Stata J 2020;20:3-29.*
35. Harrington M, Lelliott P, Paton C, Okocha C, Duffett R, Sensky T. *The results of a multi-centre audit of the prescribing of antipsychotic drugs for in-patients in the UK. Psychiatr Bull 2002;26:414-418.*
36. McMillan SS, Jacobs S, Wilson L, Theodoros T, Robinson G, Anderson C, *et al. Antipsychotic prescribing for vulnerable populations: a clinical audit at an acute Australian mental health unit at two-time points. BMC Psychiatry 2017;17:139.*
37. Sim K, Su HC, Fujii S, Yang SY, Chong MY, Ungvari G, *et al. High-dose antipsychotic use in schizophrenia: a comparison between the 2001 and 2004 Research on East Asia Psychotropic Prescription (REAP) studies. Br J Clin Pharmacol 2009;67:110-117.*
38. Takeuchi H, Kantor N, Uchida H, Suzuki T, Remington G. *Immediate vs gradual discontinuation in antipsychotic switching: a systematic review and meta-analysis. Schizophr Bull*

2017;43:862-871.

39. Begemann MJH, Thompson IA, Veling W, Gangadin SS, Geraets CNW, van 't Hag E, *et al. To continue or not to continue? Antipsychotic medication maintenance versus dose-reduction/discontinuation in first episode psychosis: HAMLETT, a pragmatic multicenter single-blind randomized controlled trial. Trials 2020;21:147*.
40. Moncrieff J, Lewis G, Freemantle N, Johnson S, Barnes TRE, Morant N, *et al. Randomised controlled trial of gradual antipsychotic reduction and discontinuation in people with schizophrenia and related disorders: the RADAR trial (Research into Antipsychotic Discontinuation and Reduction). BMJ Open 2019;9:e030912*.
41. Ayesa-Arriola R, de la Foz VO, Setién-Suero E, Ramírez-Bonilla ML, Suárez-Pinilla P, Son JM, *et al. Understanding sex differences in long-term outcomes after a first episode of psychosis. NPJ Schizophr 2020;6:33*.
42. Seeman MV. *Women who suffer from schizophrenia: critical issues. World J Psychiatry 2018;8:125-136*.
43. Gur RE, Gur RC. *Gender differences in regional cerebral blood flow. Schizophr Bull 1990;16:247-254*.
44. Mattei G, Sacchi V, Alfieri S, Bisi A, Colombini N, Ferrari S, *et al. Stakeholders' views on vocational rehabilitation programs: a call for collaboration with Occupational Health Physicians. Med Lav 2018;109:201-209*.
45. Norton J, Oude Engberink A, Gandubert C, Macgregor A, David M, Mann A, *et al. Erratum to 'Frequent attendance and the concordance between PHQ screening and GP assessment in the detection of common mental disorders' [Journal of Psychosomatic Research, Vol. 110 (2018) 1-10]. J Psychosom Res 2019;118:17*.
46. Beck A, Sanchez-Walker E, Evans LJ, Harris V, Pegler R, Cross S. *Frequent users of Mental Health Liaison Services within emergency departments. Psychiatry Res 2017;258:194-199*.
47. Naber D, Lambert M. *The CATIE and CUtLASS studies in schizophrenia: results and implications for clinicians. CNS Drugs 2009;23:649-659*.
48. Frasquilho D, Matos MG, Salonna F, Guerreiro D, Storti CC, Gaspar T, *et al. Mental health outcomes in times of economic recession: a systematic literature review. BMC Public Health 2016;16:115*.
49. Margerison-Zilko C, Goldman-Mellor S, Falconi A, Downing J. *Health impacts of the Great Recession: a critical review. Curr Epidemiol Rep 2016;3:81-91*.
50. Eurostat. *People at risk of poverty or social exclusion- statistics explained [Internet]. Eurostat; 2017 [cited at 2021 Mar 4]. Available from: https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Archive:People_at_risk_of_poverty_or_social_exclusion.*
51. Mattei G. *Fashion industry as a source of inspiration for the 'Mental Health Department 4.0': an overview. J Psychopathol 2019;25:183-189*.
52. Kanerva N, Kontto J, Erkkola M, Nevalainen J, Männistö S. *Suitability of random forest analysis for epidemiological research: Exploring sociodemographic and lifestyle-related risk factors of overweight in a cross-sectional design. Scand J Public Health 2018;46:557-564*.
53. Kimura N, Aso Y, Yabuuchi K, Ishibashi M, Hori D, Sasaki Y, *et al. Modifiable lifestyle factors and cognitive function in older people: a cross-sectional observational study. Front Neurol 2019;10:401*.