Unraveling the Influence of Age, IQ, Education, and Negative Symptoms on Neurocognitive Performance in Schizophrenia: A Conditional Inference Tree Analysis

Authors
Xenia M. Hart[1], Yasue Mitsukura[2], Robert R. Bies[3], Hiroyuki Uchida[1]

Affiliations
1 Department of Neuropsychiatry, Keio University School of Medicine, Tokyo, Japan
2 Department of System Design Engineering, Faculty of Science and Technology, Keio University Japan
3 Department of Pharmaceutical Sciences, University at Buffalo, Buffalo, NY, USA

Keywords
schizophrenia, dopamine antagonists, cognitive impairment, cognition, CATIE

received 22.08.2023
revised 23.01.2024
accepted 25.01.2024
published online 22.02.2024

Bibliography
Pharmacopsychiatry 2024; 57: 53–60
DOI 10.1055/a-2258-0379
ISSN 0176-3679

Georg Thieme Verlag, Rüdigerstraße 14,
70469 Stuttgart, Germany

Correspondence
Dr. Xenia M. Hart
Department of Neuropsychiatry, Keio University School of Medicine
160-8582 Tokyo
Japan
info@xeniahart.de

Supplementary Material is available under https://doi.org/10.1055/a-2258-0379

ABSTRACT

Introduction The complex nature of neurocognitive impairment in schizophrenia has been discussed in light of the mixed effects of antipsychotic drugs, psychotic symptoms, dopamine D_2 receptor blockade, and intelligence quotient (IQ). These factors have not been thoroughly examined before.

Methods This study conducted a comprehensive re-analysis of the CATIE data using machine learning techniques, in particular Conditional Inference Tree (CTREE) analysis, to investigate associations between neurocognitive functions and moderating factors such as estimated trough dopamine D_2 receptor blockade with risperidone, olanzapine, or ziprasidone, Positive and Negative Syndrome Scale (PANSS), and baseline IQ in 573 patients with schizophrenia.

Results The study reveals that *IQ, age*, and *education* consistently emerge as significant predictors across all neurocognitive domains. Furthermore, higher severity of *PANSS-negative symptoms* was associated with lower cognitive performance scores in several domains. CTREE analysis, in combination with a genetic algorithm approach, has been identified as particularly insightful for illustrating complex interactions between variables. Lower neurocognitive function was associated with factors such as age > 52 years, IQ < 94/95, < 12/13 education years, and more pronounced negative symptoms (score < 26).

Conclusions These findings emphasize the multifaceted nature of neurocognitive functioning in patients with schizophrenia, with the PANSS-negative score being an important predictor. This gives rise to a role in addressing negative symptoms as a therapeutic objective for enhancing cognitive impairments in these patients. Further research must examine nonlinear relationships among various moderating factors identified in this work, especially the role of D_2 occupancy.

Introduction

Neurocognitive impairment is considered one of the core features among patients with schizophrenia and is attributable to multiple causes. A blockade of dopamine D_2 receptors above approximately 80 % with antipsychotics could impair neurocognitive function, including overall neurocognitive function and vigilance [1]. Antipsychotic drugs have been associated with mixed results in terms of their effects on neurocognitive impairment due to the illness [2, 3]. The presence of psychotic symptoms has also been reported to be associated with cognitive impairment. In one longitudinal follow-

up study of patients with first-episode schizophrenia, a decrease in positive symptoms was related to improvements in neurocognitive functions, including executive function, spatial memory, concentration/speed, and global cognition [4]. Moreover, baseline and current intelligence quotient (IQ) are reported to affect neurocognitive function in patients with schizophrenia [5, 6]. Thus, neurocognitive impairment in patients with schizophrenia needs to be comprehensively interpreted from multiple angles. To the best of our knowledge, no study has investigated the associations of neurocognitive function, dopamine D_2 receptor blockade with antipsychotics, illness severity, and baseline IQ in a comprehensive manner. The Clinical Antipsychotic Trials of Intervention Effectiveness (CATIE) [7] provides the ideal dataset for such analyses since its dataset includes a large number of subjects with scores of symptomatology and neurocognitive functions as well as plasma antipsychotics concentrations that can be used to estimate dopamine D_2 receptor blockade of antipsychotics in the brain [8–10]. To further highlight this important issue, we re-analyzed the CATIE data using a machine learning technique to examine if the neurocognitive functions could be classified based on moderating factors such as dopamine D_2 receptor blockade with antipsychotics, illness severity, and baseline IQ in patients with schizophrenia.

Methods

Study Population, Assessments, and Study Design

The CATIE trial was funded by the National Institute of Mental Health to compare the effectiveness of five antipsychotic drugs in patients with schizophrenia [7]. In the present study, we used data from the CATIE study from subjects who received risperidone, olanzapine, or ziprasidone treatment. Demographic and study population characteristics are summarized in ▶ **Table 1**.

Overall, we included subjects who completed assessments for neurocognitive function and psychopathology and provided plasma samples for the assessment of plasma antipsychotic concentrations. The present study sample was chosen due to previous works that have already established nonlinear mixed-effect models for this sample [8–10].

The neurocognitive function composite scores, furthermore referred to as domain scores, were calculated from the z score of the average of the following five standardized domain scores at month two: verbal memory (the Hopkins Verbal Learning Test, N = 426), vigilance (the Continuous Performance Test, N = 389), processing speed (the Grooved Pegboard, and the Revised Wechsler Adult Intelligence Scale Digit Symbol Test, N = 427), reasoning (the Wisconsin Card Sorting Test and the Revised Wechsler Intelligence Scale for Children Mazes, N = 427), and working memory (the Letter-number test of auditory working memory and a computerized test of visuospatial working memory, N = 427). In case of missing values, we have used the available domain scores.

Nine patient-specific moderating factors were selected for further analysis, which were age, baseline IQ, years of education, Positive and Negative Syndrome Scale (PANSS) positive score, PANSS negative score, mean Simpson-Angus Scale (SAS), anticholinergic drug use, drug type (risperidone, ziprasidone or olanzapine), and estimated minimum dopamine D_2 receptor blockade. Baseline IQ was measured using the Wide Range Achievement Test-third version (WRAT-3) at baseline in the CATIE trial. PANSS scores were assessed after month one. Model-predicted trough values of plasma concentrations of antipsychotics were used to calculate the estimated minimum dopamine D_2 receptor blockade levels on the day of neurocognitive assessment by using a previously reported model [11]. During the CATIE trial, drug concentrations of risperidone plus 9-hydroxyrisperidone (active moiety), olanzapine, or ziprasidone were measured at multiple time points. Plasma antipsychotic concentrations at the trough were calculated on the day of the neurocognitive assessment using established population pharmacokinetic models and extracting the Empirical Bayes Estimates for the pharmacokinetics parameters [12, 13]. Thereafter, dopamine D_2 receptor blockade levels were estimated by incorporating the predicted plasma concentration of risperidone plus 9-hydroxyrisperidone, olanzapine, or ziprasidone in the following one-site binding model:

$$\text{Blockade}\ (\%) = a \times [\text{plasma level}/(\text{plasma level} + ED_{50})] \quad (1)$$

Here, 'a' represents the maximum receptor blockade attributable to the antipsychotic drug and 'ED_{50}' is the estimated plasma level of antipsychotic drug associated with half of the maximum receptor blockade (Risperidone plus 9-hydroxyrisperidone: $a = 88.0\,\%$, $ED_{50} = 4.9$ ng/mL; olanzapine: $a = 90.7\,\%$, $ED_{50} = 7.1$ ng/mL; and ziprasidone: $a = 88.2\,\%$, $ED_{50} = 32.9$ ng/mL) [11]. The accuracy of these predicted models was previously confirmed with 32 clinically stable outpatients with schizophrenia, as the mean (95 % confidence interval) prediction errors for the prediction of D_2 blockade were 0.64 % (−6.18 to 7.46) for risperidone and −1.76 % (−5.22 to 1.58) for olanzapine [14].

Statistical Methods

Feature selection

Feature selection was performed on the dataset consisting of one continuous target variable and nine feature variables, where there was no apparent linear correlation between the features and the outcome. To obtain a comprehensive perspective on the importance of predictors in our dataset, we utilized a three-pronged approach for feature selection that compared the outcomes of genetic algorithm (GA), random forest (RF) feature importance, and recursive feature elimination (RFE). The GA utilizes a binary string to represent each potential solution, where each position in the string corresponds to a predictor variable. The mean squared error (MSE) of an RF model was utilized to determine the fitness of each individual, taking into account the predictors represented by the binary string. GA optimization was executed for 100 generations, with a population size of 50. Convergence of the GA over time was visualized by plotting the best fitness values across generations. The second approach involved employing the built-in feature importance measure of RF. An initial RF model was trained on the complete dataset, where the feature ranking was determined by the "IncNodePurity" measure. This measure indicates the overall reduction in node impurity, measured by the Gini index, from splitting on a variable, averaged over 500 trees. After selecting features with importance values greater than zero, the final RF model was trained. Our third step involved using the RFE algorithm, which eliminates features with the least importance in a sequential manner. RFE was conducted using RF as the base model, with a repeat-

▶ **Table 1** Demographic and clinical characteristics of the study sample.

Characteristics	Sample population (N = 427)	Total sample (N = 573)*
Age, years, *mean ± SD (range)*	41.3 ± 10.6 (18–66)	41.1 ± 10.8 (18–66)
Male, *n (%)*	317 (74.2)	413 (72.5)
Ethnicity, *n (%)*		
White	262 (61.4)	352 (61.8)
Others	164 (38.4)	218 (38.3)
Duration of education, years, *mean ± SD (range)*	12.3 ± 2.0 (3–21)	12.1 ± 2.2 (3–21)
Duration of treatment, years, mean ± SD (*range*)	16.9 ± 11.1 (0–52)	16.6 ± 11.4 (0–56)
Use of anticholinergics, *n (%)*	74 (17.3)	95 (16.7)
PANSS, *mean ± SD (range)*		
Total score	69.3 ± 18.1 (32–131)	70.2 ± 18 (32–131)
Positive score	16.6 ± 5.6 (7–35)	16.6 ± 5.5 (7–35)
Negative score	18.9 ± 6.4 (7–38)	19.3 ± 6.4 (7–38)
SAS mean score, *mean ± SD (range)*	0.2 ± 0.3 (0–1.8)	0.2 ± 0.3 (0–1.8)
IQ, *mean ± SD (range)*	89.6 ± 18.0 (44–125)	89.7 ± 17.9 (44–125)
Antipsychotics		
Risperidone, *n (%)*	162 (37.9)	214 (37.5)
Trough plasma level, *mean ± SD (range)*	24.9 ± 15.5 (2.8–90.2)	24.9 ± 15.6 (2.8–90.2)
Estimated D_2 occupancy, *mean ± SD (range)*	70.4 ± 8.0 (40.9–83.4)	70.4 ± 7.9 (40.9–83.4)
Olanzapine, *n (%)*	186 (43.6)	246 (43.2)
Trough plasma level, mean ± SD (range)	32.2 ± 19.5 (7.0–119.9)	32.2 ± 18.9 (6.6–119.9)
Estimated D_2 occupancy, mean ± SD (range)	70.5 ± 9.0 (44.9–85.6)	70.6 ± 8.9 (43.8–85.6)
Ziprasidone, *n (%)*	79 (18.5)	110 (19.3)
Trough plasma level, mean ± SD (range)	50.1 ± 35.9 (8.2–228.2)	47.0 ± 33.5 (5.1–228.2)
Estimated D_2 occupancy, mean ± SD (range)	48.3 ± 13.2 (17.5–77.1)	47.1 ± 13.2 (11.8–77.1
Use of anticholinergics, *n (%)*	74 (17.3)	95 (16.7)
PANSS, *mean ± SD (range)*		
Total score	69.3 ± 18.1 (32–131)	70.2 ± 18 (32–131)
Positive score	16.6 ± 5.6 (7–35)	16.6 ± 5.5 (7–35)
Negative score	18.9 ± 6.4 (7–38)	19.3 ± 6.4 (7–38)
SAS mean score, *mean ± SD (range)*	0.2 ± 0.3 (0–1.8)	0.2 ± 0.3 (0–1.8)
IQ, *mean ± SD (range)*	89.6 ± 18.0 (44–125)	89.7 ± 17.9 (44–125)
Neurocognitive score, *mean ± SD (range)*		
Verbal memory (N = 426)	0.1 ± 1 (−2.4–2.8)	
Processing speed (N = 427)	0.1 ± 0.9 (−2.6–3.1)	
Working memory (N = 427)	0.2 ± 0.9 (−2.8–2.0)	
Reasoning (N = 427)	0.2 ± 0.9 (−2.4–2.2)	
Vigilance (N = 389)	0.2 ± 1 (−2.8–3.3)	
Neurocognitive summary score (N = 427)	0.2 ± 1 (−2.6–2.9)	

*IQ: N = 535; PANSS: N = 526; SAS: N = 527; estimated D_2 occupancy & trough plasma level: N = 528; IQ, intelligence quotient; PANSS, Positive and Negative Syndrome Scale; SD, standard deviation.

ed cross-validation method. Initially, the complete set of predictors was employed, and then, each predictor was systematically eliminated one at a time based on their importance scores. Computation and visualization of variable importance for the selected features in the final RF models from each method were undertaken to provide additional insights into the relative importance of the features. Each feature selection method utilized a 10-fold cross-validation. The cross-validation was integrated within the feature selection process in both the GA and RFE. The final model of the RF feature importance approach was validated using cross-validation to obtain an unbiased estimate of the model prediction error. By doing so, we mitigate the risk of overfitting, provide a more unbiased estimate of model performance, and ensure that our feature selection process was not overly optimistic towards our training data. Subsequently, an exploratory analysis excluding two of the main predictors, *IQ at baseline and age*, from the data was performed, and each process was repeated using the adapted dataset. Statistical analyses were performed using R version 4.2.3 [15]. The R packages 'caret' and 'randomForest' were used for model training and feature importance, 'GA' was used for feature selection via a genetic algorithm, 'Metrics' for calculating MSE, and 'ggplot2' for data visualization.

After conducting feature selection to streamline our predictive variables, we implemented a Conditional Inference Tree (CTREE) analysis and subsequently performed a comprehensive performance evaluation to assess the effectiveness and accuracy of the model. CTREE analysis was performed to construct a predictive decision tree for the neurocognitive summary score as well as for the five domain scores based on *age, baseline IQ, years of education, PANSS positive score, PANSS negative score, the mean SAS, anticholinergic drug use, drug type (risperidone, ziprasidone or olanzapine), and estimated minimum dopamine D_2 receptor blockade*. The normal distribution of these values was tested by the Kolmogorov-Smirnov test for normality. The linearity between variables and outcomes was evaluated by producing plots for visual inspection, which allowed for a direct examination of their relationship patterns. A correlation matrix was constructed to evaluate multivariate correlations.

Conditional Inference Tree analysis

CTREE analysis, which represents a non-parametric class of decision tree analysis, was applied using the 'caret' [16] and 'partykit' [17] libraries in R to establish predictive models on our dataset. CTREE is unbiased in variable selection and capable of handling both numerical and categorical data. The CTREE algorithm selects cut-offs for decision tree splits by performing statistical significance tests at each node. It identifies the most strongly associated feature with the target variable and uses a permutation test to determine the optimal binary split, adjusting for multiple testing to avoid overfitting. This process continues recursively, with splits made based on a significance level set at $p < 0.05$ until no significant association can be found or other stopping criteria are met. Specifically, cut-offs for the CTREE models in our analysis were determined using the ctree2 method in conjunction with the 'caret' package for hyperparameter tuning, facilitating the selection of optimal parameters for the decision trees. Our model tuning grid, defined using the 'expand.grid' function, comprised of different depths

ranging from 1 to 10 and mincriterion values ranging from 0.1 to 1. These parameters were applied to the 'ctree2' method, a variant of CTREE, to train our models. We aimed to maintain a model that was sufficiently complex to capture the necessary relationships in the data while remaining interpretable. The CTREE algorithm aids in this by stopping growth when the addition of another split does not significantly increase the fit of the model. The robustness of our selected cut-offs was further assessed through a 10-fold cross-validation approach during the training phase using the 'trainControl' function from 'caret' [16], which helped confirm the consistency of the results. The reported (optimal) cut-offs were then selected based on their performance in the cross-validation, with a focus on achieving the lowest possible MSE and root mean squared error (RMSE), indicative of the best model fit.

Four distinct models were trained to evaluate the effect of feature selection methods on the performance of the CTREEs. The first model utilized all available features in the dataset. The next three models employed different feature selection methods: GA, RF, and RFE. In the context of predictive modeling, these strategies are frequently employed to enhance both model interpretability and performance. To achieve this, the techniques aim to reduce dimensionality, mitigate overfitting, and improve generalization. Model performance was assessed using the MSE and RMSE, with lower values of MSE and RMSE indicating better model fit. Comparing MSE and RMSE values across different models trained on the same data set can inform about relative model performances. A model with the lowest MSE or RMSE was considered the most effective and accurate. Features in the model were considered statistically significant if the p-value was less than 0.05.

Results

Feature selection

Our findings show a substantial impact of IQ, age, and education years on all assessed neurocognitive scores across all applied feature selection methods. A detailed summary of the features selected by each method, including the MSE for each of the five neurocognitive domains and the neurocognitive summary score, is given in **supplementary Table S3**. Significant variations exist in the choice of additional variables between the methods. After removing the most important features, IQ and age, years of education remained the most significant feature, followed by the PANSS negative score, the estimated dopamine D_2 receptor blockade, and the PANSS positive score. Selected features of minor importance were the use of anticholinergic comedication, the type of the administered drug, and the mean SAS score. As expected, the MSE values, which are indicative of the model's performance, increased when the features of age and IQ were removed, thus confirming their critical role in predicting neurocognitive outcomes. GA, in general, showed better performance (lowest MSE score) when compared to the RF and the RFE approach (i. e., summary score MSE = 0.623 (GA); 0.645 (RFE); 0.646 (RF)).

Conditional Inference Tree analysis

Our analysis yielded valuable insights into the performance of CTREE models. IQ is consistently presented as a significant factor in nearly all domains and models. The impact of age was prominent in most domains. Yet, as expected, education years and the PANSS negative score were the most prevalent without these two key predictors, emphasizing their potential influence on neurocognitive scores. Detailed information on model performance (MSE, RSME) and included features are presented in ▶ **Table 2**. Also, in this analysis, models based on the GA approach, showed better performance when compared to the RF and RFE approach, and the models without previous feature selection. The final tree for the neurocognitive summary score is presented in ▶ **Fig. 1**. For each cognitive domain, the included features and final model performances of the GA-based model were as follows (Roman letters indicating tree node structure; only nodes $p < 0.05$ reported; number of participants and mean ± SD z-score for each node in brackets):

a. Verbal memory (N = 426, MSE = 0.96115 and RMSE = 0.97712):
   I. IQ ≤ 94 (N = 234; 0.225 ± 0.937) vs. IQ > 94 (N = 192; 0.458 ± 0.999), $p < 0.001$
b. Processing speed (N = 427, MSE = 0.71158 and RMSE = 0.83753):
   I. IQ ≤ 95 (N = 235; −0.234 ± 0.902) vs. IQ > 95 (N = 192; 0.433 ± 0.865), $p < 0.001$
   II. Age ≤ 41 (N = 98; 0.116 ± 0.784) vs. IQ > 41 (N = 137; −0.484 ± 0.899), $p < 0.001$
   III. Age ≤ 47 (N = 137; 0.650 ± 0.831) vs. age > 47 (N = 55; −0.110 ± 0.695), $p < 0.001$
c. Working memory (N = 427, MSE = 0.70179 and RMSE = 0.83417):
   I. IQ ≤ 94 (N = 234; −0.140 ± 0.888) vs. IQ > 94 (N = 193; 0.543 ± 0.807), $p < 0.001$
   II. Age ≤ 52 (N = 203; −0.030 ± 0.845) vs. age > 52 (N = 31; −0.856 ± 0.839), $p < 0.001$
   I. PANSS neg. ≤ 30 (N = 185; 0.581 ± 0.778) vs. PANSS neg. > 30 (N = 8; −0.341 ± 1.005), $p = 0.042$
   I. Age ≤ 49 (N = 148; 0.696 ± 0.683) vs. age > 49 (N = 37; 0.120 ± 0.956), $p = 0.022$
d. Reasoning (N = 427, MSE = 0.73589 and RMSE = 0.85346):
   I. Age ≤ 41 (N = 195; 0.535 ± 0.800) vs. age > 41 (N = 232; −0.126 ± 0.939), $p < 0.001$
   II. IQ ≤ 105 (N = 154; 0.387 ± 0.790) vs. IQ > 105 (N = 41; 1.092 ± 0.563), $p < 0.001$
   III. IQ ≤ 88 (N = 111; −0.443 ± 0.909) vs. IQ > 88 (N = 121; 0.164 ± 0.874), $p < 0.001$
   IV. IQ ≤ 64 (N = 20; −0.124 ± 0.937) vs. IQ > 64 (N = 134; 0.464 ± 0.739), $p = 0.042$
   V. Age ≤ 54 (N = 92; −0.314 ± 0.830) vs. age > 54 (N = 19; −1.064 ± 1.037), $p = 0.029$
e. Vigilance (N = 389, MSE = 0.84388 and RMSE = 0.91316):
   I. IQ ≤ 90 (N = 181; −0.139 ± 0.899) vs. IQ > 90 (N = 208; 0.488 ± 0.974), $p < 0.001$
   II. Age ≤ 52 (N = 158; −0.014 ± 0.868) vs. age > 52 (N = 23; −0.994 ± 0.603), $p = 0.007$
f. Neurocognitive Summary Score (N = 427, MSE = 0.67609 and RMSE = 0.81961):
   I. IQ ≤ 95 (N = 235; −0.196 ± 0.901) vs. IQ > 95 (N = 192; 0.637 ± 0.855), $p < 0.001$
   II. Age ≤ 52 (N = 204; −0.06 ± 0.845) vs. age > 52 (N = 31; −1.089 ± 0.736), $p < 0.001$

I. Age ≤ 47 (N = 137; 0.821 ± 0.792) vs. age > 47 (N = 55; 0.178 ± 0.839), $p < 0.001$

▶ **Table 2** An overview of CTREE analysis results.

Neurocognitive score	Method	Included features ($p < 0.05$)	MSE	RMSE	Included features without IQ and age ($p < 0.05$)	MSE	RMSE
Verbal memory	–	IQ	0.95578	0.97572	Education years, PANSS negative	0.96967	0.97963
	GA	IQ	0.96115	0.97712	–	–	–
	RFE	IQ	0.95358	0.97330	Education years	**0.96784**	**0.98077**
	RF	IQ	**0.94401**	**0.96651**	Education years	0.98277	0.98889
Processing speed	–	IQ, age, education years, PANSS negative	0.74602	0.86022	Education years, PANSS negative	0.81257	0.89785
	GA	IQ, age	0.71158	0.83753	Education years, PANSS negative	0.81843	0.89757
	RFE	IQ, age, education years, PANSS negative	0.72371	0.84485	Education years, PANSS negative	**0.80783**	**0.89257**
	RF	IQ, age, education years	**0.70018**	**0.82931**	Education years, PANSS negative	0.81342	0.89790
Working memory	–	IQ, age	0.71445	0.84385	Education years	0.79657	0.89079
	GA	IQ, age, PANSS negative	**0.70179**	**0.83417**	–	–	–
	RFE	IQ, age, PANSS negative	0.72460	0.84549	Education years, anticholinergic comedication	**0.77356**	**0.87637**
	RF	IQ, age, PANSS negative	0.71153	0.83486	Education years	0.79500	0.88817
Reasoning	–	Age, IQ	0.78274	0.88234	No model	–	–
	GA	Age, IQ	**0.73589**	**0.85346**	–	–	–
	RFE	Age, IQ	0.75616	0.86710	Education years	0.87050	0.92966
	RF	Age, IQ	0.75232	0.86181	Education years	**0.86284**	**0.92735**
Vigilance	–	IQ, age	0.89557	0.94169	Education years	**0.94561**	**0.96882**
	GA	IQ, age	**0.84388**	**0.91316**	–	–	–
	RFE	IQ, age	0.88846	0.93566	Education years	0.94823	0.96912
	RF	IQ, age	0.86142	0.91974	Education years	0.94684	0.96777
Neurocognitive summary score	–	IQ, age, education years	0.740366	0.85965	Education years	**0.86634**	**0.92461**
	GA	IQ, age, education years, anticholinergic comedication	**0.67609**	**0.81961**	–	–	–
	RFE	IQ, age	0.71209	0.84267	Education years	0.867101	0.92890
	RF	IQ, age, education years	0.71670	0.84232	Education years	0.86900	0.92921

CTREE, Conditional Inference Tree; GA; genetic algorithm, OQ, intelligence quotient; MSE, mean squared error; RMSE, root mean squared error; RF; random forest to show feature importance, RFE; random forest feature elimination.

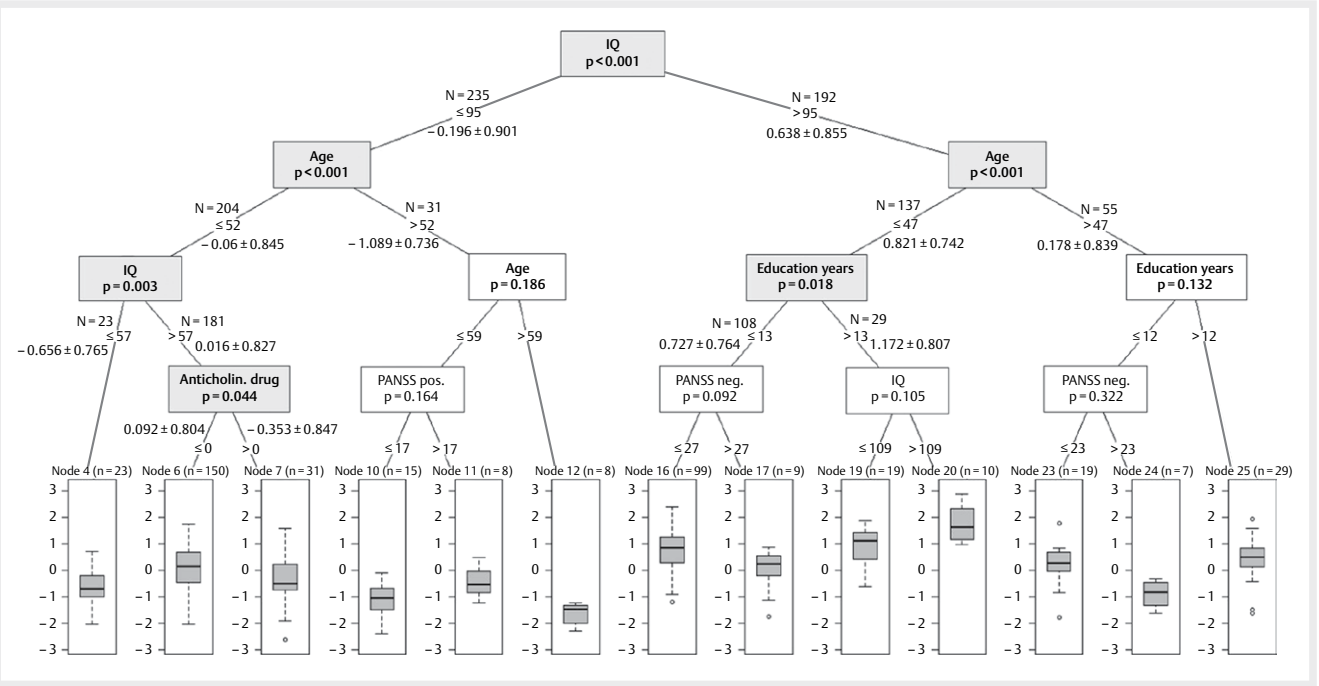
II. IQ ≤ 57 (N = 23; −0.656 ± 0.765) vs. IQ > 57 (N = 181; 0.016 ± 0.827), p = 0.003

I. Education years ≤ 13 (N = 108; 0.727 ± 0.764) vs. education years > 13 (N = 29; 1.172 ± 0.807), p = 0.018

II. Anticholinergic comedication yes (N = 31; −0.353 ± 0.847) vs. no (N = 150; 0.092 ± 0.804), p = 0.044

Discussion

Our CTREE analysis, guided by GA or RFE feature selection or by RF feature importance, revealed interesting findings regarding the influence of moderating factors on various neurocognitive domains. *IQ* represents the most important predictor among all neurocognitive domains. In this sample, the mean IQ was 90 ± 18, ranging from 44–125. Decision trees based on GA models suggested diverse cut-offs for the IQ ranging from 64 to 105. Four models suggested a quite consistent threshold of 94/95 for moderating cognitive scores. Patients with a baseline IQ below 94/95 were, in general, predicted to have significantly lower z-scores in the verbal memory, processing speed, and working memory domain. This also holds true in the model for the neurocognitive summary score.



▶ **Fig. 1** Decision tree summarizing the importance of various features on the neurocognitive summary score (N = 427) including number of patients, decision criterion and mean ± standard deviation z-score for each node.

The age of patients played a role in all neurocognitive domains except for verbal memory, where other factors (IQ and years of education) were more important. The mean patient age was 41.1 ± 10.8 years, ranging from 18 to 66 years. The decision tree models revealed several cut-offs, ranging from 41 to 54 years. An age above 52 years was suggested to be predictive for significantly lower z-scores in the working memory domain and the vigilance domain, as well as for the results of the neurocognitive summary score (based on GA models).

While being one of the main predictors of neurocognitive domain scores, the *years of education* have not been found predominant in the primary GA models (see **supplementary Figure S4** for details). However, after the exclusion of age and IQ, education years was represented in all final models independent from the method used. While the number of education years fluctuated widely among the patient sample from 3 to 21 years (mean 12.1 ± 2.2 years), the cut-off among all models consistently found a minimum number of 12/13 years being relevant in terms of better cognitive performance.

However, other factors, predominantly a higher severity of *PANSS negative symptoms*, seem to cancel out these effects, resulting in significantly lower z-scores (different models processing speed, verbal memory, working memory domains (data not shown)). For the working memory domain, two models (GA and RFE) identify the feature as significant. Here, increased negative symptom scores are associated with lower cognitive performance when compared to the patient group with the same IQ (> 94) but a negative symptom score above 30 ($p = 0.042$). In the processing speed domain, the PANSS negative score becomes significant ($p < 0.05$) in all models after removing age and IQ from the equation, thus suggesting an interplay between these variables. Patients with the same level of education (> 12 years) but a negative symptom score above 26 showed, in general, a lower cognitive performance compared to the patient group with a score below 26 or equal.

Drug type, mean SAS score, and PANSS positive score were, in general, not found to represent strong moderators for neurocognitive functioning, according to feature importance ranking and the final models.

The *concomitant use of anticholinergic medication* was represented as a significant feature in the neurocognitive summary score GA-based model (see ▶ **Fig. 1**), with the use of this drug class being in general predictive for lower z-scores in 31 patients. However, this finding must be regarded in terms of the complex interactions presented in the models.

Our findings furthermore underline the nuanced role of the *estimated minimum dopamine D_2 receptor blockade* on neurocognition as presented in a previous analysis of the same patient sample [1]. The feature of estimated D_2 occupancy was frequently selected/ranked among the significant features across all methods in the vigilance, reasoning, working memory, and verbal memory domains and the neurocognitive sum score. It was further included in the final model for the processing speed domain alongside IQ, age, education years, and PANSS negative score when no prior feature selection was applied, with an estimated D_2 occupancy threshold above 77.6 % indicating poorer performance. However, it did not become significant ($p = 0.22$). This threshold would be well in line with the findings from previous studies on olanzapine that indicate that D_2 occupancies of around 80 % are related to maximum attainable therapeutic effects (measured by schizophrenia symptom scales) [18]. Of note, D_2 occupancy showed a moderate correlation

with the type of antipsychotic being used for the treatment in the patient cohort.

To sum up, in contrast to the findings from Sakurai and colleagues [1], the role of estimated D_2 occupancy was rather negligible among the model predictions and might rather be mediated by other factors. Thus, the role of D_2 occupancy on neurocognitive performance remains elusive, pointing to the need for further study to disentangle its effects. Besides the dominant factors, age and IQ, the PANSS negative scores appear to broadly impact multiple cognitive domains. In everyday patient care, variables such as IQ, age, and years of education remain rather unmodifiable. Negative symptoms, however, offer an interesting area for potential treatment interventions. Previous studies have demonstrated a strong link between negative symptoms and impaired cognitive function in patients with schizophrenia [19]. The focus of treatment that targets negative symptoms may not only ameliorate these symptoms but may also lead to improvements in cognitive function. Several studies have found that interventions targeting negative symptoms, i. e., cognitive remediation therapy and social skills training, can positively affect cognitive abilities [20, 21]. Furthermore, clinical focus on negative symptoms to enhance cognitive function in schizophrenia is of major importance since cognitive impairments are also strongly linked to functional outcomes, such as employment and social interaction [22]. Thus, by better management of negative symptoms, clinicians may open a pathway to enhance the overall real-world performance and functional recovery of patients with schizophrenia [22, 23].

Several limitations should be taken into account when interpreting the findings of this study and considering their broader applicability. First, the study relies on specific statistical techniques and models such as GA, RF, and RFE, which could lead to biases depending on underlying assumptions and parameter tuning. The chosen feature selection methods, while comprehensive, might have overlooked interactive or nonlinear effects between variables, potentially limiting the interpretation of certain predictors like the estimated D_2 occupancy. Second, we have chosen not to set aside a separate holdout set for external validation; all data is used for both training and validation. We relied on 10-fold cross-validation to provide a robust estimate of the model's performance, avoiding any reduction in training data. The potential for over-optimistic results still exists, and the lack of a separate, independent validation sample may affect the true estimation of the model's predictive power. Furthermore, we refrained from using a weighting or enrichment scheme to maintain the integrity of the distribution in our standardized outcome data, manage the risk of overfitting, and ensure a clear interpretability of the model. Third, since the effects of dopamine D_2 blockage by antipsychotic medication on neurocognition can not only be altered by changes in age and IQ; problems may also occur, with negative symptoms being a strong predictor for the outcome domains. Negative symptoms are highly complex and may arise from various underlying neural deficits. No single receptor or pathway can be clearly pinpointed as the mediator of negative symptoms. Moreover, the impact of antipsychotic medications on negative symptoms may be indirect, arising from improvements in other symptom domains, or from interactions between various neurotransmitter systems with, i. e., serotonin 5-HT_{2A} antagonism being repeatedly highlighted in these terms. Whereas olanzapine [24] and risperidone [25] predominantly exert their antipsychotic action via D_2 and 5-HT_{2A} receptors, ziprasidone also has activity at 5-HT_{1A} receptors [26], which may contribute to its clinical effects on negative symptoms. Fourth, the exclusion of key predictors like IQ and age in the secondary analyses might raise questions about the validity of the models in generalizing to broader populations. The study's focus on specific neurocognitive scores and certain medications (e. g., risperidone, ziprasidone, or olanzapine) may limit applicability to other clinical settings or cognitive functions. Lastly, our study establishes associative links rather than causal connections, and these findings prompt further research to investigate the potential for causality.

Conclusion

The work presented shows a robust approach, employing various feature selection methods and CTREE analyses to proficiently exhibit the clinical relevance of particular factors, primarily IQ, age, the number of education years, and the severity of negative schizophrenia symptoms (assessed by PANSS negative score) across multiple cognitive domains. Overall, smaller z scores, which indicate lower neurocognitive function, were associated with advanced age (i. e., age above 52 years), lower IQ (i. e., IQ below 94/95), lower number of education years (i. e., less than 12/13 school years), and more severe negative symptoms (i. e., PANSS negative score above 26). While the verbal memory, processing speed, reasoning, and vigilance domains were dominated by age and IQ as the most relevant factors, the working memory and neurocognitive summary domains seem to be attributable to a highly multifaceted interplay of influencing factors. Our findings confirm a strong connection between negative symptoms and impaired cognitive function, as discussed in previous studies [19, 22, 23]. Personalized treatment plans might benefit from a focus on better management of negative symptoms to enhance real-world performance and functional recovery in schizophrenia patients. Of note, the presented results offer insights into the predictive importance of clinical factors on neurocognitive scores but do not infer causality. Further investigations are warranted to understand the nature of these relationships fully.

Data availability statement

The original contributions can be directed to the corresponding author by reasonable request.

Author Contributions

XML developed the first draft of the protocol and performed the analysis. YM provided assistance with machine learning analysis. RBB conducted the pharmacokinetic analysis. HU supervised the entire manuscript writing and contributed to the revision of the protocol. All authors have read and approved the final manuscript.

Acknowledgement

Dr. Hart was supported by the Overseas Research Fellow Award by the Japan Society for the Promotion of Science.

Conflict of Interest

Dr. Hart was supported by the Overseas Research Fellow Award by the Japan Society for the Promotion of Science. Dr. Bies receives grant funding from NIDA, NIAID, NICHE, USAID, FDA, Bill and Melinda Gates Foundation and serves as a consultant for Advanced Biosciences Laboratories (NIAID). In the past he has served as a consultant for Lumos Biopharma through NGT Biopharma Consultants. Dr. Uchida has received grants from Daiichi Sankyo, Eisai, Mochida, Otsuka, and Sumitomo Pharma; speaker's fees from Eisai, Janssen, Lundbeck, Meiji Seika Pharma, Otsuka, and Sumitomo Pharma; and advisory board fees from Lundbeck, Sumitomo Pharma, Takeda Pharmaceutical Company, and Boehringer Ingelheim Japan. Dr. Mitsukura reports no conflicts of interest.

References

[1] Sakurai H, Bies RR, Stroup ST et al. Dopamine D_2 receptor occupancy and cognition in schizophrenia: Analysis of the CATIE data. Schizophr Bull 2013; 39: 564–574

[2] Keefe RS, Silva SG, Perkins DO et al. The effects of atypical antipsychotic drugs on neurocognitive impairment in schizophrenia: A review and meta-analysis. Schizophr Bull 1999; 25: 201–222

[3] Minzenberg MJ, Poole JH, Benton C et al. Association of anticholinergic load with impairment of complex attention and memory in schizophrenia. Am J Psychiatry 2004; 161: 116–124

[4] Hoff AL, Sakuma M, Wieneke M et al. Longitudinal neuropsychological follow-up study of patients with first-episode schizophrenia. Am J Psychiatry 1999; 156: 1336–1341

[5] Dickinson D, Zaidman SR, Giangrande EJ et al. Distinct polygenic score profiles in schizophrenia subgroups with different trajectories of cognitive development. Am J Psychiatry 2020; 177: 298–307

[6] Akiyama K, Saito S, Saito A et al. Predictive value of premorbid IQ, negative symptoms, and age for cognitive and social functions in Japanese patients with schizophrenia: A study using the Japanese version of the brief assessment of cognition in schizophrenia. Psychiatry Res 2016; 246: 663–671

[7] Lieberman JA, Stroup TS, McEvoy JP et al. Effectiveness of antipsychotic drugs in patients with chronic schizophrenia. N Engl J Med 2005; 353: 1209–1223

[8] Bigos KL, Pollock BG, Coley KC et al. Sex, race, and smoking impact olanzapine exposure. J Clin Pharmacol 2008; 48: 157–165

[9] Feng Y, Pollock BG, Coley K et al. Population pharmacokinetic analysis for risperidone using highly sparse sampling measurements from the CATIE study. Br J Clin Pharmacol 2008; 66: 629–639

[10] Wessels AM, Bies RR, Pollock BG et al. Population pharmacokinetic modeling of ziprasidone in patients with schizophrenia from the CATIE study. J Clin Pharmacol 2011; 51: 1587–1591

[11] Uchida H, Takeuchi H, Graff-Guerrero A et al. Predicting dopamine D_2 receptor occupancy from plasma levels of antipsychotic drugs: A systematic review and pooled analysis. J Clin Psychopharmacol 2011; 31: 318–325

[12] Sheiner LB, Rosenberg B, Marathe VV. Estimation of population characteristics of pharmacokinetic parameters from routine clinical data. J Pharmacokinet Biopharm 1977; 5: 445–479

[13] Cha TA, Beall E, Irvine B et al. At least five related, but distinct, hepatitis C viral genotypes exist. Proc Natl Acad Sci U S A 1992; 89: 7144–7148

[14] Nakajima S, Uchida H, Bies RR et al. Dopamine $D_{2/3}$ receptor occupancy following dose reduction is predictable with minimal plasma antipsychotic concentrations: An open-label clinical trial. Schizophr Bull 2016; 42: 212–219

[15] R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing; Vienna, Austria: 2022

[16] Kuhn M. Building Predictive Models in R Using the caret Package. J Stat Softw 2008; 28: 1–26

[17] Hothorn T, Zeileis A. Partykit: A modular toolkit for recursive partytioning in R. J Mach Learn Res 2015; 16: 3905–3909

[18] Wesner K, Hiemke C, Bergemann N et al. Therapeutic reference range for olanzapine in schizophrenia: Systematic review on blood concentrations, clinical effects, and dopamine receptor occupancy. J Clin Psychiatry 2023; 84: 22r14626

[19] Ventura J, Hellemann GS, Thames AD et al. Symptoms as mediators of the relationship between neurocognition and functional outcome in schizophrenia: A meta-analysis. Schizophr Res 2009; 113: 189–199

[20] Wykes T, Huddy V, Cellard C et al. A meta-analysis of cognitive remediation for schizophrenia: Methodology and effect sizes. Am J Psychiatry 2011; 168: 472–485

[21] Kurtz MM, Müser KT. A meta-analysis of controlled research on social skills training for schizophrenia. J Consult Clin Psychol 2008; 76: 491–504

[22] Green MF, Kern RS, Heaton RK. Longitudinal studies of cognition and functional outcome in schizophrenia: Implications for MATRICS. Schizophr Res 2004; 72: 41–51

[23] Bowie CR, Harvey PD. Cognitive deficits and functional outcome in schizophrenia. Neuropsychiatr Dis Treat 2006; 2: 531–536

[24] Bhana N, Foster RH, Olney R et al. Olanzapine: An updated review of its use in the management of schizophrenia. Drugs 2001; 61: 111–161

[25] Schoretsanitis G, de Leon J, Haen E et al. Pharmacokinetics of risperidone in different application forms - Comparing long-acting injectable and oral formulations. Eur Neuropsychopharmacol 2018; 28: 130–137

[26] Schmidt AW, Lebel LA, Howard HR Jr. et al. Ziprasidone: A novel antipsychotic agent with a unique human receptor binding profile. Eur J Pharmacol 2001; 425: 197–201