Classification of vulnerability levels using multivariate biomarkers in schizophrenia: a machine-learning approach

Simona Caldani (simona.caldani@gmail.com)  
Hopital Universitaire Robert-Debre

François-Benoît Vialatte  
ESPCI Paris

Aurélien Baelde  
ESPCI Paris

Maria Pia Bucci  
Hopital Universitaire Robert-Debre

Narjes Bendjemaa  
Centre Hospitalier Sainte Anne

Célia Mam-Lam-Fook  
Centre Hospitalier Sainte Anne

Raphaël Gaillard  
Centre Hospitalier Sainte Anne

Marie-Odile Krebs  
Centre Hospitalier Sainte Anne

Isabelle Amado  
Centre Hospitalier Sainte Anne

Research article

Keywords: oculomotor biomarkers, coordination, sensory-motor integration, schizophrenia spectrum disorders, neurological soft signs.

Posted Date: April 9th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-15842/v2

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License
Abstract

Background: Schizophrenia is a heterogeneous neurodevelopmental disease involving cognitive and motor impairments. Motor dysfunctions, such as eye movements or neurological soft signs, are proposed as endophenotypic markers.

Methods: Supervised machine-learning methods (Support Vector Machines) applied on oculomotor performances using comprehensive testing with prosaccades, antisaccades, memory-guided saccade tasks and smooth pursuit, as well as neurological soft signs assessment, was used to discriminate patients with schizophrenia (SZ, N=53), full siblings of patients (FS, N=45) and healthy volunteers (C, N=48). 80% of patients were used in a training/validation set and 20% on a test set. The discrimination was measured using the classification error (rate of misclassified patients).

Results: The most reliable classification was between C and SZ, with only 15% and 12% of error rates for validation and test, whereas the SZ vs. FS classification provided the highest error rates (32% of error rate in both validation and test). Interestingly, neurological soft signs were selected as the best predictor, together with a combination of measures, for the two classifications: C vs. SZ, SZ vs. FS. In addition, memory-guided saccades were consistently selected among the best two multimodal features for the classifications involving the control group (C vs. SZ or FS).

Conclusions: Taken together, these results emphasize the importance of neurological soft signs and sensitive oculomotor parameters, especially memory-guided saccades. This classification provides promising avenues for improving early detection of / early intervention in psychosis.

Background

Schizophrenia is a heterogeneous neurodevelopmental disorder. Until now its aetiology has remained unclear, although there is no doubt that genetic vulnerability, interacting with environmental factors, contribute to its development (Davis et al., 2016). In the literature several studies identified endophenotypic markers of this disease that could be made by the application of statistical approaches with multivariate patterns more suitable to make individual level inferences. These approaches include machine-learning techniques (Orru et al., 2012) and support vector machines (Cortes et al., 1995). Used to analyse complex measurements and massive data, machine-learning techniques may help advance the field of better diagnosis among schizophrenia spectrum disorders (Passos et al., 2016).

For a long time, impaired eye movements have been associated with schizophrenia. Abnormalities of saccades (Calkins et al., 2008) and of smooth pursuit eye movements (Levy et al., 2010) were suggested as biomarkers for the schizophrenia spectrum. Among vulnerability markers, neurological soft signs have been widely reported in schizophrenia patients (SZ) (Bombin et al., 2005; Krebs et al., 2000) as well as in full-siblings (FS) (Gourion et al., 2003). Note that neurological soft signs correspond to several types of sensorimotor impairments (e.g. discrete motor dyscoordination, impairment in sequencing, balance, sensory integration, see Krebs & Mouchet, 2007). In the literature it has been reported that they are
Present even before the onset of the illness (Leask et al. 2002), in antipsychotic naïve patients (Krebs et al., 2000; Varambally, 2006), as well as in non-affected relatives (Gourion et al. 2003, 2004). The neurological soft signs in subjects with schizophrenia are associated with abnormalities in the cortex morphology (Gay et al. 2013) and represent a genetically transmitted vulnerability factor (Gourion et al. 2003, 2004, Krebs and Mouchet, 2007).

Recently, our group (Caldani et al., 2017a) explored memory guided saccades and antisaccades in SZ, FS and controls (C) and we found a higher error rate in memory guided saccades in SZ and FS compared to C, while the error rate in antisaccades was significantly higher only in SZ compared to C. Error rate in memory guided saccade corresponds to the number of erroneous saccades made before the extinction of the fixation point (during the memory delay period), whiled for antisaccades error rate represents the number of saccades directed toward the stimulus (direction error). Based on these findings, we suggested that memory guided saccades could be more accurate than antisaccades to detect deficient inhibitory processes. Inhibition defect are considered as a core feature of schizophrenia (Reilly et al. 2014, Krebs et al. 2010). Curtis and colleagues (2001) found a more general problem of inhibition also in relatives of patients with schizophrenia suggesting that the liability to schizophrenia is associated with vulnerability to distraction and disinhibition. More recently, Lencer et al. (2019), using the brain functional connectivity analysis during an antisaccade paradigm, reported in patients with schizophrenia an altered activity in fronto-thalamo-parietal network that is associated with inhibition deficit in these patients but not in controls.

Concerning prosaccade tasks, normal performance was reported in SZ (Gooding et al., 2008), while a deficit in accuracy was found in unmedicated SZ (Crawford et al., 1995); a reduced latency was also observed in naïve patients (Krebs et al., 2010). Regarding smooth pursuit eye movements, poor gain as well as elevated rate of catch-up saccades and intrusive saccades were reported in SZ with a large effect size (O'Driscoll et al., 2008; Franco et al., 2014). Finally, studies using fixation tasks found controversial results, with preserved performance in SZ in some studies (Gooding et al., 2000; Hutton et al., 2002), and alteration in others (Curtis et al., 2001; Raemaekers et al., 2002). Recently, our group found more intrusive saccades in smooth pursuit eye movements and fixation tasks with distractors in SZ compared to C and FS, and more abnormalities in SZ with high neurological soft signs compared to SZ with low neurological soft signs (Caldani et al., 2017b). Using multivariate modelling techniques with eye-movement tests, Benson et al. (2012) found that patients with SZ differed from C in smooth pursuit eye movements, fixation stability and free viewing tasks, with free viewing being the best single discriminator.

To our knowledge, no study explored oculomotor markers using machine-learning approach with the aim to discriminate different profiles within the schizophrenia spectrum. The support vector machines modelling strategy has several advantages over computational methods: it can be easily applied to multi-modal data analysis, and it is not constrained by a priori assumptions or abstractions on the data. Instead, machine-learning is about the construction and study of systems that can learn from data (Dreyfus, 2005). Machine-learning models such as neural networks or support vector machines can be used to design data-driven models. The model is built using the input vectors (in the present case,
multimodal features extracted from eye pursuit or saccades), and matching these vectors to expected output vectors (in our case detection of the group the subject belongs to). Once the model has been built, it is then confronted to a new independent data set to estimate its validity.

By assessing different eye movement paradigms such as memory guided saccades, antisaccades, smooth pursuit eye movements, prosaccades and fixation tasks and adding neurological soft signs characterization, the purpose of the present study was to integrate the data in a modelling strategy by means of these classifying algorithms, in three groups of subjects belonging to the schizophrenia spectrum, namely SZ, FS and healthy controls. Our aim, in this study is to use machine learning model in order to be able to discriminate at best, using different oculomotor parameters and neurological soft signs, the predictive character in schizophrenia. Our hypothesis is that this kind of multivariate analysis will elicit certain biomarkers of eye movements with a good discrimination power, and that neurological soft signs will be a major discriminator, in line with our previous studies (Caldani et al., 2017a,b, Caldani, 2017).

Supervised machine learning methodologies are particularly suited for multivariate modelling. While standard statistical analysis like ANOVA assume a linear model to model the link between a set of qualitative variables and a quantitative variable, machine learning allows to represent the relationships with nonlinear models whose complexity is statistically determined. The absence of assumption on the relationship is compensated using cross validation or leave-one-out to select the optimal complexity of the model. This allows the model to be more expressive in the modeling of the relationship between schizophrenia and oculomotor parameters and naturally account for variable interactions. Moreover, feature selection methods allow to assess the importance of particular predictive characters in unsupervised and supervised manner.

**Methods**

**Subjects**

53 patients with schizophrenia (SZ) fulfilling DSM IV-TR criteria (2003), 45 non-psychotic full-siblings of patients with SZ (FS) and 48 healthy volunteers (controls, C) were recruited in the University Department of Psychiatry at Sainte-Anne Hospital, Paris, France. Diagnostic interview for Genetic Studies (Nurnberger et al., 1994) was used to ascertain diagnosis of SZ, and exclude any DSM4 axis 1 diagnoses for siblings and controls. Patients with SZ were not considered as first episode, were mainly under atypical antipsychotics (46/53), and had been stable (meaning stability of clinical symptoms and of dosage of the treatments delivered) for more than three months. For details see Table 1.

All participants were examined using Neurological Soft Signs Examination (Krebs et al., 2000), which encompasses five dimensions of neurological soft signs (23 items) as well as assessment of extrapyramidal symptoms (Simpson and Angus, 1970), abnormal involuntary movement scale (Guy, 1976) and lateralization (adapted from Edinburgh Inventory). Lastly, patients have been clinically assessed by BPRS (Brief Psychiatric Rating Scale (Overall et al., 1962)).
For all participants in the present study the exclusion criteria were the following: history of neurological/cerebral/ophthalmological disorder, history of substance dependence during the last year and/or for a period of more than five years and recent cannabis abuse, intellectual deficiency, Simpson-Angus score > 3, abnormal involuntary movement scale score > 3 (Le Seach et al., 2012).

The principles of the Declaration of Helsinki were followed and the protocol was approved by our Institutional Human Experimentation Committee (N°2010-A00149-30). All subjects gave their written informed consent before their participation and received 50 euros.

**Oculomotor paradigms**

Stimuli were presented on a 22-inch PC screen. The stimulus was a white filled square subtending a visual angle of 0.5 deg. Eye movements were recorded using the Mobile EBT Tracker (SuriCog), a CE-marked medical eye-tracking device. Recording frequency was set up to 300 Hz. The precision of this system was 0.25°. Four paradigms were used: prosaccades, antisaccades, memory-guided saccades and smooth pursuit eye movements (for details see Caldani et al. 2017a,b; Caldani, 2017).

**Procedures**

Calibration factors for each eye were determined from the eye positions during the calibration procedure (Bucci and Seassau, 2013). For prosaccades we calculated the gain (ratio eye amplitude to target amplitude) as well as the number of anticipation (latency < 80) and of express saccades (latency 80 < x < 130 ms). For antisaccades we also calculated the error, rate, that is the number of saccades directed toward the stimulus. For memory guided saccades, we calculated the error rate (the number of erroneous saccades made before the extinction of the fixation point). The latency of erroneous saccades was classified according to four different time windows corresponding to the precise time the subject initiated the saccade (1st: 80–320 ms; 2d: 350–500 ms; 3rd: 500–750 ms; 4th: > 750 ms). Each window was established in reference to the histograms of these erroneous memory guided saccades latencies, and for each (classification) window these classes were compared to a two-sample Kolmogorov-Smirnov test. A low p-value indicates that two histograms effectively correspond to different distributions.

For smooth pursuit eye movements, we measured the number of saccades with amplitude ≥ 2° and the corresponding gain (which corresponds to the ratio between eye velocity and the target velocity).

**Data analysis**

We performed three analyses: non-supervised clustering, supervised feature selection, and supervised regression. Processing was performed using Matlab R2015b (Mathworks®). The data were subdivided into two subsets: a training/validation set, and an independent test set (containing 20% of the data); these data sets were constituted by partitioning subjects randomly with 10 thousand partitions. The homogeneity of the validation and test sets were optimized by minimizing the two-sample Kolmogorov-
Smirnov statistic between the features of each class in all partitions. The optimal partition was thereby selected.

Groups of subjects were compared in two-class comparisons: each subgroup belonging to classes \( \text{and} \) (where \( a \) and \( b \) represents SZ, FS and C) is represented as an output vector of numbers (zeros and ones) indicating their group. Each of the features (extracted from saccadic parameters, eye pursuit measurements or neurological soft signs total score) and their extended second order combinations are represented as input vectors of dimension \( m \), where \( m \) is the number of subjects belonging to classes \( a \) and \( b \), and each element is the measure of the feature (or cross-term) for subject \( s \). Supervised feature selection was used to identify the best sets of input vectors for each output vector. We analysed iteratively the best set of input vectors to model one expected output vector. This was done using the orthogonal forward regression (OFR) algorithm (Chen et al., 1989). The OFR feature selection approach follows three steps:

- All input vectors are ranked according to their distance to the output. The distance is computed as the cosine of the angle \( \theta \) between the vector and the output: \( \cos \theta \).
- The descriptor with the lowest absolute angle (maximum cosine) is ranked first. All remaining descriptors and the output are projected into the null-space of the best descriptor.
- The selected descriptor is stored and removed from the set, and the algorithm iterates on the remaining orthogonalized features.

In order to control for the relevance of the selected features, we used a probe variables approach (Stoppiglia et al., 2003). We inserted \( N \) randomly drawn vectors in the feature set. The rank distribution of these probes indicated the risk for a descriptor to contain information that could be explained by chance. We fixed a threshold of 10% of probes in our investigation, and selected only descriptors above that threshold. This method systematically produced more than three relevant input vectors, for all 2-class output vectors.

Supervised classification was performed using a support vector machine. Support vector machines are universal approximators: when good care is taken to control their complexity, they can provide better fitting than classical polynomial regressions (Hammer et al., 2003). Support vector machines constructs a hyperplane in a high-dimensional space, which can be used for classification. The original finite-dimensional space is mapped into a higher-dimensional space, with non-linearities obtained by redefining dot products in terms of a nonlinear kernel function. Subsets of of the best features selected using OFR were used to compute and estimate the output vector. The support vector machine was optimized using sequential minimal optimization. Performances were estimated using a leave-one-out approach:

- One sample was taken out of the database.
- The network was then trained on the remaining samples, and afterwards tested on the excluded sample.
The same estimation was performed iteratively for all samples of the database. The overall classification of the excluded samples is the leave-one-out error, which is a good estimate of the generalization error. We tuned the support vector machines hyperparameters (Kernel choice, Kernel parameters and soft margins) according to the leave-one-out error (minimizing the classification error between and ).

Once optimal hyperparameters were estimated, the model was retrained on the whole training/validation set. It was then applied to the test set without any retraining in order to assess the generalization capability of the model and estimate the test error on independent new data.

**Results**

Results obtained with 1, 2 and 3-feature classifications systematically led to a better model when three features were used, with lower validation errors in this condition. We afterwards estimated the generalization error of this model on the independent test subset. Low test error rates were obtained in the [12-32] % range (Table 2). The C vs. SZ classification was the most reliable, with 15% validation and 12% test error rates. The SZ vs. FS classification provided the highest error rates, with 32% of errors in both validation and test (note that a random classification in these two-class classifications would yield a 50% error). *Neurological soft signs* were selected as the best predictor, together with a combination of measures, for two classifications: C vs. SZ and SZ vs. FS. In addition, *Memory-guided saccades (4th window)* were consistently selected among the best two multimodal features for all classifications involving the control group (C vs. SZ and C vs. FS).

We computed the three-class classification resulting from the combination of these two-class classifiers (see Table 3). The detection rate is the probability for a given sample belonging to Class $c$ to be effectively detected as a member of Class $c$. In two-class classifications, the detection rate of each class corresponds to sensitivity and specificity.

The predictive value is the probability, when a sample is detected as a member of Class $c$, for this sample to indeed belong to Class $c$. In two-class classifications, the predictive power of each class corresponds to negative predictive value and positive predictive value.

Group C has the highest detection rate (74%). The SZ group had the highest predictive value (75%). Overall, the classification was reliable for all groups C, SZ and FS with both detection rates and predictive values above 50%.

**Discussion**

In this study, our aim was to develop a model of machine-learning approaches by using automatic learning method in order to discriminate three groups of subjects. Neurological soft signs were selected as the best predictor, together with a combination of measures, for two classifications: C vs. SZ, SZ vs. FS. Finally, memory-guided saccades were consistently selected within the best two multimodal features for all classifications involving the control group (C vs. SZ, C vs FS).
The success of machine-learning approaches relies on two key factors. Firstly, weak variables cannot be used to design an efficient model, as machine-learning approaches only extract the information present in the data. The model performances are therefore bounded by the statistical power of the input variables to predict the expected outcome. Secondly, uncontrolled increase of model complexity can lead to severely biased performances, when the model is overfitting / overfits the data (thereby learning non-reproducible features). As numerous indicators have to be compared, there is indeed a high risk to develop a biased model – this is the well-known “curse of dimensionality”: the more complex the data are, the fewer rules can be extracted using either data mining or classifiers. Solving this issue was done through a reduction of the input feature space, using supervised feature selection.

Low test error rates were obtained for all group comparisons, confirming the interest of neurological soft signs and saccadic measurements for the diagnostic of schizophrenia. Neurological soft signs emerged as the best predictor for the classifications. Memory guided saccades was consistently selected, with error latency occurring in the 4th window (> 750 ms), selected for almost all classifications involving the control group. Neurological soft signs appeared as very sensitive markers of vulnerability for schizophrenia, reflecting a deviance in the maturation of neurodevelopmental brain structures throughout foetal life, infancy and adolescence. This was shown at different stages of the disease (Krebs et Mouchet, 2007; Chan et al., 2017; Caldani et al., 2017a), in siblings of patients (Gourion et al., 2004), or in first episode subjects (Chan et al., 2017). In addition, neurological soft signs was also shown to be linked to ocular movement anomalies (Picard et al., 2009,2012). Memory guided saccades have been proposed as a useful endophenotyp (Calkins et al., 2008) with error rates showing a good sensitivity in patients and relatives (Landgraf et al., 2008; Caldani et al., 2017a). In our findings, errors occurring early with anticipatory saccades in antisaccades combined to neurological soft signs were good classifiers for SZ versus FS, but less powerful than memory guided saccades, in line with our previous analysis (Caldani et al., 2017a). Anticipatory or express saccades in prosaccades in combination with other features were also good classifiers for C vs. SZ. These oculomotor features suggest a defect in inhibition control, underlying the implication of crucial cerebral structures such as dorsolateral prefrontal cortex and frontal eye field (Leigh and Zee, 2015). Our findings could plausibly not be due to a consequence of antipsychotic medications, as made a strict selection in the inclusion of persons with schizophrenia, excluding persons experiencing parkinsonian side effects of the treatment in the NSS assessment (Simpson-Angus score > 3, abnormal involuntary movement scale score > 3). However, a study with a group of naïve patients is warranted.

The reader should bear in mind that the features used for these classifications contain measurement noises, which cannot be reduced to zero, therefore higher classification rates would not be possible without flaws in the model design. In three-class classifications, all classes are detected with high accuracies (random classification accuracy in a three-class classifier would be 25%). In two-class classifications, the highest error rates are obtained for SZ vs. FS (32% error) and C vs. FS (27% error), which is satisfying (random classification error in a three-class classifier would be 50%). The C vs. FS classification task is a comparison of control subjects with asymptomatic subjects with genetic risk
factors. In this condition, a low error rate would be surprising, and indeed the confusion matrix illustrates that this classification error is mainly due to FS samples detected as controls. As can be observed on the predicted FS column in Table 3, 24% of SZ are detected as FS by the model. The FS class identified by our model shares some traits in common with the SZ. On the other hand, 28% of FS are detected as control subjects (C column of the table). FS is a heterogeneous group, with oculomotor performances that could be tightly linked to the performances of their related probands (Mazhari et al., 2011; Curtis et al., 2001) via shared genetic background, but also depending on the personal history, psychopathology or personality profiles of the subjects (Morgan et al., 2015).

Multi-modal biomarkers consistently outperformed monovariate classifications. This is a well-known phenomenon in machine-learning: two apparently weak predictors can become very effective when combined (Dreyfus, 2005).

Limitations

Our sample is rather small. Our patients with schizophrenia were nearly all under treatment, and antipsychotic medications could influence the results. Further studies are warranted to confirm and replicate our results.

Conclusion

The use of different oculomotor paradigms as well as discrete neurological and clinical parameters combined in a three-class classification model has revealed the existence of different profiles, with good predictive values according to the different stages of the schizophrenia spectrum. Many of these oculomotor features support a global inhibitory control defect in the schizophrenia spectrum. The machine-learning methodology allowed us to emphasize the importance of taking into account both Neurological Soft Signs and oculomotor parameters especially memory guided saccades, in the study for early detection strategy. This study brought converging evidence that biomarkers sensitive to neurodevelopmental abnormalities could constitute useful endophenotypes in schizophrenia and could improve the detection of individuals who will convert to psychosis, in order to optimize prevention strategies in psychosis. In the literature, it has been shown that the duration of untreated psychosis could have an important impact on the course and level of remission (Perkins et al., 2005; Marshall et al., 2005), as well as on the economic costs of this disease (Knapp et al., 2014). An effective strategy in order to reduce the duration of untreated psychosis is to detect subjects at risk of psychosis before the onset of the disease (Oppetit et al., 2016). Simple and feasible assessment using oculomotor movements could provide a useful way for detection and prevention of the disease.

Declarations

Ethics approval and consent to participate
The principles of the Declaration of Helsinki were followed and the protocol was approved by our Institutional Human Experimentation Committee (N°2010-A00149-30). All subjects gave their written informed consent before their participation and received 50 euros. When the subject was a minor, consent was given to the parent, the legal guardian of patient.

**Consent to publish**

Written informed consent for publication of their clinical details was obtained from the patient/parent/guardian/relative of the patient.

**Availability of data and materials**

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

**Competing interests**

The authors have no financial and personal relationships with other people or organizations that could inappropriately influence or bias their work in this study.

**Funding**

Fondation pour la recherche en psychiatrie et santé mentale' (FRPSM 2011-2013).

French Health Ministry grant (PHRC AOM 07_118 ICAAR).

**Authors’ Contributors**

All authors participated in critically revising all sections of the manuscript, and have approved the final version. In addition, RG, MOK, IA, conceptualized and designed the study; IA, MPB and FBV directed the study implementation, including designing literature search processes and data analyses, drafting the manuscript; MOK, AI, MPB and FBV completed the literature search and study selection; NB and CMLF contributed to patient recruitment and FBV and AB to analysis tools.

**Acknowledgements**

The study was promoted by INSERM (VISA-AR, 2010-A00149-30).

This study was conducted as part of PhD thesis of Simona Caldani (Paris Sud-Orsay University). We thank the practitioners from the C’JAAD team (Drs M Kazes and J Bourgin, G Martinez, O Gay), and from Service Hospitalo-Universitaire, S14 Centre Hospitalier Sainte-Anne, for their role in participants’ enrolment. We also thank the Clinical Research Center team (M Cuenca, C Bergot), the URC Paris Centre Descartes, and DRCI. We thank the Paris Descartes University’s Centre de Langues for revising the English version of the manuscript.
Abbreviations

SZ : Patients with schizophrenia

FS : Full siblings of patients

C : Healthy volunteers

References


Tables

Table 1: Clinical characteristics of subjects

a. **Demographic and social data:** SZ, patients with schizophrenia; FS, full siblings; C, controls; F, female; M, male. Mean and standard deviation of age and Educational Level of subjects; NSS, Neurological Soft Signs, BPRS Brief Psychiatric Rating Scale.
# Table

<table>
<thead>
<tr>
<th></th>
<th>SZ = 53</th>
<th>FS = 45</th>
<th>C = 48</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sex</strong></td>
<td>11F/42M</td>
<td>25F/20M</td>
<td>22F/26M</td>
</tr>
<tr>
<td><strong>Age (Years)</strong></td>
<td>26±7</td>
<td>28±9</td>
<td>25±5</td>
</tr>
<tr>
<td><strong>Educational Level (Years)</strong></td>
<td>13±2</td>
<td>15±2</td>
<td>14±2</td>
</tr>
<tr>
<td><strong>NSS total Score</strong></td>
<td>12±8</td>
<td>6±4</td>
<td>5±4</td>
</tr>
<tr>
<td><strong>BPRS total score</strong></td>
<td>43±6</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

b. **Medical treatments for SZ**: No treatment was permitted to C and FS. In SZ no benzodiazepine was authorized, but in case of anxiety or insomnia, low dosage of hydroxyzine was permitted (dosage inferior to 50 mg). Antidepressant treatments were allowed, if the initiation had begun more than three weeks before the assessment. The ratio expressed in the table is the number of patients with schizophrenia (SZ) who were under treatment, referring to the total number of SZ, or referring to SZ taking antipsychotics.
Medical treatment | SZ = 53
---|---
Antipsychotics | 45/53
With antidepressant | 8/45
With Hydroxyzine | 1/45
Untreated | 7/53
Antidepressant | 1/53

**Table 2:** 2-class classification of results.

<table>
<thead>
<tr>
<th>Classes</th>
<th>Best features</th>
<th>Validation error</th>
<th>Test error</th>
</tr>
</thead>
<tbody>
<tr>
<td>C vs. SZ</td>
<td>Neurological soft signs, Memory guided saccades (4th window), Prosaccade anticipations</td>
<td>15%</td>
<td>12%</td>
</tr>
<tr>
<td>C vs. FS</td>
<td>Memory guided saccades (4th window), Memory guided saccades (1st window), Intrusive saccades in pursuit 36°/sec</td>
<td>26%</td>
<td>27%</td>
</tr>
<tr>
<td>SZ vs. FS</td>
<td>Neurological soft signs, Anticipation in antisaccades, Saccades expressed in prosaccades</td>
<td>32%</td>
<td>32%</td>
</tr>
</tbody>
</table>
Table 3. Overall confusion matrix: true class is indicated in lines and predicted class in columns. The central diagonal corresponds to the detection rates for each class, and the bottom line indicates predictive values for each class. A 3-class random classifier would yield detection rates and predictive value of 25%.

<table>
<thead>
<tr>
<th></th>
<th>C</th>
<th>SZ</th>
<th>FS</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>74%</td>
<td>9%</td>
<td>17%</td>
</tr>
<tr>
<td>SZ</td>
<td>30%</td>
<td>46%</td>
<td>24%</td>
</tr>
<tr>
<td>FS</td>
<td>28%</td>
<td>7%</td>
<td>65%</td>
</tr>
</tbody>
</table>

Predictive Values 56% 75% 61%

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Dataanalysis.pdf