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: March 3rd, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-15842/v1>

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 (NSS), are proposed as endophenotypic markers. **Methods** Machine-learning method applied on oculomotor performances using comprehensive testing with prosaccades, antisaccades, memory-guided saccade tasks and smooth pursuit, as well as NSS assessment, was used to discriminate patients with schizophrenia (SZ), full siblings of patients (FS) and healthy volunteers (C). **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, NSS 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 (Orri et al., 2012) and support vector machines (SVM) (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 (SPEM) (Levy et al., 2010) were suggested as biomarkers for the schizophrenia spectrum. Among vulnerability markers, neurological soft signs (NSS) 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).

Recently, our group (Caldani et al., 2017a) explored Memory Guided Saccades (MGS) and antisaccades (AS) in SZ, FS and controls (C) and we found a higher error rate in MGS in SZ and FS compared to C, while the error rate in AS was significantly higher only in SZ compared to C. Based on these findings, we suggested that MGS could be more accurate than AS to detect deficient inhibitory processes. Concerning prosaccade tasks (PS), 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 SPEM, 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 SPEM and fixation tasks with distractors in SZ compared to C and FS, and more abnormalities in SZ with high NSS compared to SZ with low NSS (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 SPEM, 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 SVM 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 SVM 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 MGS, AS, SPEM, Prosaccade and Fixation tasks and adding NSS 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 hypothesis is that this kind of multivariate analysis will elicit certain biomarkers of eye movements with a good discrimination power, and that NSS will be a major discriminator, in line with our previous studies (Caldani et al., 2017a,b; Caldani, 2017).

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 mainly under atypical antipsychotics, and had been stable 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 NSS (23 items) as well as assessment of extra-pyramidal 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 Seac'h 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 (AS), memory-guided saccades (MGS) and SPEM (for details see Caldani et al. 2017a,b).

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 AS we also calculated the error, rate, that is the number of saccades directed toward the stimulus. For MGS, 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 MGS 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 SPEM, we measured the number of saccades with amplitude $\geq 2^\circ$ and the corresponding gain (which corresponds to the ratio between eye velocity and the target velocity).

Data analysis

(see Data analysis in the Supplementary Files)

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 (P.V.) 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. NSS emerged as the best predictor for the classifications. MGS was consistently selected, with error latency occurring in the 4th window (> 750 ms), selected for almost all classifications involving the control group. NSS 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, NSS was also shown to be linked to ocular movement anomalies (Picard et al., 2009,2012). MGS 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 AS combined to NSS were good classifiers for SZ versus FS, but less powerful than MGS, 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).

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 MGS, in the study for early detection strategy. This study brought converging evidence that biomarkers sensitive to neurodevelopmental abnormalities could constitute useful endophenotypes in schizophrenia.

List Of Abbreviations

SZ : Patients with schizophrenia

FS : Full siblings of patients

C : Healthy volunteers

NSS : Neurological soft signs

SVM : Support vector machines

MGS : Memory Guided Saccades

AS : Antisaccades

PS : Prosaccades tasks

SPEM : Smooth Pursuit Eye Movements

PV : Predictive value

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.

References

American Psychiatric Association. 2003. Diagnostic and Statistical Manual of Mental Disorders. Fourth Edition. Washington: DC American Psychiatric Association, Text Revision (DSM-IV-TR).

Benson, P.J., Beedie, S.A., Shephard, E., Giegling, I., Rujescu, D., St Clair, D. (2012). Simple viewing tests can detect eye movement abnormalities that distinguish schizophrenia cases from controls with exceptional accuracy. *Biol. Psychiatry*. 72 (9) 716-24.

Bombin, I., Arango, C., Buchanan, R.W. (2005). Significance and meaning of neurological signs in schizophrenia: two decades later. *Schizophr. Bull.* 31 (4) 962-77.

Bucci, M.P., Seassau, M. (2013). Reading and visual search: a developmental study in normal children. *PLoS One*. 8 (7).

Caldani, S. (2017). Vision et Oculomotricité : Marqueurs cognitifs de dépistage chez des sujets atteints de schizophrénie. [Paris]: University of Paris Sud-Orsay.

Caldani, S., Amado, I., Bendjemaa, N., Vialatte, F., Mam-Lam-Fook, C., Gaillard, R., Krebs, M.O., Bucci, M.P. (2017b). Oculomotricity and Neurological Soft Signs: Can we refine the endophenotype? A study in subjects belonging to the spectrum of schizophrenia *Psychiatry Research*. 256:490-497.

Caldani, S., Bucci, M.P., Lamy, J.C., Seassau, M., Bendjemaa, N., Gadel, R., Gaillard, R., Krebs, M.O., Amado, I. (2017a). Saccadic eye movements as markers of schizophrenia spectrum: Exploration in at-risk mental states. *Schizophrenia Research*. 181: 30–37.

Calkins, M.E., Iacono W.G., Ones, D.S. (2008). Eye movement dysfunction in first-degree relatives of patients with schizophrenia: a meta-analytic evaluation of candidate endophenotypes. *Brain. Cogn.* 68: 436–461.

Chan, R.C.K., Cui, H.R., Chu, M.Y., Zhang, T.H., Wang, Y., Wang, Y., Li, Z., Lui, S.S.Y., Wang, J.J., Cheung, E.F.C. (2017). Neurological soft signs precede the onset of schizophrenia: a study of individuals with schizotypy, ultra-high-risk individuals, and first-onset schizophrenia. *Eur. Arch. Psychiatry Clin. Neurosci.* DOI 10.1007/s00406-017-0828-4.

Chen, S., Billings, S.A., and Luo, W. (1989). Orthogonal least squares methods and their application to non-linear system identification, *International Journal of Control*. 50 (5) 1873-1896.

Cortes, C., Vapnik, V. (1995). Support-Vector Networks. *Machine Learning* 20(3):273–297.

Crawford, T.J., Haegar, B., Kennard, C., Reveley, M.A., Henderson, L. (1995). Saccadic abnormalities in psychotic patients, I: neuroleptic-free psychotic patients. *Psychol. Med.* 1995 25:461–471

Curtis, C.E., Calkins, M.E., Grove, W.M., Feil, J.K., Iacono, W.J. (2001). Saccadic Disinhibition in Patients With Acute and Remitted Schizophrenia and Their First Degree Biological Relatives. *Am. J. Psychiatry* 158:100–106.

- Davis, J., Eyre, H., Jacka, F.N., Dodd, S., Dean, O., McEwen, S., Debnath, M., McGrath, J., Maes, M., Amminger, P., McGorry, P.D., Pantelis, C., Berk, M. (2016). A review of vulnerability and risks for schizophrenia: Beyond the two hit hypothesis. *Neurosci. Biobehav. Rev.* 65:185-94.
- Dreyfus, G. (2005). *Neural Networks, Methodology and Applications*. Berlin: Springer.
- Franco, J.G., De Pablo, J., Gaviria, A.M., Sepulveda, E., Vilella, E. (2014). Smooth pursuit eye movements and schizophrenia: Literature review *Arch. Spanish Society of Ophthalmol.* 89:361– 367.
- Gooding, D.C. and Basso, M.A. (2008). The Tell-Tale Tasks: A Review of Saccadic Research in Psychiatric Patient Populations. *Brain Cogn.* 68 (3) 371–390.
- Gooding, D.C., Grabowski, J.A., Hendershot, C.S. (2000). Fixation stability in schizophrenia, bipolar, and control subjects. *Psychiatry Research.* 97:119–112.
- Gourion, D., Goldberger, C., Bourdel, M.C., Bayle, F.J., Millet, B., Olie, J.P., Krebs, M.O. (2003). Neurological soft-signs and minor physical anomalies in schizophrenia: Differential transmission within families. *Schizophr. Res.* 63:181–187.
- Gourion, D., Gourevitch, R., Le provost, J.B., Olié, J.P., Loo, H., Krebs, M.O. (2004). L'hypothèse neurodéveloppementale dans la schizophrénie. *L'Encéphale.* 109-18.
- Guy, W. (1976). *Abnormal Involuntary Movement Scale*. ECDEU Assessment Manual for Psychopharmacology. Washington, DC: US Government Printing Office.
- Hammer, B., Gersmann, K. (2003). A Note on the Universal Approximation Capability of Support Vector Machines. *Neural Processing Letters* 17(1):43-53.
- Hutton, S.B., Joyce, E.M., Barnes, T.R., Kennard, C. (2002). Saccadic distractibility in first episode schizophrenia. *Neuropsychologia* 40:1729–1736.
- Krebs, M.O., Bourdel, M.C., Cherif, Z.R., Bouhours, P., Lôo, H., Poirier, M.F., Amado, I. (2010). Deficit of inhibition motor control in untreated patients with schizophrenia: Further support from visually guided saccade paradigms. *Psych. Res.* 179 : 279–284.
- Krebs, M.O., Gut-Fayand, A., Bourdel, M., Dischamps, J., Olié, J. (2000). Validation and factorial structure of a standardized neurological examination assessing neurological soft signs in schizophrenia". *Schizophr. Res.* 45 (3) 245-260.
- Krebs, M.O., Mouchet, S. (2007). Neurological soft signs and schizophrenia: a review of current knowledge. *Rev. Neurol.* 163 (12) 1157-1168.
- Landgraf, S., Amado, I., Bourdel, M.C., Leonardi, S., Krebs, M.O. (2008). Memory-guided saccade abnormalities in schizophrenic patients and their healthy, full biological siblings. *Psychol. Med.* 38 (6)

861-870.

Le Seac'h, A., Picard, H., Gorsane, M.A., Vidal, P.P., Amado, I., Krebs, M.O. (2012). A step toward an objective quantification of subtle neurological signs in schizophrenia. *Psychiatry Res.* 198 (2):230–234.

Leigh, R.G., Zee, D.S. (2015). *The Neurology of Eye Movements*. 5th Edition Oxford : University Press.

Levy, D.L., Sereno, A.B., Gooding, C.D., and O'Driscoll, A.G. (2010). Eye Tracking Dysfunction in Schizophrenia: Characterization and Pathophysiology. *Curr. Top. Behav. Neurosci.* 4: 311–347.

Mazhari, S., Price, G., Dragovic, M., Waters, F.A., Clissa, P., Jablensky, A. (2011). Revisiting the suitability of antisaccade performance as an endophenotype in schizophrenia. *Brain Cogn.* 77:223–230.

Morgan, C.J., Lenzenweger, M.F., Levy, D.L. (2015). Is Poor Antisaccade Performance in Healthy First-Degree Relatives of Schizophrenics an Artifact of Study Design? Poster presented at McLean Research Day. Belmont, Massachusetts.

Nurnberger, J.I., Blehar, M.C., Kaufmann, C.A., York-Cooler, C., Simpson, S.G., Harkavy-Friedman, J., Severe, J.B., Malaspina, D., Reich, T. (1994). Diagnostic Interview for Genetic Studies (DIGS). Rationale, unique features, and training. *Arch. Gen. Psychiatry.* 51:849–859.

O'Driscoll, G.A., Callahan, B.L. (2008). Smooth pursuit in schizophrenia: A meta-analytic review of research since. *Brain Cogn.* 68:359–370.

Orru, G., Pettersson-Yeo, W., Marquand, A.F., Sartori, G., Mechelli, A. (2012). Using support vector machine to identify imaging biomarkers of neurological and psychiatric disease: a critical review. *Neurosci. Biobehav. Rev.* 36 (4) 1140–1152.

Overall, J.E., and Gorham, D.R. (1962). The Brief Psychiatry Rating Scale. *Psychological Reports.* 10 799-812.

Overall, J.E., and Gorham, D.R. (1962). The Brief Psychiatry Rating Scale. *Psychological Reports.* 10:799-812.

Passos, I.C., Mwangi, B., Cao, B., Hamilton, J.E., Wu, M-J., Zhang, X.Y., Zunta-Soares, G.B., Quevedo, J., Kauer-Sant'Anna, M., Kapczinski, F., Soares, J.,C. (2016). Identifying a clinical signature of suicidality among patients with mood disorders: a pilot study using a machine learning approach. *J Affect Disord.* 193:109–116.

Picard, H., LeSeac'h, A., Amado, I., Gaillard, R., Krebs, M.O., Beauvillain, C. (2012). Impaired saccadic adaptation in schizophrenic patients with high neurological soft sign scores. *Psych. Res.* 199: 12–18.

Picard, H.J., Amado, I., Bourdel, M.C., Landgraf, S., Olié, J.P., Krebs, M.O. (2009). Correlation between neurological soft signs and saccadic parameters in schizophrenia. *Prog. Neuropsychopharmacol. Biol.*

Psych. 1533 (4) 676-681.

Raemaekers, M., Jansma, J.M., Cahn, W., Van der Geest, J.N., Van der Linden, J.A., Kahn, R.S., Ramsey, N.F. (2002). Neuronal substrate of the saccadic inhibition deficit in schizophrenia investigated with 3-dimensional event-related functional magnetic resonance imaging. Archives of General Psychiatry. 59:313–320.

Simpson, G.M., Angus, J.W.S. (1970). A rating scale for extrapyramidal side effects. Acta Psychiatr. Scand. 212 11–19.

Stoppiglia, G., Dreyfus, R., Dubois, and Y. Oussar. (2003). Ranking a random feature for variable and feature selection. J.M.L.R . 3 1399–1414

Tables

Table 1: Clinical characteristics of subjects

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.

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

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.

SZ n=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.

Classes	Best features	Validation error	Test error
C vs. SZ	Neurological soft signs, Memory guided saccades (4 th window), Prosaccade anticipations (PS_Anticipation)	15%	12%
C vs. FS	Memory guided saccades (4 th window), Memory guided saccades (1 st window), Intrusive saccades in pursuit 36°/sec	26%	27%
SZ vs. FS	Neurological soft signs, Anticipation in AS, Saccades expressed in prosaccades	32%	32%

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%.

	C	SZ	FS
C	74%	9%	17%
SZ	30%	46%	24%
FS	28%	7%	65%
P.V.	56%	75%	61%

Supplementary Files

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

- [Dataanalysis.pdf](#)