

1 A multivariate approach to investigate the associations of electrophysiological 2 indices with schizophrenia clinical and functional outcome

3 Luigi Giuliani¹, Nikolaos Koutsouleris², Giulia Maria Giordano¹, Thomas Koenig³, Armida
4 Mucci¹, Andrea Perrottelli¹, Anne Reuf², Mario Altamura⁴, Antonello Bellomo⁴, Roberto
5 Brugnoli⁵, Giulio Corrivetti⁶, Giorgio Di Lorenzo⁷, Paolo Girardi⁸, Palmiero Monteleone⁹, Cinzia
6 Niolu⁷, Silvana Galderisi^{1*}, Mario Maj¹, Italian Network for Research on Psychoses[†]

7

8 ¹Department of Psychiatry, University of Campania Luigi Vanvitelli, Naples, Italy;

9 ²University Hospital, Department of Psychiatry and Psychotherapy, Ludwig Maximilian University of Munich, Munich, Germany;

10 ³Translational Research Center, University Hospital of Psychiatry, University of Bern, Bern, Switzerland;

11 ⁴Psychiatry Unit, Department of Medical Sciences, University of Foggia, Foggia, Italy;

12 ⁵Neurosciences, Mental Health, and Sensory Organs (NESMOS) Department, Faculty of Medicine and Psychology, Sapienza
13 University, Rome, Italy;

14 ⁶Department of Mental Health of ASL (Local Health Company) of Salerno, Salerno, Italy;

15 ⁷Department of Systems Medicine, Psychiatry and Clinical Psychology Unit, Tor Vergata University of Rome, Rome, Italy;

16 ⁸Department of Neurosciences, Mental Health and Sensory Organs, Suicide Prevention Center, Sant'Andrea Hospital, Sapienza
17 University of Rome, Rome, Italy;

18 ⁹Department of Medicine, Surgery and Dentistry "Scuola Medica Salernitana", Section of Neuroscience, University of Salerno,
19 Salerno, Italy.

20 [†]The members of the Italian Network for Research on Psychoses are listed in the Acknowledgments

21 *Corresponding Author: Silvana Galderisi, Department of Psychiatry, University of Campania Luigi Vanvitelli,
22 Largo Madonna delle Grazie, 80138 Naples, Italy; email: silvana.galderisi@gmail.com

23

24 **Running Title:** EEG markers of clinical and functional outcome in schizophrenia

This peer-reviewed article has been accepted for publication but not yet copyedited or typeset, and so may be subject to change during the production process. The article is considered published and may be cited using its DOI.

This is an Open Access article, distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives licence (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is unaltered and is properly cited. The written permission of Cambridge University Press must be obtained for commercial re-use or in order to create a derivative work.

25 **ABSTRACT**

26 **Background:** Different electrophysiological (EEG) indices have been investigated as possible
27 biomarkers of schizophrenia. However, these indices have a very limited use in clinical practice,
28 since their associations with clinical and functional outcome remain unclear.

29 The present study aims to investigate the associations of multiple electrophysiological markers
30 with clinical variables and functional outcome in subjects with schizophrenia (SCZs).

31 **Methods:** Resting-state EEGs (frequency bands and microstates) and auditory ERPs (MMN-P3a
32 and N100-P3b) were recorded in 113 SCZs and 57 healthy controls (HCs) at baseline. Illness- and
33 functioning-related variables were assessed both at baseline and at 4-year follow-up in 61 SCZs.
34 We generated a machine learning classifier for each EEG parameter (frequency bands, microstates,
35 N100-P300 task and MMN-P3a task) to identify potential markers discriminating SCZs from HCs,
36 and a global classifier. Associations of the classifiers' decision scores with illness- and
37 functioning-related variables at baseline and follow-up were then investigated.

38 **Results:** The global classifier discriminated SCZs from HCs with an accuracy of 75.4% and its
39 decision scores significantly correlated with negative symptoms, depression, neurocognition and
40 real-life functioning at 4-year follow-up.

41 **Conclusions:** These results suggest that a combination of multiple EEG alterations is associated
42 to poor functional outcome and its clinical and cognitive determinants in SCZs. These findings
43 need replication, possibly looking at different illness stages in order to implement EEG as a
44 possible tool for the prediction of poor functional outcome.

45

46

47

48 **Keywords:** EEG, machine learning, functional outcome, schizophrenia

49

50 1. INTRODUCTION

51 Despite the continuous advances in pharmacological and psychosocial treatments, schizophrenia
52 still remains one of the most severe mental disorders, characterized by a chronic relapsing course
53 and marked disability in a substantial proportion of patients (1). Although the reduction of
54 symptoms severity contributes to functional recovery, several studies revealed that subjects with
55 schizophrenia in a chronic stage, with remission of psychotic symptoms, still have serious
56 impairment in different areas of real-life functioning, including independent living, work activities
57 and social relationships (2, 3). In fact, the functional recovery is influenced by the interaction of
58 multiple factors, which represent major determinants of impairment in the above-mentioned real-
59 life functioning areas, beyond psychotic symptoms (4-10).

60 The identification of objective neurophysiological indices associated with the determinants of
61 functional outcome might represent a crucial step towards the implementation of personalized
62 treatments and the identification of new treatment strategies, aiming at improving the functional
63 recovery of subjects with schizophrenia (11-14). Indeed, so far, we are not able to predict
64 individual's outcome across different stages of the illness (15, 16). In addition, most studies
65 investigating determinants of poor functional outcome, such as negative symptoms and cognitive
66 impairment, did not contribute to any increase in knowledge concerning the underlying
67 neurobiological processes (17-19).

68 Identifying biological markers of factors associated with functional outcome, and of the outcome
69 itself, may contribute to the generation of detailed and specific pathophysiological models,
70 resulting in more accurate predictions, as well as to the development of innovative treatment
71 interventions (20).

72 Electrophysiological indices have been largely investigated as possible biomarkers of
73 schizophrenia (21-24).

74 A number of quantitative resting-state EEG and Event-Related Potentials (ERP) alterations have
75 been reported in subjects with schizophrenia in different stages of the illness and many of them
76 are associated with psychopathology, cognitive impairment and functional outcome (25-29).

77 In particular, different studies showed that gamma band activity and Mismatch Negativity (MMN)
78 are associated with functional impairment and may predict the course of the illness in chronic (30-
79 32) and in first-episode psychosis patients, as well as in subjects at clinical high-risk of psychosis
80 (28, 33, 34). Conflicting evidence has been reported for other EEG bands and ERPs (35, 36). As

81 to determinants of functional outcome, cognitive impairment was found to be associated with
82 alterations in multiple resting-state frequency bands (29, 37), abnormalities of P300 amplitude and
83 latency (27, 37, 38), deficit in both N100 amplitude and sensory gating (29, 37, 39), and lower
84 MMN amplitude (29, 37, 40-42). As regard to psychopathology, the severity of negative symptoms
85 was found to be related to increased slower rhythms in resting-state recordings and reduced N100
86 amplitude (25, 43, 44). Conflicting findings were reported about the relationship between negative
87 symptoms and other ERPs (44).

88 However, none of these EEG indices has been implemented in clinical practice, probably due to
89 the variability of the methodology across studies (sample size, illness phase, experimental
90 paradigms) and the paucity of relevant studies investigating several outcome determinants and
91 multiple EEG indices.

92 Indeed, the majority of the studies focused only on the associations between EEG indices and
93 specific clinical or functional outcome measures, rarely assessing more than one or few outcome
94 determinants. This represents an important obstacle to the comprehension of the neurobiological
95 mechanisms associated with the outcome of schizophrenia (45). In fact, as previously reported, the
96 pathways to functional recovery are extremely complex, involving different factors which directly
97 and indirectly influence the real-life functioning of subjects with schizophrenia (4-8). Recent
98 studies considering candidate EEG biomarkers of schizophrenia and several disease-related
99 variables, such as cognitive impairment and negative symptoms, demonstrated multiple
100 contribution of different EEG indices to cognitive deficits and negative symptoms, leading to poor
101 functional outcome (45). In addition, considering that schizophrenia presents a high rate of
102 variability also in terms of pathophysiology (46, 47), the investigation of one or only few EEG
103 indices, instead of a combination of them, is too limiting for the evaluation of the prognostic value
104 of EEG in schizophrenia. Therefore, the association of these potential EEG markers of
105 schizophrenia with the functional outcome still remains unclear (48). Lastly, the possibility of
106 implementing EEG indices in clinical routine as prognostic markers of schizophrenia is also related
107 to the ability of formulating outcome predictions beyond group-level prognostication (15, 49).

108 In order to achieve this goal, in the last decade, different approaches, such as machine learning,
109 deep learning or “multiverse” approaches, were adopted to identify combinations of
110 neurophysiological indices associated with different characteristics of the disease, accounting for
111 the complexity and the heterogeneity of the pathophysiological pathways towards the functional

112 outcome of schizophrenia (21, 50-52). The multiverse approach indicated no associations among
113 multiple EEG features discriminating patients from controls, suggesting that each feature might
114 subtend a different aspect, thus reflecting the heterogeneity of the syndrome at the
115 phenomenological and pathophysiological level (51). As a matter of fact, even in the same illness
116 phase, e.g., chronic stage, schizophrenia is characterized by heterogeneity as to the course and
117 functional outcome (5-11).

118 In the light of these observations, our study aimed to identify patterns of EEG indices, among those
119 discriminating subjects with schizophrenia from controls, which might predict the functional
120 outcome of the disease. Therefore, we first identified the EEG markers which best discriminated
121 subjects with schizophrenia from controls, without preselection of the parameters, and then we
122 investigated the relationships of these patterns with the functional outcome and the
123 psychopathological and neuropsychological determinants of the functional outcome, e.g., negative
124 symptoms and neurocognitive deficits. We decided to use machine learning techniques which are
125 able to learn statistical functions from multidimensional data, recognize data patterns and use those
126 identified patterns to make prediction about individuals (49, 53).

127 To these aims, we analyzed a well characterized population of community dwelling chronic and
128 clinically stable subjects with schizophrenia and matched healthy controls.

129 EEGs were recorded in resting-state condition and during two different tasks, in order to obtain
130 different neurophysiological measures. The EEG indices to analyze as possible prognostic markers
131 of schizophrenia were chosen according to the literature on the topic (23, 29, 37, 54). Indeed, we
132 selected the neurophysiological indices which have been found to be frequently altered in subjects
133 with schizophrenia and those showing the strongest association with the functional outcome (23,
134 29, 37, 54). Therefore, multiple frequency bands and microstates parameters were obtained from
135 the resting-state EEG recording; MMN and P3a were obtained from the EEG recorded during a
136 passive auditory paradigm (in which the subjects had no task), and N100 and P3b were obtained
137 from the EEG recorded during an auditory oddball task. We used a machine learning approach to
138 identify the EEG patterns which better discriminated subjects with schizophrenia from healthy
139 controls and we assessed the associations of these patterns with symptom dimensions, cognitive
140 impairment and real-life functioning in subjects with schizophrenia.

141

142 2. METHODS

143 **2.1 Study Participants**

144 The study has been conducted as part of the add-on EEG study of the Italian Network for Research
145 on Psychoses (4-8). One hundred and forty-eight subjects with schizophrenia (SCZs) and 70
146 healthy controls (HCs) were recruited for the cross-sectional study, at five research sites in Naples,
147 Foggia, Rome “Tor Vergata”, Rome “Sapienza” and Salerno. All 148 SCZs recruited for the cross-
148 sectional study were asked to participate in the longitudinal study, after 4 years of follow-up.

149

150 Baseline

151 The group composed by SCZs included individuals consecutively seen at the outpatient units of
152 the five mentioned Italian university psychiatric clinics. Inclusion criteria for SCZs were a
153 diagnosis of schizophrenia according to DSM-IV, confirmed with the Structured Clinical
154 Interview for DSM IV — Patient version (SCID-I-P), and an age between 18 and 65 years. HCs
155 were recruited from the community at the same sites mentioned above. Inclusion criterion for HCs
156 was the absence of a current or lifetime Axis I or II psychiatric diagnosis. Exclusion criteria for
157 both groups were: (a) history of head trauma with loss of consciousness; (b) history of moderate
158 to severe mental retardation or of neurological diseases; (c) history of alcohol and/or substance
159 abuse in the last six months; (d) current pregnancy or lactation; (e) inability to provide an informed
160 consent. Other exclusion criteria for SCZs were treatment modifications and/or hospitalization due
161 to symptom exacerbation in the last three months. The electrophysiological add-on EEG study was
162 approved by the Ethics Committee of the involved institutions and the study was performed in
163 accordance with the ethical standards laid down in the 1964 Declaration of Helsinki. All
164 participants signed a written informed consent to participate after receiving a detailed explanation
165 of the study procedures and goals.

166

167 Follow-up

168 Only SCZs participated to the 4-year longitudinal study. The inclusion criterion was a diagnosis
169 of schizophrenia according to DSM-IV, confirmed by the Structured Clinical Interview for DSM-
170 IV - Patient version (SCID-I-P). Exclusion criteria were: (a) history of head trauma with loss of
171 consciousness in the four-years interval between baseline and follow-up; (b) progressive cognitive
172 deterioration possibly due to dementia or other neurological illness diagnosed in the last 4 years;
173 (c) history of alcohol and/or substance abuse in the last 6 months; (d) current pregnancy or

174 lactation; I inability to provide an informed consent; (f) treatment modifications and/or
175 hospitalization due to symptom exacerbation in the last 3 months. The longitudinal study was
176 approved by the Local Ethics Committees of the participating centers. All patients signed a written
177 informed consent to participate, after receiving a comprehensive explanation of the study
178 procedures and goals.

179

180 **2.2 Assessment instruments**

181 *Baseline*

182 At baseline, all subjects were evaluated for socio-demographic variables such as age, education
183 and gender, through a clinical form filled using every available source of information. The Positive
184 and Negative Syndrome Scale (PANSS) was used to rate severity of positive, negative and
185 disorganization symptoms in SCZs (55). Scores for these dimensions were calculated based on the
186 consensus 5-factor solution proposed by Wallwork et al (for negative dimension we use the
187 Wallwork criteria except for the item “G7 – motor retardation”, which was excluded from the
188 calculation of this dimension) (56). A semi-structured interview, the Brief Negative Symptom
189 Scale (BNSS) was used to assess negative symptoms in SCZs (57). According to literature (57,
190 58), the domains evaluated by this instrument loaded on two factors: “experiential domain”,
191 consisting of anhedonia, asociality and avolition, and “expressive deficit”, including blunted affect
192 and alogia. We also assessed depressive symptoms using the Calgary Depression Scale for
193 Schizophrenia (CDSS) (59) and extrapyramidal symptoms using the St. Hans Rating Scale (SHRS)
194 for Extrapyramidal Syndromes (60). Neurocognitive functions were rated using the Measurement
195 and Treatment Research to Improve Cognition in Schizophrenia (MATRICS) Consensus
196 Cognitive Battery (MCCB) (61). This battery assesses seven distinct cognitive domains:
197 processing speed, attention/vigilance, working memory, verbal learning, visual learning, social
198 cognition, and reasoning and problem solving. Raw scores on the MCCB were standardized to T-
199 scores, corrected for age and gender, based on the Italian normative sample of community
200 participants. For summary score of cognitive domains including more than one measure and for
201 Neurocognitive and Overall composite scores, we calculated T-score summing the T-scores of the
202 tests included in each domain and then standardizing the sum to a T-score (62).

203 We assessed real-life functioning using the Specific Level of Functioning Scale (SLOF), a hybrid
204 instrument which evaluates many aspects of functioning and is based on the key caregiver’s

205 judgment on behavior and functioning of the patient (63). It is composed of 43 items and includes
206 the following domains: physical efficiency, skills in self-care, interpersonal relationships, social
207 acceptability, community activities (e.g., shopping, using public transportation), and working
208 abilities. In our study we interviewed the key relative, usually the individual most frequently and
209 closely in contact with the patient.

210

211 *Follow-up*

212 At follow-up, a clinical form was filled with data about the course of the disease and treatment
213 information during the previous 4 years, using every available source of information (patients,
214 relatives, medical records and mental health workers). All the variables which had been measured
215 at baseline were tested also at follow-up, using the same assessment tools.

216

217 **2.3 EEG recording procedures**

218 EEGs were recorded only at baseline, using two highly comparable EEG recording systems:
219 EASYS2 (Brainscope, Prague) and Galileo MIZAR-Sirius (EBNeuro, Florence). Before starting
220 the study, a harmonization of the amplifier settings and recording procedure was performed to
221 ensure the same recording settings in all the centers. EEGs were recorded using a cap electrode
222 system with 29 unipolar leads (Fpz, Fz, Cz, Pz, Oz, F3, F4, C3, C4, FC5, FC6, P3, P4, O1, O2,
223 Fp1, Fp2, F7, F8, T3, T4, T5, T6, AF3, AF4, PO7, PO8, Right Mastoid and Left Mastoid), placed
224 following the 10–20 system. All the leads were referenced to linked earlobes (a resistor of 10 k Ω
225 was interposed between the earlobe leads). A ground electrode was placed on the forehead. The
226 following neurophysiological indices were analyzed: frequency bands activity and microstates
227 extracted from the resting-state EEG recording, four ERP components registered during the two
228 different auditory tasks (MMN, P3a and N100, P3b). Further details on the recording procedure
229 and data preprocessing are provided in the Supplementary materials.

230

231 **2.5 Statistical analyses**

232 Two sample t-test and χ^2 test were used for group comparison (SCZs vs HCs). The same analysis
233 was conducted to compare subjects who took part in the longitudinal study with subjects who did
234 not. For the SCZs sample, within-subject comparisons at baseline and follow-up were performed

235 using paired-sample t-test and χ^2 test. Bonferroni-Holm correction was applied to comparisons in
236 order to control for type-I error inflation.

237 Matlab release 2019b was used for all the above-described analyses.

238

239 In order to discriminate SCZs from HCs we generated four different machine learning classifiers,
240 one for each EEG parameter (frequency bands, microstates, N100-P300 and MMN-P3a) and a
241 global classifier resulting from the combination of the four unimodal classifiers' output. The
242 machine learning platform NeuroMiner version 1.0 ([https:// github.com/neurominer-git](https://github.com/neurominer-git);
243 MATLAB release 2019b), was employed to set up a machine learning strategy for testing the
244 classification performance (SCZs vs HCs) of the four EEG unimodal classifiers and, later, of the
245 global classifier (Figure 1).

246 The goal of this approach was to investigate whether, using all the information coming from
247 classifiers using different EEG features, could lead to a higher classification accuracy, compared
248 to the single classifiers' ones. Statistical significance ($p < 0.05$) of individual and global classifiers
249 was assessed with permutation testing, using 1000 permutations of the labels.

250 The detailed machine learning pipeline is reported in supplementary materials and is synthesized
251 in Figure 1.

252 A post-hoc analyses was conducted to compare the individual classifier with the best accuracy and
253 the global classifier (McNemar test). T-tests for independent samples were performed for the 10%
254 most frequently selected features of each individual classifier according to the parameter "selection
255 probability", and Person's correlations were performed on the same EEG indices to estimate the
256 amount of shared information contained in the variables that were used to distinguish SCZs and
257 HCs. Moreover, we performed a Pearson's correlation between the chlorpromazine equivalent
258 doses and the output of the global classifier, the classifier's decision scores, in order to account for
259 the possible impact of antipsychotic medications on the patients' EEG.

260

261 In order to investigate the correlations of the classifiers' decision scores with illness-related
262 variables and real-life functioning in SCZs at baseline and follow-up, we firstly projected
263 baseline variables to four factors using a Non-Negative Matrix Factorization (NNMF). We chose
264 NNMF instead of other dimensionality reduction methods because it produces clearly separated
265 and well-defined variance components, enhancing results' interpretability (64). The number of

266 factors was chosen selecting the optimal dimension that allowed the encoding of data variability
267 while discarding noise. In order to do this, we calculated the variation of the residual error of the
268 data approximation with the variation of the number of estimated components, determining the
269 optimal number of factors by detecting the inflection point of the slope of the reconstruction
270 error (65). The resulting sparse factor matrices were inspected, and the factors were interpreted
271 according to the variables showing non-negative loadings on a given factor. After that, we
272 projected the same illness-related variables and real-life functioning indices, measured at 4 years
273 follow-up, to four factors using the same NMF algorithm, in order to confirm if the obtained
274 baseline latent variables remained stable from baseline to follow-up. The obtained factor scores
275 were used to compute factor trajectories from baseline to follow-up and paired-sample t-test was
276 used to assess the significance of the changes. Pearson's correlations were performed between
277 classifiers' decision scores and the scores of each of the four factors resulting from NMF at
278 baseline and follow-up. All the correlation analyses were corrected for multiple comparisons.
279 Matlab release 2019b was used for NMF and Pearson's correlation analysis.

280

281 3. RESULTS

282 3.1 Sociodemographic and clinical characteristics of the study sample

283 One hundred and forty-eight SCZs and 70 HCs were originally enrolled in the baseline study.
284 Thirty-three SCZs and 13 HCs were excluded because they were found to have more than 25% of
285 missing values in at least one modality data (frequency bands, microstates, MMN & P3a and N100
286 & P300). Two subjects from the SCZs group were excluded after visual inspection of the EEG
287 recordings for an excess of artifacts. Therefore, 113 SCZs and 57 HCs were included in the
288 analysis. As regard the EEG recording systems, the EEGs of 88 SCZs (77.9%) and 40 HCs (70.2%)
289 were recorded using the Galileo MIZAR-Sirius system, while the EEGs of the remaining subjects
290 were recorded using the Easys2 system. There was no group difference in the percentage of
291 subjects recorded with the Galileo MIZAR or the Easys2 system ($\chi^2=1.21$; $p=0.27$). Demographic
292 characteristics and illness related variables are provided in Table 1. We did not find significant
293 group differences for age ($t = 1.05$; $p = .30$). Gender distribution significantly differed between
294 groups ($\chi^2 = 7.02$; $p < .01$), with a higher percentage of males in the patient compared to the control
295 group. Patients had significantly lower education level than controls ($t = -3.49$; $p < .01$). The
296 average duration of illness in the patient group was 12.75 ± 8.29 years. SCZs were characterized

297 by absent to mild positive and disorganization symptom severity (PANSS mean score < 9 for both
298 dimensions) and mild to moderate negative symptom severity (PANSS negative dimension mean
299 score of 15.58 and BNSS total score of 34.88). They had a low mean level of depression (CDSS
300 total score < 4) and parkinsonism (SHRS Parkinsonism score < 1). SCZs, compared to HCs,
301 showed worse performance on cognitive tests (neurocognitive composite score: $t = -10.13$ and $p <$
302 $.001$ overall composite score including social cognition: $t = -9.53$ and $p < .001$) and worse
303 functioning (SLOF-Personal care skills: $t = -5.40$ and $p < .001$; SLOF-Interpersonal relationships:
304 $t = -12.84$ and $p < .001$; SLOF-Social acceptability: $t = -5.32$ and $p < .001$; SLOF-Everyday life
305 skills: $t = -8.44$ and $p < .001$; SLOF-Work skills: -9.47 and $p < .001$). Sixty-one SCZs from the
306 113 patients who had taken part in the baseline study, participated in the four-year follow-up study.
307 Table 2 shows comparisons on demographic characteristics and illness-related variables between
308 follow-up participants ($N=61$) and the rest of the original SCZs ($N=52$) sample. Patients who
309 participated in the follow-up study did not significantly differ from the rest of the sample on
310 baseline socio-demographic characteristics and illness-related variables, except for global
311 parkinsonism ($t = 3.15$; $p = .002$) (Table 2). This mean difference in parkinsonism was relatively
312 small and not clinically significant; thus, the 61 patients participating in the follow-up study can
313 be considered representative of the original sample. The mean values and SDs of all variables
314 included in the analysis at baseline and follow-up are reported in Table 3. In the overall sample of
315 61 subjects participating in the follow-up study, improvements in severity of disorganization, the
316 experiential domain of BNSS negative symptoms, and global parkinsonism were found.
317 Neurocognition was stable, while the overall cognitive performance improved after 4 years. We
318 did not find significant changes in real-life functioning from baseline to follow-up. The NNMF
319 analysis showed four stable factors during different time point (baseline and follow-up): one factor
320 captured functioning and cognitive impairments, a second factor positive symptoms and
321 parkinsonism, a third factor captured negative symptoms, in particular the “expressive deficit”
322 subdomain, and the fourth factor captured depression (Figure 2). Exploring the NNMF factors
323 trajectories, only the factor capturing functioning and cognitive impairment significantly changed
324 ($p = 0.005$) from baseline to follow-up (eTable 1 in supplementary materials).

325

326 **3.2 SCZs vs HCs classification performance**

327 Since there was a gender imbalance between the two sample groups (SCZs and HCs), in order to
328 control for the possible confounding effect of this factor, we created a gender classifier, using EEG
329 variables as predictors. We found that this classifier correctly discriminated males from females
330 with a balanced accuracy of 52.6% and was not significant ($p = 0.25$). Moreover, we created a
331 EEG classifier with all the features together entered as input in the algorithm independently from
332 the data modality, including gender among predictors. Thus, we compared this model with an
333 identical classifier without gender among predictors. We found no significant differences in the
334 accuracy of the two classifier (eTable 2 in supplementary materials). So, we concluded that EEG
335 indices are not influenced by the gender, and we did not correct the other analyses for this variable.
336 Also education was different between SCZs and HCs, but we did not use it as a covariate in the
337 analyses because lower education level is a well-known consequence of schizophrenia.

338 As regard to EEG classifiers, detailed statistics of all classifiers are reported in Table 4. The
339 balanced accuracy was highest for the frequency bands classifier and lowest for the microstate
340 one. Figure 3 shows the 10% most frequently selected features for each classifier. The results of
341 the group comparisons on these EEG features and the correlations among these same indices are
342 reported in the supplementary materials (eTable 3 supplement; eFigure 1 supplement). The global
343 classifier discriminated SCZs from HCs with a balanced accuracy of 75.4% ($p < 0.01$), which was
344 statistically different from the frequency band classifier's accuracy ($\chi^2 = 7.111$; $p = 0.008$). As
345 expected, the decisions generated by frequency bands classifier ($\rho = 0.54$) were the most important
346 for the final classification, followed by N100-P3b ($\rho = 0.46$) and MMN-P3a ($\rho = 0.44$). The
347 decision generated by microstates classifier was the less important for the classification ($\rho = 0.21$)
348 (Figure 4).

349 We did not find significant correlation between the chlorpromazine equivalent doses of
350 antipsychotic medications and the global classifier's decision scores ($r = 0.160$; $p = 0.171$).

351

352 **3.3 Association of classifiers' output with illness-related variables and real-life functioning**

353 No significant association were found between the classifiers' decision scores and the NNMF
354 factors obtained from illness-related variables and real-life functioning measured at baseline. On
355 the contrary, our results showed significant correlations of the global classifier output with
356 depression, negative symptoms, functioning and cognitive impairment at 4 years follow-up (Table
357 5). The direction of the correlations indicated that higher global classifier's decision score at

358 baseline were associated with more severe negative symptoms, depression and cognitive
359 impairment, and lower real-life functioning at follow-up. The results of the Pearson's correlations
360 between the individual classifiers' output and the NMF factors' scores at follow-up are reported
361 in Table 5.

362

363 **4. DISCUSSION**

364 Our results showed that each classifier, using different EEG indices, can identify patterns of neural
365 alterations which are able to significantly distinguish SCZs from HCs at individual level.
366 Combining those patterns of EEG indices recorded under different conditions the classification
367 accuracy significantly increases up to 75.4%. The resulting combination of EEG alterations, in
368 chronic patients with schizophrenia, was associated with real-life functioning and with illness-
369 related variables which have an impact on functional outcome, such as cognitive impairment,
370 depression and negative symptoms, at 4-years follow-up (4, 6). Previous research identified
371 alterations of several EEG indices in SCZs, which are related to different brain functions and
372 associated with different illness features influencing the outcome or with the outcome itself (25-
373 29). However, despite the results of these studies, no EEG index has been implemented in clinical
374 practice.

375 In this study we evaluated multiple EEG indices, recorded under different conditions, and we used
376 a machine learning approach in order to identify patterns of electrophysiological alterations which
377 could better predict illness outcome. Using this strategy we tried to improve the precision in
378 detecting the relationships of electrophysiological alterations with clinical features, and the
379 knowledge of the pathophysiological pathways involved in schizophrenia outcome. Indeed,
380 schizophrenia is a heterogeneous syndrome with a high variability in brain structure influenced by
381 gene-environment interactions (66-68). Moreover, the pathways toward the outcome are extremely
382 complex, with several factors influencing real-life functioning of people with schizophrenia (4-6,
383 8, 69). A combination of factors more than any single of them is probably involved in determining
384 individual subjects outcome, and the identification of reproducible, objective indicators might
385 facilitate the implementation of translational studies results, improving the knowledge about the
386 relative pathophysiological mechanisms. Previous studies used different approaches to investigate
387 multiple EEG alterations in schizophrenia and the correlations between these neurophysiological
388 alterations and illness-related variables (51, 70, 71). The majority of these studies demonstrated

389 that a weighted combination of electrophysiological features provides better information about the
390 characteristics of the disorder than any single index. However, only a limited number of parameters
391 for each electrophysiological index were included and varied among studies. Within this
392 framework, machine learning methods have the advantage of learning statistical functions from
393 multidimensional data in order to make prediction about individuals. Therefore, in the present
394 study they allowed us to recognize, among a huge amount of parameters (e.g., band activity or
395 ERP amplitude at multiple electrode sites) of different markers, an EEG pattern which was able to
396 discriminate single subjects with schizophrenia from controls. Furthermore, the summary index of
397 this EEG pattern, represented by the decision scores of the global classifier, could be used to
398 investigate the association of such specific combination of neurophysiological markers with the
399 functional outcome, as well as with the clinical and neuropsychological determinants of functional
400 outcome. Indeed, we found that the most selected features by each classifier were poorly correlated
401 to each other, except for the microstates parameters which were significantly associated to theta
402 and alpha activity. These results, in line with those obtained with the multiverse approach (51)
403 demonstrate that combining multiple EEG parameters associated with different characteristics of
404 the disease could lead to a better recognition of the heterogeneous pathophysiological mechanisms,
405 allowing more accurate predictions of the SCZs outcome.

406 Among the different EEG indices investigated in the present study, resting-state frequency bands
407 activity turned out to be the most important feature for the classification of SCZs and HCs, while
408 microstates parameters seem to be redundant with the frequency bands oscillations, adding very
409 little information to the global classifier. According to previous findings, we found that slower
410 bands activity alterations were the most specific of schizophrenia, and, in particular, decreased
411 alpha 2 activity and increased theta 1 and theta 2 activity (23). The alterations in theta and alpha
412 activity are associated with gray and white matter volume reduction in SCZs. Theta activity is
413 associated with learning and its alterations are present in first-degree relatives of SCZs, are
414 independent from antipsychotic medications and are associated with biological vulnerability to
415 schizophrenia (72, 73). Genetic analyses showed that theta activity is correlated with two different
416 genetic components, comprising genes participating extensively in brain development,
417 neurogenesis and synaptogenesis (74). Theta abnormalities were also mediated by gene clusters
418 involved in glutamic acid pathways, cadherin and synaptic contact-based cell adhesion processes.
419 Alpha rhythm is functionally related to memory and attention (75), and is associated with the

420 default mode network activity, involved in cognitive functioning (18). Some genome-wide and
421 positional gene-based analysis showed correlations between alpha activity and tissue-specific
422 single nucleotide polymorphism (SNP), codifying for protein involved in signal transmission,
423 inflammation, and other biological functions (76). These associations were found principally at
424 cortical level (hippocampus, frontal cortex, anterior cingulate cortex) and in putamen (76).
425 According to these findings, it is possible to assume that slower band activity in SCZs reflects
426 alterations of cortical functions linked to specific genetic patterns.

427 Correlation analyses revealed that global classifier's decision scores were associated with real-life
428 functioning and different illness-related variables (cognitive impairment, depression, and negative
429 symptoms) at follow-up. On the opposite, we did not find any association with positive symptoms
430 and disorganization. Negative symptoms and cognitive impairment are core features of
431 schizophrenia, and are present, respectively, in more than 50% and 80% of patients (77, 78).

432 Available evidence indicates that, differently from positive symptoms and disorganization, they
433 are largely present at the onset of the disorder and during the prodromal stages of the disease (77,
434 79). Moreover, in more than half of the cases, negative symptoms have a continuous or relapsing
435 course and cognitive deficit is relatively stable throughout the course of the illness, unlike positive
436 symptoms, which usually have variable severity (77, 80). Both cognitive dysfunction and negative
437 symptoms are associated in chronic patients to poor functional outcome (4-8, 81).

438 No correlations were found with the same features measured at baseline. Our hypothesis is that
439 neurophysiological alterations occur before their related clinical manifestations and reflect the
440 severity of these manifestations measured months or years after the neurophysiological findings.

441 The study has a number of limitations. The first one is the sample size, which is larger compared
442 to previous electrophysiological studies, but relatively small considering the complexity of the
443 machine learning structure. Moreover, in order to make our findings more generalizable, the
444 above-reported classifiers should be applied to an independent sample. Additionally, the study
445 sample is composed only of SCZs and HCs. In order to improve the specificity of the
446 electrophysiological model, it is necessary to include also patients with other psychiatric
447 syndromes. Moreover, our sample is composed of chronic patients, with an average duration of
448 illness of 12.75 years and a median age of 36.34 years. Schizophrenia is particularly prevalent in
449 young adults between 20 and 30 years of age and the onset follows years of prodromal symptoms
450 and leads to disability in about half of the patients (82). Furthermore, different studies

451 demonstrated that the early intervention leads to a better prognosis (83, 84). Therefore, the main
452 goal of any prognostic tool should be the early recognition of the illness and the possibility to make
453 outcome prediction at the onset of the syndrome. To do this, our model needs to be tested also in
454 first-episode psychotic and at-risk subjects. Furthermore, the prognostic information obtained from
455 the analysis does not allow to make predictions about individuals, but it only describes the
456 associations between electroencephalographic patterns and outcome measures at a group level.
457 These results suggest that a combination of different EEG alterations found in SCZs and associated
458 with main determinants of functional outcome and the outcome itself, could be able to predict the
459 course of schizophrenia. To assess whether this neurophysiological pattern can be implemented as
460 prognostic marker of schizophrenia in clinical practice, further studies are required including
461 validation samples and subjects at different stages of the disorder.

462 **Acknowledgements**

463 Members of the add-on EEG study of the Italian Network for Research on Psychoses participating
464 in this study include: Giuseppe Piegari, Eleonora Merlotti, Francesco Brando, Pasquale Pezzella
465 (University of Campania “Luigi Vanvitelli”); Raffaella Carnevale, Maddalena La Montagna
466 (University of Foggia); Fabiola Ferrentino, Alberto Siracusano (University of Rome “Tor
467 Vergata”); Anna Comparelli, Valentina Corigliano (University of Rome “La Sapienza”);
468 Giammarco Cascino, Gianfranco Del Buono (University of Salerno).

469

470 **Author contributions**

471 The project idea was initiated by SG, AM and MM. NK created NeuroMiner software used in the
472 work. LG, NK and AM planned the experimental procedures; LG performed the data analyses and
473 wrote the first draft of the manuscript. NK, GMG, AM, SG and MM contributed to the supervision
474 of the manuscript. All Authors were responsible for the interpretation of the analysis and
475 contributed to critically revising the content.

476

477 **Funding and other support**

478 The study was funded by the Italian Society of Psychopathology (SOPSI), the Italian Society of
479 Biological Psychiatry (SIPB), Roche, Switzerland Lilly, United States AstraZeneca, United
480 Kingdom Lundbeck foundation, Denmark and Bristol-Myers Squibb, United Kingdom. These
481 entities had no role in study design; in the collection, analysis and interpretation of data; in the
482 writing of the report and in the decision to submit the paper for publication.

483

484 **Conflict of interest**

485 All authors declared no conflict of interest

486

487 **Data availability statement**

488 All data generated or analyzed during this study are included in this published article.

489

490 **References**

491

492

- 493
- 494 1. Fusar-Poli P, Estrade A, Stanghellini G, Venables J, Onwumere J, Messas G, Gilardi L,
495 Nelson B, Patel V, Bonoldi I, Aragona M, Cabrera A, Rico J, Hoque A, Otaiku J, Hunter N,
496 Tamelini MG, Maschiao LF, Puchivailo MC, Piedade VL, Keri P, Kpodo L, Sunkel C, Bao J,
497 Shiers D, Kuipers E, Arango C, Maj M. The lived experience of psychosis: a bottom-up review
498 co-written by experts by experience and academics. *World Psychiatry*. 2022;21(2):168-88.
- 499 2. Killaspy H, Harvey C, Brasier C, Brophy L, Ennals P, Fletcher J, Hamilton B.
500 Community-based social interventions for people with severe mental illness: a systematic review
501 and narrative synthesis of recent evidence. *World Psychiatry*. 2022;21(1):96-123.
- 502 3. Disease GBD, Injury I, Prevalence C. Global, regional, and national incidence,
503 prevalence, and years lived with disability for 328 diseases and injuries for 195 countries, 1990-
504 2016: a systematic analysis for the Global Burden of Disease Study 2016. *Lancet*.
505 2017;390(10100):1211-59.
- 506 4. Galderisi S, Rossi A, Rocca P, Bertolino A, Mucci A, Bucci P, Rucci P, Gibertoni D,
507 Aguglia E, Amore M, Bellomo A, Biondi M, Brugnoli R, Dell'Osso L, De Ronchi D, Di Emidio
508 G, Di Giannantonio M, Fagiolini A, Marchesi C, Monteleone P, Oldani L, Pinna F, Roncone R,
509 Sacchetti E, Santonastaso P, Siracusano A, Vita A, Zeppegno P, Maj M, Italian Network For
510 Research on P. The influence of illness-related variables, personal resources and context-related
511 factors on real-life functioning of people with schizophrenia. *World Psychiatry*. 2014;13(3):275-
512 87.
- 513 5. Galderisi S, Rucci P, Kirkpatrick B, Mucci A, Gibertoni D, Rocca P, Rossi A, Bertolino
514 A, Strauss GP, Aguglia E, Bellomo A, Murri MB, Bucci P, Carpiniello B, Comparelli A, Cuomo
515 A, De Berardis D, Dell'Osso L, Di Fabio F, Gelao B, Marchesi C, Monteleone P, Montemagni C,
516 Orsenigo G, Pacitti F, Roncone R, Santonastaso P, Siracusano A, Vignapiano A, Vita A,
517 Zeppegno P, Maj M, Italian Network for Research on P. Interplay Among Psychopathologic
518 Variables, Personal Resources, Context-Related Factors, and Real-life Functioning in Individuals
519 With Schizophrenia: A Network Analysis. *JAMA Psychiatry*. 2018;75(4):396-404.
- 520 6. Galderisi S, Rucci P, Mucci A, Rossi A, Rocca P, Bertolino A, Aguglia E, Amore M,
521 Bellomo A, Bozzatello P, Bucci P, Carpiniello B, Collantoni E, Cuomo A, Dell'Osso L, Di Fabio
522 F, di Giannantonio M, Gibertoni D, Giordano GM, Marchesi C, Monteleone P, Oldani L,
523 Pompili M, Roncone R, Rossi R, Siracusano A, Vita A, Zeppegno P, Maj M, Italian Network for

- 524 Research on P. The interplay among psychopathology, personal resources, context-related
525 factors and real-life functioning in schizophrenia: stability in relationships after 4 years and
526 differences in network structure between recovered and non-recovered patients. *World*
527 *Psychiatry*. 2020;19(1):81-91.
- 528 7. Giuliani L, Giordano GM, Bucci P, Pezzella P, Brando F, Galderisi S. Improving
529 Knowledge on Pathways to Functional Outcome in Schizophrenia: Main Results From the Italian
530 Network for Research on Psychoses. *Front Psychiatry*. 2021;12:791117.
- 531 8. Mucci A, Galderisi S, Gibertoni D, Rossi A, Rocca P, Bertolino A, Aguglia E, Amore M,
532 Bellomo A, Biondi M, Blasi G, Brasso C, Bucci P, Carpiniello B, Cuomo A, Dell'Osso L,
533 Giordano GM, Marchesi C, Monteleone P, Niolu C, Oldani L, Pettorruso M, Pompili M,
534 Roncone R, Rossi R, Tenconi E, Vita A, Zeppegno P, Maj M, Italian Network for Research on P.
535 Factors Associated With Real-Life Functioning in Persons With Schizophrenia in a 4-Year
536 Follow-up Study of the Italian Network for Research on Psychoses. *JAMA Psychiatry*.
537 2021;78(5):550-9.
- 538 9. Ventura J, Helleman GS, Thames AD, Koellner V, Nuechterlein KH. Symptoms as
539 mediators of the relationship between neurocognition and functional outcome in schizophrenia: a
540 meta-analysis. *Schizophr Res*. 2009;113(2-3):189-99.
- 541 10. Peralta V, Gil-Berrozpe GJ, Sanchez-Torres A, Cuesta MJ. Clinical relevance of general
542 and specific dimensions in bifactor models of psychotic disorders. *World Psychiatry*.
543 2021;20(2):306-7.
- 544 11. Maj M, van Os J, De Hert M, Gaebel W, Galderisi S, Green MF, Guloksuz S, Harvey
545 PD, Jones PB, Malaspina D, McGorry P, Miettunen J, Murray RM, Nuechterlein KH, Peralta V,
546 Thornicroft G, van Winkel R, Ventura J. The clinical characterization of the patient with primary
547 psychosis aimed at personalization of management. *World Psychiatry*. 2021;20(1):4-33.
- 548 12. Salazar de Pablo G, Studerus E, Vaquerizo-Serrano J, Irving J, Catalan A, Oliver D,
549 Baldwin H, Danese A, Fazel S, Steyerberg EW, Stahl D, Fusar-Poli P. Implementing Precision
550 Psychiatry: A Systematic Review of Individualized Prediction Models for Clinical Practice.
551 *Schizophr Bull*. 2021;47(2):284-97.
- 552 13. McGorry PD, Mei C, Chanen A, Hodges C, Alvarez-Jimenez M, Killackey E. Designing
553 and scaling up integrated youth mental health care. *World Psychiatry*. 2022;21(1):61-76.

- 554 14. Hickie IB. Implementing 21st century "end-to-end" and technology-enhanced care for
555 young people. *World Psychiatry*. 2022;21(1):79-81.
- 556 15. Fusar-Poli P, Hijazi Z, Stahl D, Steyerberg EW. The Science of Prognosis in Psychiatry:
557 A Review. *JAMA Psychiatry*. 2018;75(12):1289-97.
- 558 16. Leichsenring F, Steinert C, Rabung S, Ioannidis JPA. The efficacy of psychotherapies
559 and pharmacotherapies for mental disorders in adults: an umbrella review and meta-analytic
560 evaluation of recent meta-analyses. *World Psychiatry*. 2022;21(1):133-45.
- 561 17. Kaur R, Sood A, Lang DK, Goyal R, Saini B. Prospecting the Intricate Role of Novel and
562 Potent Biomarkers in Schizophrenia. *Curr Top Med Chem*. 2021;21(16):1441-56.
- 563 18. Menon V. Brain networks and cognitive impairment in psychiatric disorders. *World*
564 *Psychiatry*. 2020;19(3):309-10.
- 565 19. Stein DJ, Shoptaw SJ, Vigo DV, Lund C, Cuijpers P, Bantjes J, Sartorius N, Maj M.
566 Psychiatric diagnosis and treatment in the 21st century: paradigm shifts versus incremental
567 integration. *World Psychiatry*. 2022;21(3):393-414.
- 568 20. Carpenter WT. Primary psychosis: more to know, much more to do. *World Psychiatry*.
569 2021;20(1):1-2.
- 570 21. Barros C, Silva CA, Pinheiro AP. Advanced EEG-based learning approaches to predict
571 schizophrenia: Promises and pitfalls. *Artif Intell Med*. 2021;114:102039.
- 572 22. Earls HA, Curran T, Mittal V. A Meta-analytic Review of Auditory Event-Related
573 Potential Components as Endophenotypes for Schizophrenia: Perspectives From First-Degree
574 Relatives. *Schizophr Bull*. 2016;42(6):1504-16.
- 575 23. Galderisi S, Mucci A, Volpe U, Boutros N. Evidence-based medicine and
576 electrophysiology in schizophrenia. *Clin EEG Neurosci*. 2009;40(2):62-77.
- 577 24. Kotov R, Jonas KG, Carpenter WT, Dretsch MN, Eaton NR, Forbes MK, Forbush KT,
578 Hobbs K, Reininghaus U, Slade T, South SC, Sunderland M, Waszczuk MA, Widiger TA,
579 Wright AGC, Zald DH, Krueger RF, Watson D, Hi TOPUW. Validity and utility of Hierarchical
580 Taxonomy of Psychopathology (HiTOP): I. Psychosis superspectrum. *World Psychiatry*.
581 2020;19(2):151-72.
- 582 25. Giordano GM, Brando F, Perrottelli A, Di Lorenzo G, Siracusano A, Giuliani L, Pezzella
583 P, Altamura M, Bellomo A, Cascino G, Del Casale A, Monteleone P, Pompili M, Galderisi S,
584 Maj M, Italian Network for Research on P. Tracing Links Between Early Auditory Information

- 585 Processing and Negative Symptoms in Schizophrenia: An ERP Study. *Front Psychiatry*.
586 2021;12:790745.
- 587 26. Giordano GM, Koenig T, Mucci A, Vignapiano A, Amodio A, Di Lorenzo G, Siracusano
588 A, Bellomo A, Altamura M, Monteleone P, Pompili M, Galderisi S, Maj M, add-on
589 EEGsotINfRoP. Neurophysiological correlates of Avolition-apathy in schizophrenia: A resting-
590 EEG microstates study. *Neuroimage Clin*. 2018;20:627-36.
- 591 27. Giordano GM, Perrottelli A, Mucci A, Di Lorenzo G, Altamura M, Bellomo A, Brugnoli
592 R, Corrivetti G, Girardi P, Monteleone P, Niolu C, Galderisi S, Maj M, The Italian Network For
593 Research On P. Investigating the Relationships of P3b with Negative Symptoms and
594 Neurocognition in Subjects with Chronic Schizophrenia. *Brain Sci*. 2021;11(12).
- 595 28. Haigh SM, Coffman BA, Salisbury DF. Mismatch Negativity in First-Episode
596 Schizophrenia: A Meta-Analysis. *Clin EEG Neurosci*. 2017;48(1):3-10.
- 597 29. Perrottelli A, Giordano GM, Brando F, Giuliani L, Mucci A. EEG-Based Measures in At-
598 Risk Mental State and Early Stages of Schizophrenia: A Systematic Review. *Front Psychiatry*.
599 2021;12:653642.
- 600 30. Featherstone RE, Melnychenko O, Siegel SJ. Mismatch negativity in preclinical models
601 of schizophrenia. *Schizophr Res*. 2018;191:35-42.
- 602 31. Giordano GM, Giuliani L, Perrottelli A, Bucci P, Di Lorenzo G, Siracusano A, Brando F,
603 Pezzella P, Fabrazzo M, Altamura M, Bellomo A, Cascino G, Comparelli A, Monteleone P,
604 Pompili M, Galderisi S, Maj M, The Italian Network For Research On P. Mismatch Negativity
605 and P3a Impairment through Different Phases of Schizophrenia and Their Association with Real-
606 Life Functioning. *J Clin Med*. 2021;10(24).
- 607 32. Koshiyama D, Miyakoshi M, Thomas ML, Joshi YB, Molina JL, Tanaka-Koshiyama K,
608 Sprock J, Braff DL, Swerdlow NR, Light GA. Unique contributions of sensory discrimination
609 and gamma synchronization deficits to cognitive, clinical, and psychosocial functional
610 impairments in schizophrenia. *Schizophr Res*. 2021;228:280-7.
- 611 33. Higgins A, Lewandowski KE, Liukasemsarn S, Hall MH. Longitudinal relationships
612 between mismatch negativity, cognitive performance, and real-world functioning in early
613 psychosis. *Schizophr Res*. 2021;228:385-93.

- 614 34. Tada M, Nagai T, Kirihara K, Koike S, Suga M, Araki T, Kobayashi T, Kasai K.
615 Differential Alterations of Auditory Gamma Oscillatory Responses Between Pre-Onset High-
616 Risk Individuals and First-Episode Schizophrenia. *Cereb Cortex*. 2016;26(3):1027-35.
- 617 35. Jhung K, Cho SH, Jang JH, Park JY, Shin D, Kim KR, Lee E, Cho KH, An SK. Small-
618 world networks in individuals at ultra-high risk for psychosis and first-episode schizophrenia
619 during a working memory task. *Neurosci Lett*. 2013;535:35-9.
- 620 36. van Tricht MJ, Ruhrmann S, Arns M, Muller R, Bodatsch M, Velthorst E, Koelman JH,
621 Bour LJ, Zurek K, Schultze-Lutter F, Klosterkötter J, Linszen DH, de Haan L, Brockhaus-
622 Dumke A, Nieman DH. Can quantitative EEG measures predict clinical outcome in subjects at
623 Clinical High Risk for psychosis? A prospective multicenter study. *Schizophr Res*. 2014;153(1-
624 3):42-7.
- 625 37. Perrottelli A, Giordano GM, Brando F, Giuliani L, Pezzella P, Mucci A, Galderisi S.
626 Unveiling the Associations between EEG Indices and Cognitive Deficits in Schizophrenia-
627 Spectrum Disorders: A Systematic Review. *Diagnostics (Basel)*. 2022;12(9).
- 628 38. Parker DA, Trotti RL, McDowell JE, Keedy SK, Hill SK, Gershon ES, Ivleva EI,
629 Pearlson GD, Keshavan MS, Tamminga CA, Clementz BA. Auditory Oddball Responses Across
630 the Schizophrenia-Bipolar Spectrum and Their Relationship to Cognitive and Clinical Features.
631 *Am J Psychiatry*. 2021;178(10):952-64.
- 632 39. Boutros NN, Brockhaus-Dumke A, Gjini K, Vedeniapin A, Elfakhani M, Burroughs S,
633 Keshavan M. Sensory-gating deficit of the N100 mid-latency auditory evoked potential in
634 medicated schizophrenia patients. *Schizophr Res*. 2009;113(2-3):339-46.
- 635 40. Xiong YB, Bo QJ, Wang CM, Tian Q, Liu Y, Wang CY. Differential of Frequency and
636 Duration Mismatch Negativity and Theta Power Deficits in First-Episode and Chronic
637 Schizophrenia. *Front Behav Neurosci*. 2019;13:37.
- 638 41. Sehatpour P, Avissar M, Kantrowitz JT, Corcoran CM, De Baun HM, Patel GH, Girgis
639 RR, Brucato G, Lopez-Calderon J, Silipo G, Dias E, Martinez A, Javitt DC. Deficits in Pre-
640 attentive Processing of Spatial Location and Negative Symptoms in Subjects at Clinical High
641 Risk for Schizophrenia. *Front Psychiatry*. 2020;11:629144.
- 642 42. Qu X, Liukasemsarn S, Tu J, Higgins A, Hickey TJ, Hall MH. Identifying Clinically and
643 Functionally Distinct Groups Among Healthy Controls and First Episode Psychosis Patients by
644 Clustering on EEG Patterns. *Front Psychiatry*. 2020;11:541659.

- 645 43. Boutros NN, Mucci A, Vignapiano A, Galderisi S. Electrophysiological aberrations
646 associated with negative symptoms in schizophrenia. *Curr Top Behav Neurosci.* 2014;21:129-56.
- 647 44. Boutros NN, Mucci A, Diwadkar V, Tandon R. Negative symptoms in schizophrenia.
648 *Clin Schizophr Relat Psychoses.* 2014;8(1):28-35B.
- 649 45. Koshiyama D, Thomas ML, Miyakoshi M, Joshi YB, Molina JL, Tanaka-Koshiyama K,
650 Sprock J, Braff DL, Swerdlow NR, Light GA. Hierarchical Pathways from Sensory Processing
651 to Cognitive, Clinical, and Functional Impairments in Schizophrenia. *Schizophr Bull.*
652 2021;47(2):373-85.
- 653 46. First MB, Gaebel W, Maj M, Stein DJ, Kogan CS, Saunders JB, Poznyak VB, Gureje O,
654 Lewis-Fernandez R, Maercker A, Brewin CR, Cloitre M, Claudino A, Pike KM, Baird G, Skuse
655 D, Krueger RB, Briken P, Burke JD, Lochman JE, Evans SC, Woods DW, Reed GM. An
656 organization- and category-level comparison of diagnostic requirements for mental disorders in
657 ICD-11 and DSM-5. *World Psychiatry.* 2021;20(1):34-51.
- 658 47. Insel TR. Rethinking schizophrenia. *Nature.* 2010;468(7321):187-93.
- 659 48. Nieman DH, Ruhrmann S, Dragt S, Soen F, van Tricht MJ, Koelman JH, Bour LJ,
660 Velthorst E, Becker HE, Weiser M, Linszen DH, de Haan L. Psychosis prediction: stratification
661 of risk estimation with information-processing and premorbid functioning variables. *Schizophr*
662 *Bull.* 2014;40(6):1482-90.
- 663 49. Chekroud AM, Bondar J, Delgadillo J, Doherty G, Wasil A, Fokkema M, Cohen Z,
664 Belgrave D, DeRubeis R, Iniesta R, Dwyer D, Choi K. The promise of machine learning in
665 predicting treatment outcomes in psychiatry. *World Psychiatry.* 2021;20(2):154-70.
- 666 50. Park SM, Jeong B, Oh DY, Choi CH, Jung HY, Lee JY, Lee D, Choi JS. Identification of
667 Major Psychiatric Disorders From Resting-State Electroencephalography Using a Machine
668 Learning Approach. *Front Psychiatry.* 2021;12:707581.
- 669 51. Gordillo D, da Cruz JR, Chkonia E, Lin WH, Favrod O, Brand A, Figueiredo P,
670 Roinishvili M, Herzog MH. The EEG multiverse of schizophrenia. *Cereb Cortex.* 2022.
- 671 52. Barros C, Roach B, Ford JM, Pinheiro AP, Silva CA. From Sound Perception to
672 Automatic Detection of Schizophrenia: An EEG-Based Deep Learning Approach. *Front*
673 *Psychiatry.* 2021;12:813460.
- 674 53. Dwyer DB, Falkai P, Koutsouleris N. Machine Learning Approaches for Clinical
675 Psychology and Psychiatry. *Annu Rev Clin Psychol.* 2018;14:91-118.

- 676 54. Boutros NN, Arfken C, Galderisi S, Warrick J, Pratt G, Iacono W. The status of spectral
677 EEG abnormality as a diagnostic test for schizophrenia. *Schizophr Res.* 2008;99(1-3):225-37.
- 678 55. Kay SR, Fiszbein A, Opler LA. The positive and negative syndrome scale (PANSS) for
679 schizophrenia. *Schizophr Bull.* 1987;13(2):261-76.
- 680 56. Wallwork RS, Fortgang R, Hashimoto R, Weinberger DR, Dickinson D. Searching for a
681 consensus five-factor model of the Positive and Negative Syndrome Scale for schizophrenia.
682 *Schizophr Res.* 2012;137(1-3):246-50.
- 683 57. Kirkpatrick B, Strauss GP, Nguyen L, Fischer BA, Daniel DG, Cienfuegos A, Marder
684 SR. The brief negative symptom scale: psychometric properties. *Schizophr Bull.*
685 2011;37(2):300-5.
- 686 58. Strauss GP, Keller WR, Buchanan RW, Gold JM, Fischer BA, McMahon RP, Catalano
687 LT, Culbreth AJ, Carpenter WT, Kirkpatrick B. Next-generation negative symptom assessment
688 for clinical trials: validation of the Brief Negative Symptom Scale. *Schizophr Res.* 2012;142(1-
689 3):88-92.
- 690 59. Addington J, Shah H, Liu L, Addington D. Reliability and validity of the Calgary
691 Depression Scale for Schizophrenia (CDSS) in youth at clinical high risk for psychosis.
692 *Schizophr Res.* 2014;153(1-3):64-7.
- 693 60. Gerlach J, Korsgaard S, Clemmesen P, Lauersen AM, Magelund G, Noring U, Povlsen
694 UJ, Bech P, Casey DE. The St. Hans Rating Scale for extrapyramidal syndromes: reliability and
695 validity. *Acta Psychiatr Scand.* 1993;87(4):244-52.
- 696 61. Nuechterlein KH, Green MF, Kern RS, Baade LE, Barch DM, Cohen JD, Essock S,
697 Fenton WS, Frese FJ, 3rd, Gold JM, Goldberg T, Heaton RK, Keefe RS, Kraemer H, Mesholam-
698 Gately R, Seidman LJ, Stover E, Weinberger DR, Young AS, Zalcman S, Marder SR. The
699 MATRICS Consensus Cognitive Battery, part 1: test selection, reliability, and validity. *Am J*
700 *Psychiatry.* 2008;165(2):203-13.
- 701 62. Mucci A, Galderisi S, Green MF, Nuechterlein K, Rucci P, Gibertoni D, Rossi A, Rocca
702 P, Bertolino A, Bucci P, Helleman G, Spisto M, Palumbo D, Aguglia E, Amodeo G, Amore M,
703 Bellomo A, Brugnoli R, Carpiniello B, Dell'Osso L, Di Fabio F, di Giannantonio M, Di Lorenzo
704 G, Marchesi C, Monteleone P, Montemagni C, Oldani L, Romano R, Roncone R, Stratta P,
705 Tenconi E, Vita A, Zeppegno P, Maj M, Italian Network for Research on P. Familial aggregation

- 706 of MATRICS Consensus Cognitive Battery scores in a large sample of outpatients with
 707 schizophrenia and their unaffected relatives. *Psychol Med.* 2018;48(8):1359-66.
- 708 63. Mucci A, Rucci P, Rocca P, Bucci P, Gibertoni D, Merlotti E, Galderisi S, Maj M, Italian
 709 Network for Research on P. The Specific Level of Functioning Scale: construct validity, internal
 710 consistency and factor structure in a large Italian sample of people with schizophrenia living in
 711 the community. *Schizophr Res.* 2014;159(1):144-50.
- 712 64. Sotiras A, Resnick SM, Davatzikos C. Finding imaging patterns of structural covariance
 713 via Non-Negative Matrix Factorization. *Neuroimage.* 2015;108:1-16.
- 714 65. Frigyesi A, Hoglund M. Non-negative matrix factorization for the analysis of complex
 715 gene expression data: identification of clinically relevant tumor subtypes. *Cancer Inform.*
 716 2008;6:275-92.
- 717 66. Alnaes D, Kaufmann T, van der Meer D, Cordova-Palomera A, Rokicki J, Moberget T,
 718 Bettella F, Agartz I, Barch DM, Bertolino A, Brandt CL, Cervenka S, Djurovic S, Doan NT,
 719 Eisenacher S, Fatouros-Bergman H, Flyckt L, Di Giorgio A, Haatveit B, Jonsson EG, Kirsch P,
 720 Lund MJ, Meyer-Lindenberg A, Pergola G, Schwarz E, Smeland OB, Quarto T, Zink M,
 721 Andreassen OA, Westlye LT, Karolinska Schizophrenia Project C. Brain Heterogeneity in
 722 Schizophrenia and Its Association With Polygenic Risk. *JAMA Psychiatry.* 2019;76(7):739-48.
- 723 67. Guloksuz S, Pries LK, Delespaul P, Kenis G, Luykx JJ, Lin BD, Richards AL, Akdede B,
 724 Binbay T, Altinyazar V, Yalincetin B, Gumus-Akay G, Cihan B, Soygur H, Ulas H, Cankurtaran
 725 E, Kaymak SU, Mihaljevic MM, Petrovic SA, Mirjanic T, Bernardo M, Cabrera B, Bobes J, Saiz
 726 PA, Garcia-Portilla MP, Sanjuan J, Aguilar EJ, Santos JL, Jimenez-Lopez E, Arrojo M,
 727 Carracedo A, Lopez G, Gonzalez-Penas J, Parellada M, Maric NP, Atbasog Lu C, Ucok A,
 728 Alptekin K, Saka MC, Genetic R, Outcome of Psychosis i, Arango C, O'Donovan M, Rutten
 729 BPF, van Os J. Examining the independent and joint effects of molecular genetic liability and
 730 environmental exposures in schizophrenia: results from the EUGEI study. *World Psychiatry.*
 731 2019;18(2):173-82.
- 732 68. Lahey BB, Moore TM, Kaczkurkin AN, Zald DH. Hierarchical models of
 733 psychopathology: empirical support, implications, and remaining issues. *World Psychiatry.*
 734 2021;20(1):57-63.
- 735 69. Moura BM, Isvoranu AM, Kovacs V, Van Rooijen G, Van Amelsvoort T, Simons CJP,
 736 Bartels-Velthuis AA, Bakker PR, Marcelis M, De Haan L, Schirmbeck F. The Puzzle of

- 737 Functional Recovery in Schizophrenia-Spectrum Disorders-Replicating a Network Analysis
738 Study. *Schizophr Bull.* 2022;48(4):871-80.
- 739 70. Light GA, Swerdlow NR, Thomas ML, Calkins ME, Green MF, Greenwood TA, Gur
740 RE, Gur RC, Lazzeroni LC, Nuechterlein KH, Pela M, Radant AD, Seidman LJ, Sharp RF,
741 Siever LJ, Silverman JM, Sprock J, Stone WS, Sugar CA, Tsuang DW, Tsuang MT, Braff DL,
742 Turetsky BI. Validation of mismatch negativity and P3a for use in multi-site studies of
743 schizophrenia: characterization of demographic, clinical, cognitive, and functional correlates in
744 COGS-2. *Schizophr Res.* 2015;163(1-3):63-72.
- 745 71. Price GW, Michie PT, Johnston J, Innes-Brown H, Kent A, Clissa P, Jablensky AV. A
746 multivariate electrophysiological endophenotype, from a unitary cohort, shows greater research
747 utility than any single feature in the Western Australian family study of schizophrenia. *Biol*
748 *Psychiatry.* 2006;60(1):1-10.
- 749 72. Hong LE, Summerfelt A, Mitchell BD, O'Donnell P, Thaker GK. A shared low-
750 frequency oscillatory rhythm abnormality in resting and sensory gating in schizophrenia. *Clin*
751 *Neurophysiol.* 2012;123(2):285-92.
- 752 73. Narayanan B, O'Neil K, Berwise C, Stevens MC, Calhoun VD, Clementz BA, Tamminga
753 CA, Sweeney JA, Keshavan MS, Pearlson GD. Resting state electroencephalogram oscillatory
754 abnormalities in schizophrenia and psychotic bipolar patients and their relatives from the bipolar
755 and schizophrenia network on intermediate phenotypes study. *Biol Psychiatry.* 2014;76(6):456-
756 65.
- 757 74. Narayanan B, Soh P, Calhoun VD, Ruano G, Kocherla M, Windemuth A, Clementz BA,
758 Tamminga CA, Sweeney JA, Keshavan MS, Pearlson GD. Multivariate genetic determinants of
759 EEG oscillations in schizophrenia and psychotic bipolar disorder from the BSNIP study. *Transl*
760 *Psychiatry.* 2015;5:e588.
- 761 75. Klimesch W. EEG alpha and theta oscillations reflect cognitive and memory
762 performance: a review and analysis. *Brain Res Brain Res Rev.* 1999;29(2-3):169-95.
- 763 76. Smit DJA, Wright MJ, Meyers JL, Martin NG, Ho YYW, Malone SM, Zhang J, Burwell
764 SJ, Chorlian DB, de Geus EJC, Denys D, Hansell NK, Hottenga JJ, McGue M, van Beijsterveldt
765 CEM, Jahanshad N, Thompson PM, Whelan CD, Medland SE, Porjesz B, Lacono WG,
766 Boomsma DI. Genome-wide association analysis links multiple psychiatric liability genes to
767 oscillatory brain activity. *Hum Brain Mapp.* 2018;39(11):4183-95.

- 768 77. Galderisi S, Mucci A, Buchanan RW, Arango C. Negative symptoms of schizophrenia:
769 new developments and unanswered research questions. *Lancet Psychiatry*. 2018;5(8):664-77.
- 770 78. Moritz S, Silverstein SM, Dietrichkeit M, Gallinat J. Neurocognitive deficits in
771 schizophrenia are likely to be less severe and less related to the disorder than previously thought.
772 *World Psychiatry*. 2020;19(2):254-5.
- 773 79. Ventura J. Computer-based virtual reality assessment of functional capacity in primary
774 psychosis. *World Psychiatry*. 2022;21(3):464-5.
- 775 80. Klingberg T, Judd N, Sauce B. Assessing the impact of environmental factors on the
776 adolescent brain: the importance of regional analyses and genetic controls. *World Psychiatry*.
777 2022;21(1):146-7.
- 778 81. Lysaker PH, Hasson-Ohayon I. Metacognition in psychosis: a renewed path to
779 understanding of core disturbances and recovery-oriented treatment. *World Psychiatry*.
780 2021;20(3):359-61.
- 781 82. Murray CJ, Lopez AD. Evidence-based health policy--lessons from the Global Burden of
782 Disease Study. *Science*. 1996;274(5288):740-3.
- 783 83. Malla A, Boksa P, Joober R. Meeting the challenges of the new frontier of youth mental
784 health care. *World Psychiatry*. 2022;21(1):78-9.
- 785 84. Danese A. Bridging between youth psychiatry and child and adolescent psychiatry.
786 *World Psychiatry*. 2022;21(1):83-5.
- 787
788
789
790

791 **Tables**

792

793 Table 1. Socio-demographic, illness-related and real-life functioning variables at baseline.

	HCs (N=57)	SCZs (N=113)	t/X2	p
Age (mean ± SD)	34.56 ± 12.58	36.34 ± 9.16	1.05	0.30
Gender (M/F)	28/29	80/33	7.02	0.008*
Education (mean ± SD)	14.14 ± 4.15	12.18 ± 3.04	-3.49	< 0.001*
Duration of illness (mean ± SD)		12.75 ± 8.29		
PANSS positive (mean ± SD)		7.88 ± 4.31		
PANSS negative (mean ± SD)		15.58 ± 5.96		
PANSS disorganization (mean ± SD)		8.56 ± 3.52		
BNSS total score (mean ± SD)		34.88 ± 16.21		
BNSS — Experiential Domain (mean ± SD)		21.17 ± 8.81		
BNSS — Expressive deficit (mean ± SD)		11.41 ± 7.39		
CDSS total score (mean ± SD)		3.31 ± 4.00		
SHRS – Parkinsonism (mean ± SD)		0.79 ± 1.13		
Neurocognitive composite score (mean ± SD)	51.17 ± 9.98	29.85 ± 12.04	-10.13	< 0.001*
Overall composite score (mean ± SD)	49.28 ± 9.29	27.94 ± 11.93	-9.53	< 0.001*
SLOF – Physical functioning (mean ± SD)	24.85 ± 0.40	24.48 ± 1.08	-2.51	0.01
SLOF – Personal care skills (mean ± SD)	34.98 ± 0.13	32.44 ± 3.49	-5.40	< 0.001*
SLOF – Interpersonal relationships (mean ± SD)	33.87 ± 2.14	23.35 ± 5.88	-12.84	< 0.001*
SLOF – Social acceptability (mean ± SD)	34.91 ± 0.40	32.27 ± 3.67	-5.32	< 0.001*
SLOF – Everyday life skills (mean ± SD)	54.80 ± 0.66	46.89 ± 6.86	-8.44	< 0.001*
SLOF – Work skills (mean ± SD)	28.71 ± 2.10	20.86 ± 5.96	-9.47	< 0.001*

794

795 SCZs – Patients with Schizophrenia, HCs – Healthy Controls, PANSS – Positive and Negative
796 Syndrome Scale, BNSS – Brief Negative Symptom Scale, CDSS – Calgary Depression Scale for
797 Schizophrenia, SHRS – St. Hans Rating Scale, SLOF – Specific Level of Functioning scale

798

799 *significant t-test after Bonferroni-Holm correction

800

801 Table 2. Differences in baseline variables between subjects included and not included in follow-
 802 up study.

803

	FU included (N=61)	FU not-included (N=52)	t/X2	p
Age (mean ± SD)	36.70 ± 9.16	35.90 ± 9.24	0.46	0.65
Gender (M/F)	43/18	37/15	5.82	0.02
Education (mean ± SD)	12.31 ± 3.00	12.02 ± 3.11	0.50	0.61
Duration of illness (mean ± SD)	12.95 ± 8.58	12.45 ± 7.94	0.30	0.77
PANSS positive (mean ± SD)	8.07 ± 4.80	7.66 ± 3.67	0.49	0.62
PANSS negative (mean ± SD)	15.70 ± 5.57	15.42 ± 6.47	0.25	0.80
PANSS disorganization (mean ± SD)	8.48 ± 3.40	8.66 ± 3.68	-0.27	0.78
BNSS total score (mean ± SD)	34.75 ± 16.22	35.04 ± 16.37	-0.09	0.93
BNSS — Experiential Domain (mean ± SD)	21.11 ± 9.16	21.24 ± 8.44	-0.08	0.94
BNSS — Expressive deficit (mean ± SD)	11.21 ± 7.07	11.65 ± 7.84	-0.31	0.76
CDSS total score (mean ± SD)	3.70 ± 4.07	2.82 ± 3.90	1.16	0.25
SHRS – Parkinsonism (mean ± SD)	1.08 ± 1.26	0.43 ± 0.82	3.15	0.0021*
Neurocognitive composite score (mean ± SD)	29.98 ± 12.88	29.67 ± 10.91	0.13	0.90
Overall composite score (mean ± SD)	28.13 ± 12.36	27.41 ± 10.91	0.24	0.81
SLOF – Physical functioning (mean ± SD)	24.61 ± 0.74	24.31 ± 1.39	1.42	0.16
SLOF – Personal care skills (mean ± SD)	32.51 ± 3.77	32.35 ± 3.15	0.24	0.81
SLOF – Interpersonal relationships (mean ± SD)	23.23 ± 5.71	23.50 ± 6.13	-0.24	0.81
SLOF – Social acceptability (mean ± SD)	32.15 ± 3.74	32.42 ± 3.61	-0.39	0.70
SLOF – Everyday life skills (mean ± SD)	47.13 ± 6.73	46.60 ± 7.07	0.40	0.69
SLOF – Work skills (mean ± SD)	20.26 ± 6.24	21.58 ± 5.57	-1.16	0.25

804

805 SCZs – Patients with Schizophrenia, HCs – Healthy Controls, PANSS – Positive and Negative
 806 Syndrome Scale, BNSS – Brief Negative Symptom Scale, CDSS – Calgary Depression Scale for
 807 Schizophrenia, SHRS – St. Hans Rating Scale, SLOF – Specific Level of Functioning scale

808

809 *significant t-test after Bonferroni-Holm correction

810

811 Table 3. Differences in variables measured at baseline and follow-up.

	Baseline (N=61)	Follow-up (N=61)	t/X2	p
PANSS positive (mean ± SD)	8.07 ± 4.80	6.54 ± 3.51	2.69	0.009
PANSS negative (mean ± SD)	15.70 ± 5.57	12.74 ± 6.79	2.86	0.006e
PANSS disorganization (mean ± SD)	8.48 ± 3.40	6.31 ± 3.30	4.22	< 0.001*
BNSS total score (mean ± SD)	34.75 ± 16.22	24.05 ± 16.98	4.21	< 0.001*
BNSS – Experiential Domain (mean ± SD)	21.11 ± 9.16	14.23 ± 9.20	4.77	< 0.001*
BNSS — Expressive deficit (mean ± SD)	11.21 ± 7.07	8.52 ± 7.38	2.42	0.019
CDSS total score (mean ± SD)	3.70 ± 4.07	2.11 ± 3.31	2.38	0.02
Parkinsonism (mean ± SD)	1.08 ± 1.26	0.52 ± 1.06	3.39	0.001*
Neurocognitive composite score (mean ± SD)	29.98 ± 12.88	33.93 ± 14.80	-2.82	0.07
Overall composite score (mean ± SD)	28.13 ± 12.36	33.66 ± 14.25	-4.14	< 0.001*
SLOF – Physical functioning (mean ± SD)	24.61 ± 0.74	24.56 ± 0.92	0.32	0.75
SLOF – Personal care skills (mean ± SD)	32.51 ± 3.77	32.39 ± 3.45	0.22	0.82
SLOF – Interpersonal relationships (mean ± SD)	23.23 ± 5.71	22.48 ± 6.85	0.74	0.46
SLOF – Social acceptability (mean ± SD)	32.15 ± 3.74	31.23 ± 4.09	1.45	0.15
SLOF – Everyday life skills (mean ± SD)	47.13 ± 6.73	48.31 ± 7.91	-1.23	0.23
SLOF – Work skills (mean ± SD)	20.26 ± 6.24	20.85 ± 6.23	-0.67	0.51

812

813 SCZs – Patients with Schizophrenia, HCs – Healthy Controls, PANSS – Positive and Negative
814 Syndrome Scale, BNSS – Brief Negative Symptom Scale, CDSS – Calgary Depression Scale for
815 Schizophrenia, SHRS – St. Hans Rating Scale, SLOF – Specific Level of Functioning scale

816

817 *significant t-test after Bonferroni-Holm correction

818

Table 4. Classification performance (SCZs vs HCs) of Machine Learning models.

Classification SCZs vs HC	Number of variables	TN	TP	FN	FP	Sensitivity	Specificity	Balanced Accuracy	PPV	NPV	NND	PLR	Diagnostic odds ratio	P value
Frequency bands	290	40	82	31	17	72.6	70.2	71.4	82.8	56.3	2.3	2.4	5.9	<0.001
Microstates	43	29	73	40	28	64.6	50.9	57.7	72.3	42.0	6.5	1.3	1.7	0.03
MMN – P3a	40	33	85	28	24	75.2	57.9	66.6	78.0	54.1	3.0	1.8	3.2	0.03
N100 – P300	24	38	85	28	19	75.2	66.7	70.9	81.7	57.6	2.4	2.3	5.1	<0.001
Stacking-based classifier	/	40	91	22	17	80.5	70.2	75.4	84.3	64.5	2.0	2.7	7.3	<0.001

TN=True Negative; TP=True Positive; FN=False Negative; FP=False Positive; PPV=Positive Predictive Value; NND=Number Needed to Diagnosis; PLR=Positive likelihood ratio

Table 5. Correlations between classifier decision scores and Non-Negative Matrix Factorization (NNMF) factor scores at follow-up in SCZs.

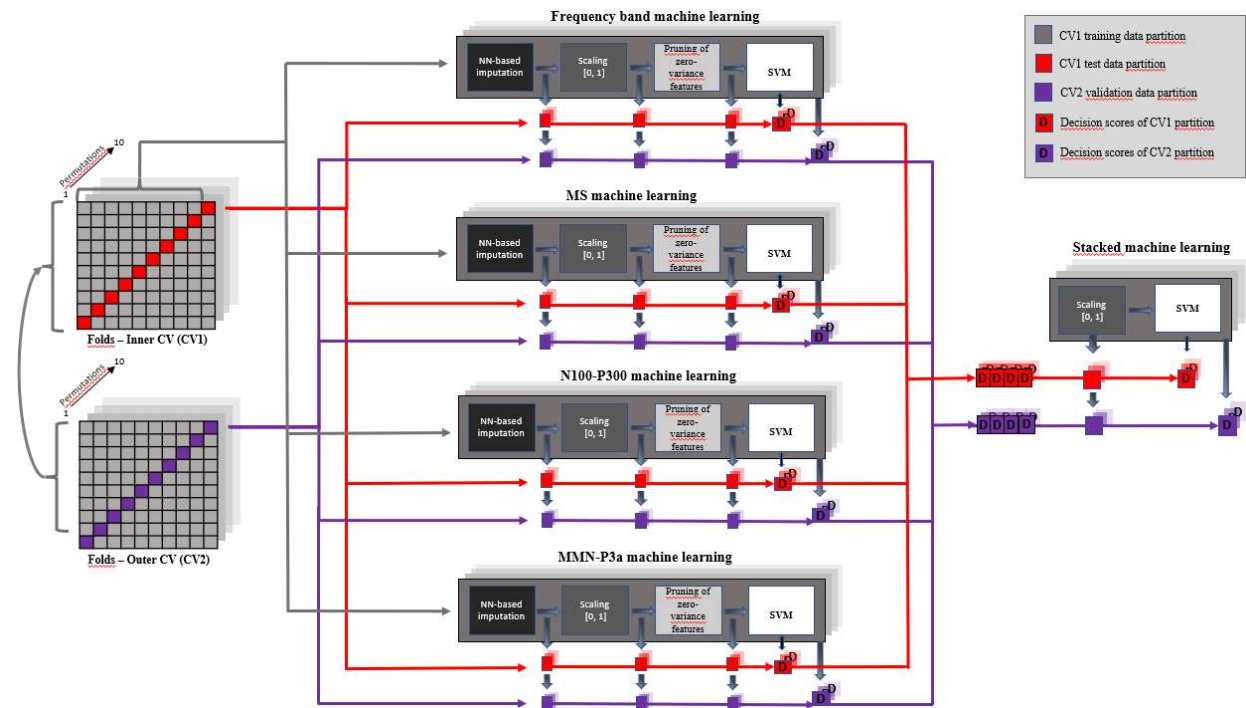
Classifiers' decision scores	Positive symptoms and parkinsonism (r;p)	Negative symptoms (r;p)	Depression (r;p)	Functioning and cognitive disturbances (r;p)
Global classifier	0.014; 0.937	0.399; 0.002*	0.429; <0.001*	-0.332; 0.009*
Frequency bands classifier	-0.018; 0.890	0.271; 0.034*	0.435; <0.001*	-0.229; 0.077
Microstates classifier	-0.132; 0.311	0.092; 0.479	0.282; 0.028	-0.020; 0.880
MMN-P3a classifier	0.120; 0.361	0.341; 0.011*	0.110; 0.399	-0.262; 0.041
N100-P3b classifier	-0.007; 0.955	0.210; 0.104	0.082; 0.530	-0.179; 0.168

*p-value survived correction for multiple tests ($p < 0.013$)

Figures

Figure 1. Experimental design of the machine learning pipelines used to train and cross-validate the unimodal and stacked classifiers.

We used nested, repeated cross-validation to train and validate the 4 individual machine learning classifiers, consisting of an outer 10-fold cross-validation cycle (CV2), which provided validation participants for computing an unbiased estimate of predictor generalisability to new patients, and an inner 10-fold cross-validation cycle (CV1), which delivered training participants to the multivariate pattern analysis pipeline as well as test participants for feature and parameter optimisation. The same nested cross-validation structure was applied to the stacked machine learning classifier, obtained combining unimodal classifiers' outputs within the machine learning environment.



Abbreviations: CV=Cross-validation, SVM=Support Vector Machine, NN=Nearest neighbor

Figure 2. Projection of illness-related and functioning variables, measured at baseline (left) and follow-up (right), to four factors, using Non-Negative Matrix Factorization.

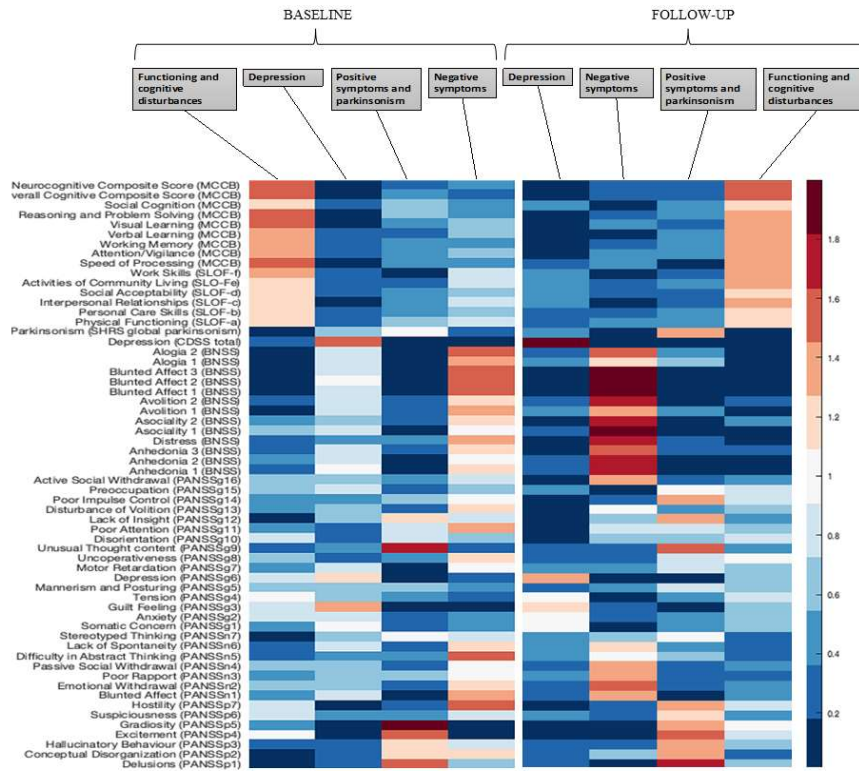


Figure 3. Composition of predictive variable sets selected by the unimodal machine learning classifiers: frequency bands (A), microstates (B), MMN-P3a (C) and N100-P3b (D).

The features were first ranked according to the selection probability measured across all inner-cycle training partitions. Variables ranking among the top 10% of selected features were marked with red and listed with their selection probability (*psel*) and correlation with the classifier's outcome (Spearman's ρ).

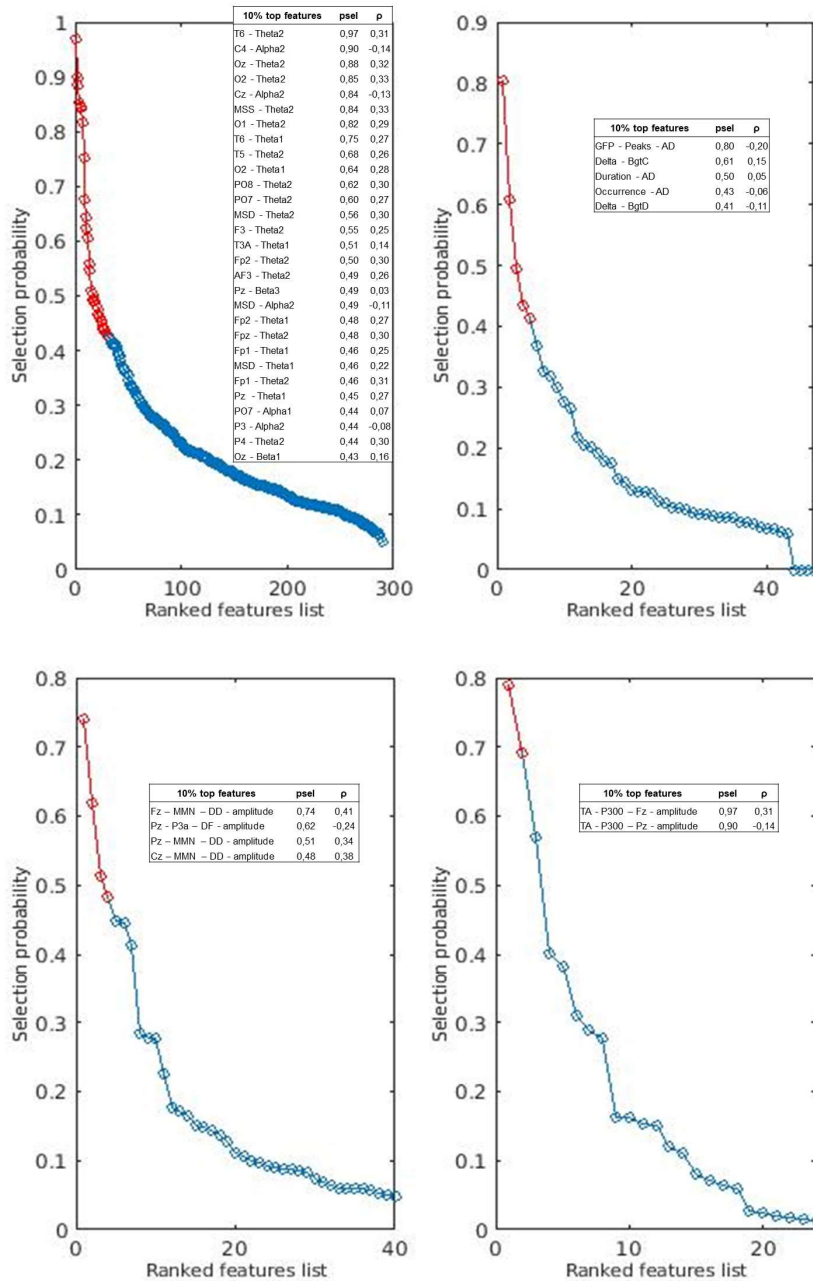


Figure 4. Contribution (Spearman's ρ) of each individual EEG data modality to the global classifier's decisions.

