The EEG multiverse of schizophrenia

Authors: Janir Ramos da Cruz[1,2,a]*, Dario Gordillo[1,a], Eka Chkonia[3,4], Wei-Hsiang Lin[1], Ophélie Favrod[1], Andreas Brand[1], Patrícia Figueiredo[2], Maya Roinishvili[4,5], and Michael H. Herzog[1]

Affiliations:

[1] Laboratory of Psychophysics, Brain Mind Institute, École Polytechnique Fédérale de Lausanne (EPFL), Switzerland

[2] Institute for Systems and Robotics – Lisbon (LARSys) and Department of Bioengineering, Instituto Superior Técnico, Universidade de Lisboa, Portugal

[3] Department of Psychiatry, Tbilisi State Medical University, Tbilisi, Georgia

[4] Institute of Cognitive Neurosciences, Free University of Tbilisi, Tbilisi, Georgia

[5] Laboratory of Vision Physiology, Beritashvili Centre of Experimental Biomedicine, Tbilisi, Georgia

[a] These authors contributed equally

*__Corresponding author:__

Janir Ramos da Cruz, Laboratory of Psychophysics, Brain Mind Institute, School of Life Sciences, École Polytechnique Fédérale de Lausanne (EFPL), CH-1015 Lausanne, Switzerland

Phone number: +41 21 693 17 42

Email: janir.ramos@epfl.ch

Abstract

Research on schizophrenia typically focuses on one paradigm, for which clear-cut abnormalities between patients and controls are established. Great care is taken to understand the underlying genetical, neurophysiological, and cognitive mechanisms, which eventually may explain the clinical outcome. This approach has led to many important hypotheses about the causes of schizophrenia. One tacit assumption of these *deep rooting* approaches is that paradigms tap into common and representative characteristics of the disease. Here, we analyzed resting-state electroencephalogram (EEG) of 121 schizophrenia patients and 75 age-matched controls, from which we extracted 194 features. Sixty-nine of these features showed significant group differences with medium to large effect sizes, indicating important abnormalities. To understand to what extent these features tap into the same aspects of the disorder, we computed both Pearson and partial least squares correlations. Surprisingly, correlations were very low, except for very similar features, where correlations were high, suggesting that most features are sensitive to different abnormalities. Using partial least squares regression, we show that combining the 69 features increases predictability of clinical outcomes by more than a factor of 5. We propose that complementing *deep* with *shallow* rooting approaches, where many roughly independent features are extracted from one paradigm (or several paradigms), will strongly improve diagnosis and potential treatment of schizophrenia.

Introduction

Schizophrenia patients show strong abnormalities in many domains including personality, cognition, perception, and even immunology. In many experimental paradigms, the differences between patients and controls have large effect sizes, indicating that important aspects of the disease are detected. This provokes two questions: what do these abnormalities have in common, and how representative are they for the disease? For example, patients exhibit strong deficits in cognitive tasks, such as working memory tasks[1], which are attributed to abnormalities of cortico-cerebellar-thalamic-cortical circuits[2]. Patients show also diminished skin flushing in niacin skin tests[3], which is attributed to dysfunctional phospholipase A2 arachidonic acid signaling[4]. How do these working memory deficits correspond to deficits in skin functioning? Very few studies have correlated deficits with each other[5–10]. The Consortium on the Genetics of Schizophrenia studied neurocognitive and neurophysiological abnormalities in schizophrenia patients with a battery of 15 paradigms[9]. They found that neurocognitive measures shared a significant amount of variance while neurophysiological measures were almost entirely independent. Price and colleagues[8] studied four candidate electrophysiological endophenotypes of schizophrenia (mismatch negativity, P50, P300, and antisaccades). Even though patients *and* their family members showed deficits in each of these endophenotypes, the features were largely uncorrelated. Here, we took another road. Instead of comparing different paradigms, we analyzed the very same data of the very same patients with different electroencephalogram (EEG) analysis methods, including many that have shown strong atypical patterns in patients[11–18].

Results

For 121 patients (22 females, 35.8 ± 9.2 years old, 13.3 ± 2.6 years of education) and 75 age-matched healthy controls (39 females, 35.1 ± 7.7 years old, 15.1 ± 2.9 years of education; Table 1), we recorded 64-channel resting-state EEG for 5 minutes. During this period participants had their eyes closed, and did not engage in any task. We extracted in total 194 features from the EEG recordings, including time-domain features, frequency-domain and connectivity features, both in electrode and source space, and nonlinear dynamical features (Supplementary Table 1). Among the 194 EEG features, 69 showed significant differences between patients and controls with medium to large effect sizes (Cohen's d varied from 0.463 to 1.037, Figure 1).

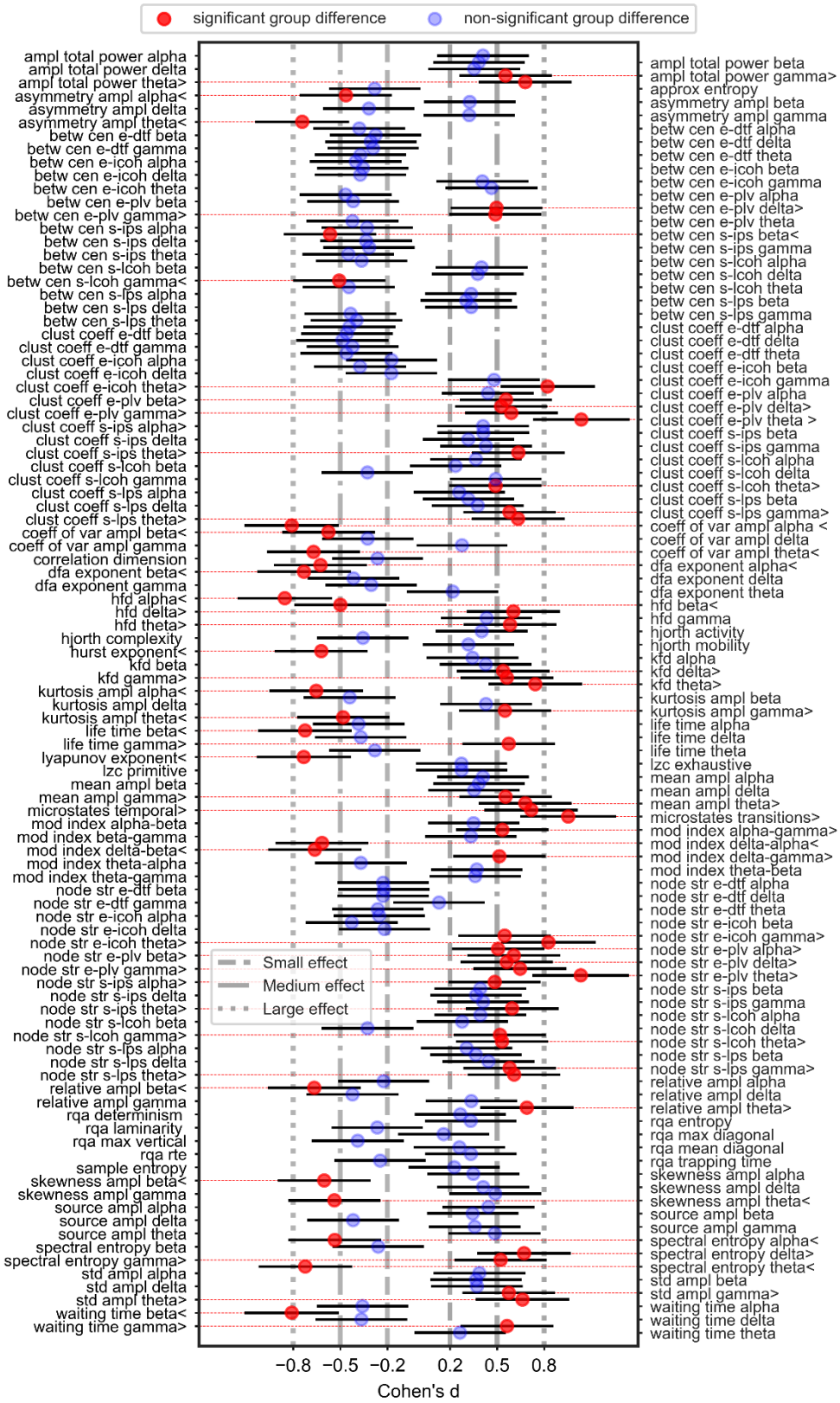

Figure 1. Effect size (Cohen's d) of the group differences between patients and controls for each of the 194 EEG features. We tested 157 features at the electrode level (each computed for 64 electrodes), 35 features at the source level (each computed for 80 brain regions), and 2 microstate features (each computed for 12 parameters). For each

feature, we conducted group comparisons with ANCOVAs (with gender as a factor and education as a co-variate), for each electrode, brain region, or microstate parameter. Multiple comparisons were corrected with false discovery rate (FDR). For each feature, we then took the values of the electrode, brain region, or microstate parameter with the largest effect size according to Cohen's d (η^2 values were converted to Cohen's d) to be the representative variable for this feature. Positive Cohen's d values mean that patients show higher feature values than controls, while negative Cohen's d values mean that controls show higher feature values than patients. Significant group differences, after correction for multiple comparisons, are depicted in red, with dotted red vertical lines serving as a guide to their labels. > and < were added to the feature labels to indicate if patients had significantly higher or lower values than controls, respectively. The non-significant effects are shown in blue. A list with the abbreviations and the corresponding name of each feature is presented in Supplementary Table 1. Error bars represent 95% confidence intervals (C.I.).

Next, we wondered to what extent features that showed significant group differences are sensitive to the same aspects of the disorder. For this, we first computed Pearson's correlations between pairs of features (Figure 2). As the representative variable for each feature, we took the values of the electrode, brain region, or microstate parameter that showed the largest group difference according to Cohen's d (Figure 1). Surprisingly, we found that in the patients group only 36.02% of the pairwise correlations were significant at a level of 0.05 (without correcting for multiple comparisons). For the control group, only 26.64% of the correlations were significant. Since significance depends on the sample size, here, we focus on the magnitude of the correlation coefficients (r-values). In general, the magnitudes of the r-values were very low in both patients (0.055, 0.123, 0.250, for the 25th, 50th, and 75th percentiles, respectively) and controls (0.058, 0.130, 0.243, for the 25th, 50th, and 75th percentiles, respectively; Figure 2). High correlations were found mainly for pairs of very closely related features, such as waiting-time statistics of gamma bursts (*waiting time gamma*) and life-time statistics of gamma bursts (*life time gamma*; r=0.836 and r=0.926, in patients and controls, respectively).

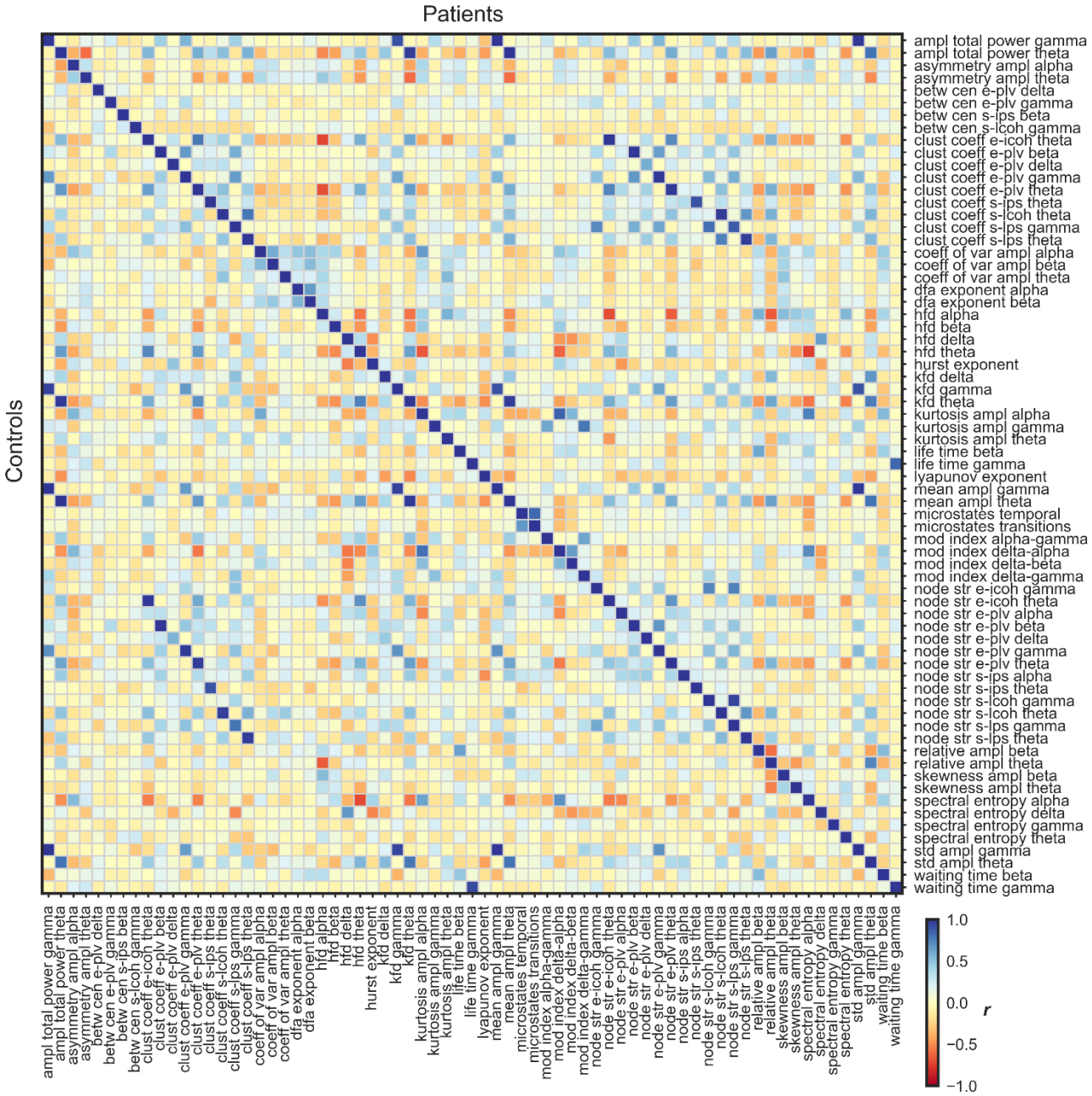

Figure 2. Pairwise correlations between the 69 EEG features that showed significant group differences between patients and controls. Patients' r-values are presented in the upper triangle and controls' r-values are shown in the lower triangle. Strong negative and positive r-values are depicted in red and blue, respectively, and r-values around 0 in yellow. For each feature, we used the values of the electrode, brain region, or microstate parameter that showed the largest effect size as the representative variable for the correlations. In general, both patients and controls showed low pairwise correlations between the different EEG features. A list with the abbreviations and corresponding name of each feature is shown in Supplementary Table 1.

Next, we wanted to have a closer look at the overall *shared information* between pairs of EEG features, which showed significant group differences, by taking not only variables with the largest effect size into account but all variables of the features. To this end, we used partial least squares correlation (PLSC), which is a multivariate statistical technique widely used in neuroscience[19]. PLSC generalizes the principle of correlation between two variables to the *correlation* between

two matrices (here, pairs of EEG features)[20–22]. In brief, PLSC is based on the singular value decomposition of a matrix containing the dot product of two matrices comprising the normalized variables of features $\mathbf{X}$ and $\mathbf{Z}$, respectively. The *shared information* between the two features is given by the inertia ($\mathfrak{I}$), which is the sum of the singular values. The higher the inertia, the higher the amount of *shared information.* We normalized the inertias by the square-root of the product of the number of variables in the features $\mathbf{X}$ and $\mathbf{Z}$ (relative inertia; $\mathfrak{I}_{relative}$). In this case, the relative inertias range from 0 (the two features are completely unrelated) to 1 (the two features move together with a fixed proportion), which is akin to a correlation coefficient. We assessed the statistical significance of the inertia using permutation tests[20]. For details, see section Statistical Analyses. In patients, 56.05% of the pairwise inertias were significant (without correcting for multiple comparisons) and for controls, 40.32%. In general, relative inertias were not very high in both patients (0.255, 0.329, 0.409, for the 25th, 50th, and 75th percentiles, respectively) and controls (0.306, 0.388, 0.472, for the 25th, 50th, and 75th percentiles, respectively; Figure 3). As in the Pearson's correlation results, features that showed high associations were mainly similar features, such as the same network statistics for different connectivity measures in the theta band, for example, at the electrode level: clustering coefficient connectivity estimated with phase locking value (*clust coeff e-plv theta*) and with imaginary part of coherence (*clust coeff e-icoh theta*; $\mathfrak{I}_{relative}$=0.804 and $\mathfrak{I}_{relative}$=0.826, in patients and controls, respectively).

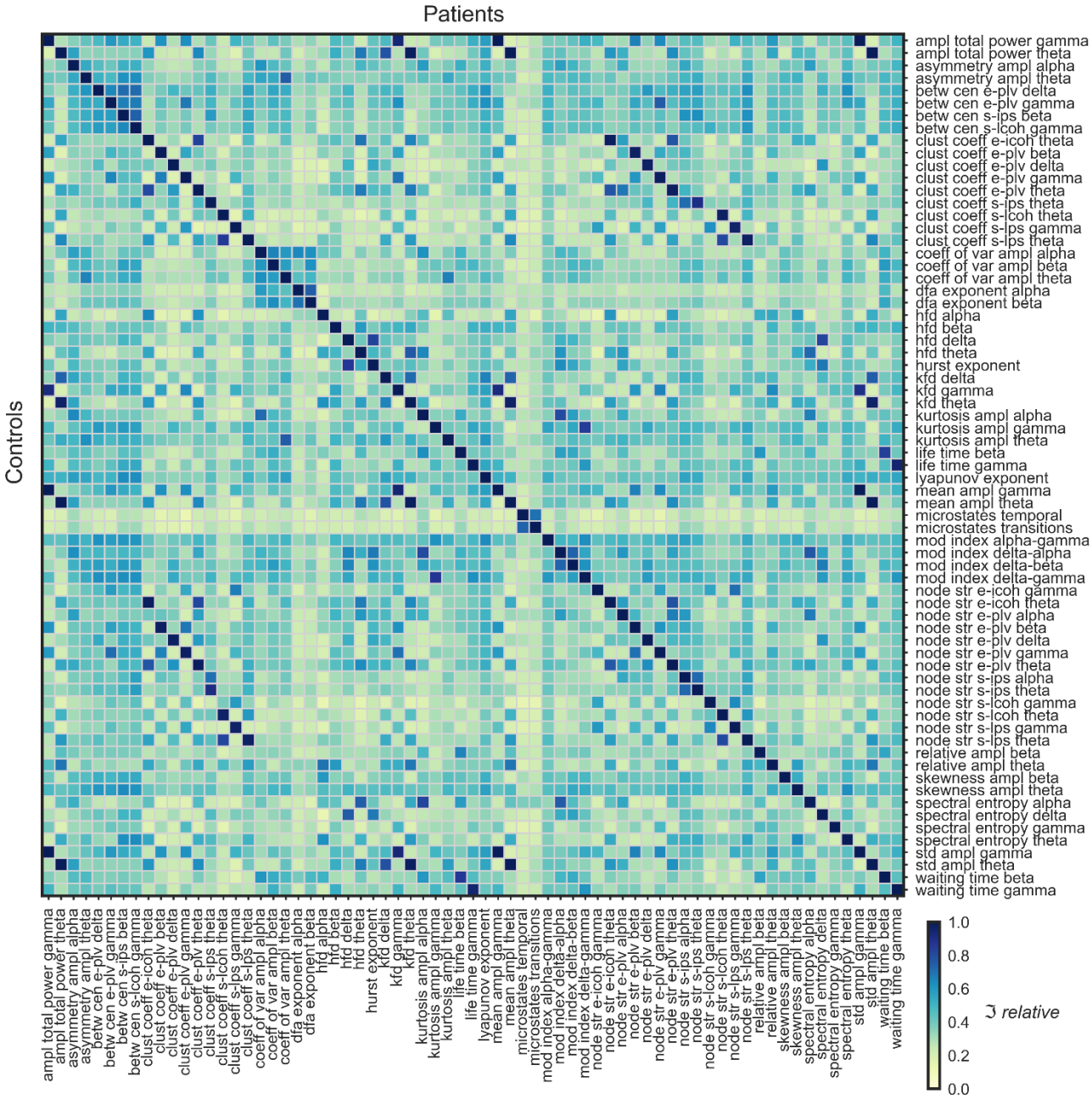

Figure 3. *Shared information* between the 69 EEG features that showed significant group differences, as measured by the relative inertia ($\Im_{relative}$) computed with partial least squares correlations (PLSC). The relative inertia ranges from 0 (two features are completely unrelated) to 1 (the two features' values move together by the exact same percentage). Patients' relative inertias are presented in the upper triangle and controls' relatives are shown in the lower triangle. In general, both patients and controls, showed relatively low relative inertias. A list with the abbreviations and corresponding name of each feature is shown in Supplementary Table 1.

The next question was whether combining EEG features can improve predictability of clinical outcomes. We used partial least squares regression (PLSR) to predict the clinical outcomes as determined by the Scale for the Assessment of Positive Symptoms (SAPS) and the Scale for the Assessment of Negative Symptoms (SANS), which target positive (hallucinations, delusions, bizarre behavior, and positive formal thought disorder) and negative (affective flattening, alogia, apathy, anhedonia, and attention) symptoms, respectively. PLSR is an alternative to linear regression in situations where the

number of predictors is relatively large compared to the number of variables and when the predictors are highly correlated[19,20]. PLSR projects both the predicted variables (stored in $\mathbf{Y}$) and the predictors (stored in $\mathbf{X}$) into a new space formed by latent variables that simultaneously model $\mathbf{X}$ and predict $\mathbf{Y}$. PLSR benefits from dimensionality reduction since it only uses a few latent variables. For details see section Statistical Analyses. In general, each of the 69 EEG features showed relatively weak predictive power of the clinical outcomes as determined by the Pearson's correlation and the root mean square error (RMSE) calculated between the observed and predicted scores (SANS: r=0.163$\pm$0.117, RMSE=5.097$\pm$0.124; SAPS: r=0.146$\pm$0.108, RMSE=3.111$\pm$0.072; Supplementary Tables 2 and 3). Since our results suggest that the EEG features are largely independent or weakly correlated, we expected that a combination of features would greatly improve the prediction of SAPS and SANS. We found that, indeed, the *multiverse* of EEG features led to a much higher predictive power (SANS: r=0.837, RMSE=2.835; SAPS: r=0.832, RMSE=1.752), namely a 5.114 and 5.711 times increase compared to the mean r-values and a 1.798 and 1.776 times decrease compared to the mean RMSE values, for SANS and SAPS, respectively.

Discussion

Traditionally, schizophrenia research focuses on a single experimental paradigm and analysis method showing significant differences between patients and controls, and then tries to derive the underlying genetic or neurophysiological causes of the disorder. This approach has been quite successful in the formulation of hypotheses, such as the dopamine hypothesis[23], the social brain hypothesis[24], the glutamate hypothesis[25], or the dysconnection hypothesis[26], just to name a few. Here, we examined to what extent abnormalities quantified by different EEG analysis features correlate with each other. Many of the investigated features were previously linked to different abnormalities of brain processes in schizophrenia, and, here, we reproduced many of these results, such as imbalance in microstates dynamics[13,27], decreased long-range temporal correlation in the alpha and beta bands[16], decreased life- and waiting-times in the beta band[17], increased spectral amplitude in the theta band[12], increased connectivity in the theta band at the source level[11,14], decreased Lyapunov exponent[15], among others. With our systematic analysis, we found also abnormalities in EEG features, which, to the best of our knowledge, have not been reported yet, namely, delta-phase gamma-amplitude coupling, delta-phase alpha-amplitude coupling, range EEG coefficient of variation and asymmetry in the theta and alpha bands, etc. In some way, deeper analysis of each feature may have warranted an in-depth study and a potential publication. However, we did not want to elaborate on these methods individually because we wanted to understand how all EEG features relate to each other in their entirety. The surprising insight from our analysis is that, even though we are probing the same signals from the same participants, we found only weak correlations between the 69 significant features. The only high correlations were between features that are similar from the outset, thereby resembling test-retests. This suggests that none of the features is truly representative for the disease, but rather that all these features pick up more or less independent aspects of schizophrenia. Hence, the traditional approach of focusing on a single experimental paradigm and analysis method has its limitations. These results remind us that schizophrenia is indeed a very heterogeneous disease, a well-known fact, which is however not always taken seriously enough because, as mentioned above, most research tries to find the one or a few causes of schizophrenia within one well described paradigm by digging as deep as possible into the underlying neurophysiological and genetic mechanisms. In analogy to botany, one may call these approaches "deep rooting" approaches.

We propose that it may be useful to complement these deep rooting approaches with "shallow rooting" approaches, representing schizophrenia within a high-dimensional space, where many tests and analysis outcomes are the basis variables. The outcomes should ideally have large effect sizes, low mutual correlations, and a "flat" factor structure. Whether this is possible is an open question and depends very much on the underlying causes of schizophrenia. On the lowest complexity level, there may be only a few independent causes (or even only one), which were not found yet. Given

the heterogeneity of the disease, including abnormalities in the cognitive[2] but also the skin functioning domain[4], the causes need to be on a rather general level, likely subcellular, present in all human functioning. Alternatively, schizophrenia may be an approximatively "additive" disease, where many small abnormalities add up to severe symptoms. In an even more complex scenario, only certain combinations of redundant functions, each coming with at least two variants, cause the disease. For instance, if one function is up-regulated and another one down-regulated in an individual, there are no abnormalities. Deficits manifest only when all or most functions are either up- or down-regulated. In such a combinatorial scenario, it would be difficult to find the underlying causes since each variant itself does not lead to a deficit; only certain combinations do. Our correlation analyses (Figure 2 and Figure 3) provide some evidence for the additive scenario. This may be good news because combining measures can provide much better predictions on clinical outcomes than relying on a single feature (Supplementary Table 2). By combining all 69 features, we could indeed increase predictability of the positive and negative symptoms from 0.434 (for the best feature) to 0.837 (all 69 features) – notably with the very same resting-state paradigm, thus reducing the testing burden for patients. In the next steps, it will be important to find the right set of tests, which may include cognitive tests but also potentially immunological markers - of which each may contribute with a variety of analysis methods.

Our results are in line with recent results in magnetic resonance imaging[28] and genetics[29] studies, which have shown that combining measures, such as in polygenic risk scores, can largely improve predictions[28]. However, even though we used cross-validation techniques to avoid overfitting, we cannot exclude that the PLSR results are too optimistic since we do not have an independent testing set. Also, there are demographic differences between patients and controls, which might affect our group comparisons. However, we attempted to minimize these demographic effects by using education as a covariate and gender as factor in the analyses. Similarly, we cannot exclude effects of medication in our results. Nonetheless, we find similar patterns of correlations between EEG features, i.e., weak associations, in both patients and controls, suggesting that if there is an effect of medication, it is small.

Our results may explain a deep mystery in schizophrenia research. Schizophrenia has an estimated heritability of 70 to 85%[30]. For example, the chance to also suffer from schizophrenia for monozygotic twins is about 33% when the partner twin has the disease[31]. Furthermore, about 0.25 to 0.75% people of a population suffer from schizophrenia and related psychotic disorders[32–34]. These values are rather stable across cultures[35]. Given that schizophrenia patients have less offspring[36–39], this provokes the question why schizophrenia has not been extinguished during the course of evolution[38,40]. In the above-mentioned combinatorial scenario with many redundant functions this may simply happen because evolution operates on the individual single-nucleotide polymorphism (SNP) level and not on the combinatorial one. As long as most of the population shows average functioning, there will be no change of the allele distributions. In the additive scenario, evolution may extinct harmful alleles, of which each constitutes only a little risk, very slowly and these may be replaced by harmful de novo mutations[38]. To what extent such considerations hold true will be shown by shallow rooting approaches using a plethora of paradigms and a multiverse of analysis methods. In a nutshell, deep rooting will help to understand the different aspects of the disorder, while shallow rooting will help to better diagnose schizophrenia by finding subpopulations, leading to more personalized treatment.

Methods

Participants

Two groups of participants joined the experiment: schizophrenia patients (n=121) and healthy controls (n=75). All participants took part in a battery of tests comprising perceptual and cognitive tasks as well as EEG recordings. Data of 101 patients and 75 controls have already been published in different contexts[13,41–43], while data of the other 20 participants have not been published yet. Patients were recruited from the Tbilisi Mental Health Hospital or the psycho-

social rehabilitation center. Patients were invited to participate in the study when they had recovered sufficiently from an acute psychotic episode. Thirty-five were inpatients and 86 were outpatients. Patients were diagnosed using the Diagnostic and Statistical Manual of Mental Disorders Fourth Edition (DSM-IV) by means of an interview based on the Structured Clinical Interview for DSM-IV, Clinical Version, information from staff, and study of patients' records. Psychopathology of patients was assessed by an experienced psychiatrist using the Scale for the Assessment of Negative Symptoms (SANS) and the Scale for the Assessment of Positive Symptoms (SAPS). Out of the 121 patients, 106 were receiving neuroleptic medication. Chlorpromazine (CPZ) equivalents are indicated in Table 1. Controls were recruited from the general population in Tbilisi, aiming to match patients' demographics as closely as possible. All controls were free from psychiatric axis I disorders and had no family history of psychosis. General exclusion criteria were alcohol or drug abuse, severe neurological incidents or diagnoses, developmental disorders (autism spectrum disorder or intellectual disability), or other somatic mind-altering illnesses, assessed through interview by certified psychiatrists. All participants were no older than 55 years. Group characteristics are presented in Table 1. Since patients and controls differed in terms of gender and education, gender was used as factor while education was used as a covariate in subsequent group comparisons.

All participants signed informed consent and were informed that they could quit the experiment at any time. All procedures complied with the Declaration of Helsinki (except for pre-registration) and were approved by the Ethical Committee of Institute of Postgraduate Medical Education and Continuous Professional Development (Georgia). Protocol number: 09/07. Title: "Genetic polymorphisms and early information processing in schizophrenia".

Table 1 - Group average statistics (±SD)

	Patients	Controls	Statistics
Gender (F/M)	22/99	39/36	χ^2 (1)=24.702, p=6.690e-7[b]
Age (years)	35.8 ± 9.2	35.1 ± 7.7	t(194)=0.519, p=0.604[c]
Education (years)	13.3 ± 2.6	15.1 ± 2.9	t(194)=-4.418, p=1.657e-5[c]
Handedness (L/R)	6/115	4/71	χ^2 (1)=0.013, p=0.908[b]
Illness duration (years)	10.8 ± 8.7		
SANS	10.1 ± 5.2		
SAPS	8.6 ± 3.2		
CPZ equivalent[a]	561.1 ± 389.4		

SANS - Scale for the Assessment of Negative Symptoms, SAPS - Scale for Assessment of Positive, CPZ - chlorpromazine
[a]Average CPZ equivalents calculated over the 106 Patients receiving neuroleptic medication
[b]Pearson's chi-squared test
[c]Two-sided independent samples *t*-test

EEG recording and data processing

Participants were sitting in a dim lit room. They were instructed to keep their eyes closed and to relax for 5 minutes. Resting-state EEG was recorded using a BioSemi Active Two Mk2 system (Biosemi B.V., The Netherlands) with 64 Ag-AgCl sintered active electrodes, referenced to the common mode sense electrode. The recording sampling rate was 2048 Hz. Offline data were downsampled to 256 Hz and preprocessed using an automatic pipeline[44]. See Supplementary Methods for details.

Statistical Analyses

First, we compared patients' and controls' scores for each of the 194 EEG features. For each of J variables (64 electrodes, 80 brain regions, or 12 microstate parameters depending on the number of variables of each EEG feature) of a given feature, we performed a two-way ANCOVA, with Group (patients and controls) and Gender (male and female) as factors and Education as a covariate. *P*-values for the effect of Group were corrected for J comparisons using False Discovery Rate (FDR; with an error rate of 5%). Group effects' η^2 were converted to Cohen's d.

Second, for each EEG feature that showed at least one variable with a significant effect (after correcting for multiple comparisons), we used the variable (electrode, brain region, or microstate parameter) with the biggest effect size to be the representative variable for that feature. We then used Pearson correlations to estimate the pairwise correlations between the representative variables of each of the significant EEG features. The correlation analysis was done for patients and controls separately.

Third, to quantify the overall relationship, i.e., the amount of *shared information*, between pairs of multivariate EEG features, we used Partial Least Squares Correlation (PLSC). PLSC is the generalization of the correlation between two variables to two matrices[21,22]. Let $\mathbf{X}$ be an $N \times J$ matrix, containing the data of N participants (121 patients or 75 controls) for all J variables (64 electrodes, 80 brain regions, or 12 microstate parameters) of a certain EEG feature (alpha, beta, etc.), and $\mathbf{Z}$ be an $N \times K$ matrix, containing data of the N participants for all K variables (64 electrodes, 80 brain regions, or 12 microstate parameters) of another EEG feature. With both $\mathbf{X}$ and $\mathbf{Z}$ mean-centered and normalized, the pattern of relationship between the columns of $\mathbf{X}$ and $\mathbf{Z}$ can be stored in a $K \times J$ cross-product correlation matrix, denoted $\mathbf{R}$, computed as:

$$\mathbf{R} = \mathbf{Z}^{\mathrm{T}}\mathbf{X} \tag{1}$$

The goal of PLSC is to analyze the *shared information* between $\mathbf{X}$ and $\mathbf{Z}$, which is stored in the matrix $\mathbf{R}$. This is done by deriving two sets of latent variables, one for $\mathbf{X}$ and another for $\mathbf{Z}$, that are linear combinations of the respective original variables. These latent variables are computed in order to obtain the maximal covariance between $\mathbf{X}$ and $\mathbf{Z}$. The original variables are described by their *saliences*, which are similar to loadings in principal components analysis [19]. This is achieved by the singular value decomposition (SVD) of the correlation matrix $\mathbf{R}$:

$$\mathbf{R} = \mathbf{U}\Delta\mathbf{V}^{\mathrm{T}} \tag{2}$$

where $\mathbf{U}$ is the $J \times L$ matrix of $\mathbf{X}$-saliences and $\mathbf{V}$ is the $K \times L$ matrix of $\mathbf{Z}$-saliences, while Δ is the $L \times L$ diagonal matrix of the L singular values (with L being the rank of $\mathbf{R}$).

The quantity of *shared information* between $\mathbf{X}$ and $\mathbf{Z}$ can be directly quantified as the *inertia* common to the two features[19]. The inertia, denoted $\mathfrak{I}$, is defined as:

$$\mathfrak{I} = \sum_{l=1}^{L} \delta_l \tag{3}$$

where δ_l is the l^{th} diagonal element, i.e., singular value, of Δ, and L is the number of non-zero singular values of $\mathbf{R}$, i.e., the rank of the correlation matrix.

The statistical significance of the inertia is assessed using a permutation test[20,45]. A permutation sample is created by shuffling the rows of $\mathbf{X}$ (i.e., the participants) while keeping $\mathbf{Z}$ fixed. Then PLSC is used to recompute a new value of inertia for the permutated sample. This procedure is repeated 10,000 times, which produces a null distribution of inertias that can be used for null hypothesis testing. The *p*-values are given by counting how many times the permutated inertias where larger than the original inertia and dividing by the number of permutations (10,000).

Here, since some EEG features have different numbers of variables (64 electrodes, 80 brain regions, or 12 microstates parameters), within and across the pairwise comparisons, which results in different orders of the $\mathbf{R}$ matrix, we normalized the inertias for better comparability across pairwise comparisons of EEG features. In essence, we divided the computed inertias by the square-root of the product of dimensions of the $\mathbf{R}$ matrix ($\sqrt{K \times J}$) [46], resulting in relative inertias ($\Im_{relative}$). In this case, the inertias range from 0 ($\mathbf{X}$ and $\mathbf{Z}$ are completely unrelated) to 1 ($\mathbf{X}$ and $\mathbf{Z}$ are basically the same).

Last, to predict patients' psychopathology scores (SANS and SAPS) based on the EEG features, we used Partial Least Squares Regression (PLSR). PLSR is a multivariate high-dimensional regression method that has been widely used in predicting behavioral scores based on neuroimaging data[19]. Since PLSR can handle regression problems where the number of predictors is relatively large compared to the number of samples as well as multicollinearity (i.e., when the predictors are not linearly independent) by dimensionality reduction, PLSR is a very versatile tool to study brain-behavior relationships[19,45]. The predictors (in our case, the variables of a given EEG feature) are stored in a $N \times J$ matrix $\mathbf{X}$ and the predicted variables (in our case, SANS or SAPS) are stored in a $N \times M$ matrix $\mathbf{Y}$, where N is the number of participants (121 patients), J is the number of predictors (64 electrodes, 80 brain regions, 12 microstates parameters, or the variables of all features together), and M is the number of predicted variables. Here, since we predicted SANS and SANS separately, M is 1. $\mathbf{X}$ is mean-centered and normalized. PLSR projects both the predictors and the predicted variables into a new space formed by latent variables stored in a matrix $\mathbf{T}$ that simultaneously models $\mathbf{X}$ and predicts $\mathbf{Y}$. This is expressed as a double decomposition of $\mathbf{X}$ and the predicted $\mathbf{Y}$ ($\widehat{\mathbf{Y}}$):

$$\mathbf{X} = \mathbf{T}\mathbf{P}^{\mathrm{T}} \text{ and } \widehat{\mathbf{Y}} = \mathbf{T}\mathbf{B}\mathbf{C}^{\mathrm{T}} \quad (4)$$

where $\mathbf{P}$ and $\mathbf{C}$ are loadings, while $\mathbf{B}$ is a diagonal matrix. If we let $\mathbf{B}_{\mathrm{PLS}} = \mathbf{P}^{\mathrm{T}+}\mathbf{B}\mathbf{C}$, where $\mathbf{P}^{\mathrm{T}+}$ is the pseudo-inverse of $\mathbf{P}^{\mathrm{T}}$ and $\mathbf{B}_{\mathrm{PLS}}$ is a $J \times M$ matrix equivalent to the regression weights of a multiple regression, $\widehat{\mathbf{Y}}$ can be expressed as a regression model:

$$\widehat{\mathbf{Y}} = \mathbf{X}\mathbf{B}_{\mathrm{PLS}} \quad (5)$$

In PLSR, the latent variables are computed by applying SVD iteratively. At each iteration of the SVD, orthogonal latent variables and corresponding coefficients are produced. There are many algorithms that solve the PLSR problem[47]. Here, we used the Nonlinear Iterative Partial Least Squares (NIPALS) as implemented in Scikit-learn 0.21.3[48].

Since in our study the number of predictors is relatively large compared to the number of observations, especially when we aggregate the variables of all the EEG features, PLSR will most likely overfit the data (i.e., perform well in training data but poorly in new observations). To avoid overfitting, we resorted to leave-one-out cross-validation (LOOCV)[49]. We conducted the LOOCV in two steps: 1) select the number of latent variables, and 2) select the predictors. In the LOOCV procedure, the testing set consists of only one participant. Each participant is removed from $\mathbf{X}$ and $\mathbf{Y}$, and a PLSR model is computed for the remaining participants. Then the PLSR model is used to predict the left-out participant's $\mathbf{Y}$ value from their $\mathbf{X}$ values. The N predicted values are stored in $\widetilde{\mathbf{Y}}$. The quality of the prediction on unseen data is given by the Predicted Residual Estimated Sum of Squares (PRESS), which is formally defined as[47]:

$$\mathrm{PRESS} = \left\|\mathbf{Y} - \widetilde{\mathbf{Y}}\right\|^2 \quad (6)$$

where $\|\ \ \|^2$ is the sum of squares of all elements in this matrix (the smaller the PRESS the better). The number of latent variables at which the PRESS starts increasing gives an indication of the optimal number of latent variables to be kept in the model[47]. After fixing the number of latent variables to be kept, we used recursive feature elimination (RFE)[50] with LOOCV to optimize the model performance on unseen data. Note that, here, the features in RFE are what we refer to as variables or predictors. In RFE, we build a model on the entire set of predictors to get an importance score for each predictor, in our case, the regression coefficients in the PLSR model. Then, the least important predictor, i.e., with the smallest absolute regression coefficient, is removed, the model is re-built and the importance scores are re-computed. By using PRESS as a criterion, this procedure is repeated until the PRESS starts increasing. The subset of predictors with the smallest PRESS is then used to train the final model. Finally, we reported the predictive performance of the PLSR model as the Pearson's correlation coefficient and the root-mean square error (RMSE) between the observed $\mathbf{Y}$ and the leave-one-out predicted $\widetilde{\mathbf{Y}}$.

Data Availability

The data that support the findings of this study are available upon reasonable request.

Code Availability

The code that support the findings of this study are available upon request.

Acknowledgments

This work was partially funded by the Fundação para a Ciência e a Tecnologia under grant FCT PD/BD/105785/2014 and the National Centre of Competence in Research (NCCR) Synapsy financed by the Swiss National Science Foundation under grant 51NF40-185897. We would like to thank Marc Repnow for his comments on the manuscript and Ben Lönnqvist for proofreading the manuscript.

Author Contributions

M.H.H., E.C., A.B., and M.R. designed the research; M.R. and E.C. performed the research; J.R.C., D.G., W.H.L., and O.F. analyzed data; J.R.C., D.G., O.F. A.B., P.F., and M.H.H. wrote the paper.

Competing Interests

The authors declare no competing interests.

References

1. Meyer-Lindenberg, A. *et al.* Evidence for Abnormal Cortical Functional Connectivity During Working Memory in Schizophrenia. *Am. J. Psychiatry* **158**, 1809–1817 (2001).
2. Andreasen, N. C., Paradiso, S. & O'Leary, D. S. 'Cognitive Dysmetria' as an Integrative Theory of Schizophrenia: A Dysfunction in Cortical-Subcortical-Cerebellar Circuitry? *Schizophr. Bull.* **24**, 203–218 (1998).
3. Rybakowski, J. & Weterle, R. Niacin test in schizophrenia and affective illness. *Biol. Psychiatry* **29**, 834–836 (1991).

4. Messamore, E. Niacin subsensitivity is associated with functional impairment in schizophrenia. *Schizophr. Res.* **137**, 180–184 (2012).
5. Braff, D. L., Freedman, R., Schork, N. J. & Gottesman, I. I. Deconstructing Schizophrenia: An Overview of the Use of Endophenotypes in Order to Understand a Complex Disorder. *Schizophr. Bull.* **33**, 21–32 (2007).
6. Braff, D. L., Light, G. A. & Swerdlow, N. R. Prepulse Inhibition and P50 Suppression Are Both Deficient but not Correlated in Schizophrenia Patients. *Biol. Psychiatry* **61**, 1204–1207 (2007).
7. Dickinson, D., Goldberg, T. E., Gold, J. M., Elvevag, B. & Weinberger, D. R. Cognitive Factor Structure and Invariance in People With Schizophrenia, Their Unaffected Siblings, and Controls. *Schizophr. Bull.* **37**, 1157–1167 (2011).
8. Price, G. W. *et al.* A Multivariate Electrophysiological Endophenotype, from a Unitary Cohort, Shows Greater Research Utility than Any Single Feature in the Western Australian Family Study of Schizophrenia. *Biol. Psychiatry* **60**, 1–10 (2006).
9. Seidman, L. J. *et al.* Factor structure and heritability of endophenotypes in schizophrenia: Findings from the Consortium on the Genetics of Schizophrenia (COGS-1). *Schizophr. Res.* **163**, 73–79 (2015).
10. Toomey, R. *et al.* Association of neuropsychological vulnerability markers in relatives of schizophrenic patients. *Schizophr. Res.* **31**, 89–98 (1998).
11. Andreou, C. *et al.* Resting-state theta-band connectivity and verbal memory in schizophrenia and in the high-risk state. *Schizophr. Res.* **161**, 299–307 (2015).
12. Boutros, N. N. *et al.* The status of spectral EEG abnormality as a diagnostic test for schizophrenia. *Schizophr. Res.* **99**, 225–237 (2008).
13. da Cruz, J. R. *et al.* EEG microstates are a candidate endophenotype for schizophrenia. *Nat. Commun.* **11**, (2020).
14. Di Lorenzo, G. *et al.* Altered resting-state EEG source functional connectivity in schizophrenia: the effect of illness duration. *Front. Hum. Neurosci.* **9**, (2015).
15. Kim, D.-J. *et al.* An estimation of the first positive Lyapunov exponent of the EEG in patients with schizophrenia. *Psychiatry Res. Neuroimaging* **98**, 177–189 (2000).
16. Nikulin, V. V., Jönsson, E. G. & Brismar, T. Attenuation of long-range temporal correlations in the amplitude dynamics of alpha and beta neuronal oscillations in patients with schizophrenia. *NeuroImage* **61**, 162–169 (2012).
17. Sun, J. *et al.* Abnormal Dynamics of EEG Oscillations in Schizophrenia Patients on Multiple Time Scales. *IEEE Trans. Biomed. Eng.* **61**, 1756–1764 (2014).
18. Uhlhaas, P. J. & Singer, W. Abnormal neural oscillations and synchrony in schizophrenia. *Nat. Rev. Neurosci.* **11**, 100–113 (2010).
19. Krishnan, A., Williams, L. J., McIntosh, A. R. & Abdi, H. Partial Least Squares (PLS) methods for neuroimaging: A tutorial and review. *NeuroImage* **56**, 455–475 (2011).
20. Abdi, H. & Williams, L. J. Partial Least Squares Methods: Partial Least Squares Correlation and Partial Least Square Regression. in *Computational Toxicology* (eds. Reisfeld, B. & Mayeno, A. N.) vol. 930 549–579 (Humana Press, 2013).
21. McIntosh, A. R., Bookstein, F. L., Haxby, J. V. & Grady, C. L. Spatial Pattern Analysis of Functional Brain Images Using Partial Least Squares. *NeuroImage* **3**, 143–157 (1996).
22. Tucker, L. R. An inter-battery method of factor analysis. *Psychometrika* **23**, 111–136 (1958).
23. Howes, O. D. & Kapur, S. The Dopamine Hypothesis of Schizophrenia: Version III--The Final Common Pathway. *Schizophr. Bull.* **35**, 549–562 (2009).
24. Burns, J. The social brain hypothesis of schizophrenia. *World Psychiatry Off. J. World Psychiatr. Assoc. WPA* **5**, 77–81 (2006).
25. Hu, W., MacDonald, M. L., Elswick, D. E. & Sweet, R. A. The glutamate hypothesis of schizophrenia: evidence from human brain tissue studies. *Ann. N. Y. Acad. Sci.* **1338**, 38–57 (2015).
26. Friston, K., Brown, H. R., Siemerkus, J. & Stephan, K. E. The dysconnection hypothesis (2016). *Schizophr. Res.* **176**, 83–94 (2016).

27. Rieger, K., Diaz Hernandez, L., Baenninger, A. & Koenig, T. 15 Years of Microstate Research in Schizophrenia – Where Are We? A Meta-Analysis. *Front. Psychiatry* **7**, (2016).
28. Cole, J. H. Multimodality neuroimaging brain-age in UK biobank: relationship to biomedical, lifestyle, and cognitive factors. *Neurobiol. Aging* **92**, 34–42 (2020).
29. Stefansson, H. *et al.* Common variants conferring risk of schizophrenia. *Nature* **460**, 744–747 (2009).
30. Burmeister, M., McInnis, M. G. & Zöllner, S. Psychiatric genetics: progress amid controversy. *Nat. Rev. Genet.* **9**, 527–540 (2008).
31. Hilker, R. *et al.* Heritability of Schizophrenia and Schizophrenia Spectrum Based on the Nationwide Danish Twin Register. *Biol. Psychiatry* **83**, 492–498 (2018).
32. Kessler, R. C. *et al.* The prevalence and correlates of nonaffective psychosis in the National Comorbidity Survey Replication (NCS-R). *Biol. Psychiatry* **58**, 668–676 (2005).
33. Moreno-Küstner, B., Martín, C. & Pastor, L. Prevalence of psychotic disorders and its association with methodological issues. A systematic review and meta-analyses. *PLOS ONE* **13**, e0195687 (2018).
34. Saha, S., Chant, D., Welham, J. & McGrath, J. A systematic review of the prevalence of schizophrenia. *PLoS Med.* **2**, e141 (2005).
35. Simeone, J. C., Ward, A. J., Rotella, P., Collins, J. & Windisch, R. An evaluation of variation in published estimates of schizophrenia prevalence from 1990–2013: a systematic literature review. *BMC Psychiatry* **15**, (2015).
36. Avila, M., Thaker, G. & Adami, H. Genetic epidemiology and schizophrenia: a study of reproductive fitness. *Schizophr. Res.* **47**, 233–241 (2001).
37. Bassett, A. S., Bury, A., Hodgkinson, K. A. & Honer, W. G. Reproductive fitness in familial schizophrenia. *Schizophr. Res.* **21**, 151–160 (1996).
38. Keller, M. C. & Miller, G. Resolving the paradox of common, harmful, heritable mental disorders: Which evolutionary genetic models work best? *Behav. Brain Sci.* **29**, 385–404 (2006).
39. MacCabe, J. H., Koupil, I. & Leon, D. A. Lifetime reproductive output over two generations in patients with psychosis and their unaffected siblings: the Uppsala 1915–1929 Birth Cohort Multigenerational Study. *Psychol. Med.* **39**, 1667 (2009).
40. Liu, C., Everall, I., Pantelis, C. & Bousman, C. Interrogating the Evolutionary Paradox of Schizophrenia: A Novel Framework and Evidence Supporting Recent Negative Selection of Schizophrenia Risk Alleles. *Front. Genet.* **10**, (2019).
41. da Cruz, J. R. *et al.* Neural Compensation Mechanisms of Siblings of Schizophrenia Patients as Revealed by High-Density EEG. *Schizophr. Bull.* **46**, 1009–1018 (2020).
42. Favrod, O. *et al.* Electrophysiological correlates of visual backward masking in patients with first episode psychosis. *Psychiatry Res. Neuroimaging* **282**, 64–72 (2018).
43. Garobbio, S. A. *et al. Electrophysiological correlates of visual backward masking in patients with bipolar disorder*. http://biorxiv.org/lookup/doi/10.1101/2020.05.12.090407 (2020) doi:10.1101/2020.05.12.090407.
44. da Cruz, J. R., Chicherov, V., Herzog, M. H. & Figueiredo, P. An automatic pre-processing pipeline for EEG analysis (APP) based on robust statistics. *Clin. Neurophysiol.* **129**, 1427–1437 (2018).
45. McIntosh, A. R., Chau, W. K. & Protzner, A. B. Spatiotemporal analysis of event-related fMRI data using partial least squares. *NeuroImage* **23**, 764–775 (2004).
46. Srebro, N. & Shraibman, A. Rank, Trace-Norm and Max-Norm. in *Learning Theory* (eds. Auer, P. & Meir, R.) 545–560 (Springer, 2005). doi:10.1007/11503415_37.
47. Abdi, H. Partial least squares regression and projection on latent structure regression (PLS Regression). *WIREs Comput. Stat.* **2**, 97–106 (2010).
48. Pedregosa, F. *et al.* Scikit-learn: Machine Learning in Python. *J. Mach. Learn. Res.* (2011).
49. Lachenbruch, P. A. & Mickey, M. R. Estimation of Error Rates in Discriminant Analysis. *Technometrics* **10**, 1–11 (1968).

50. Guyon, I., Weston, J., Barnhill, S. & Vapnik, V. Gene Selection for Cancer Classification using Support Vector Machines. *Mach. Learn.* **46**, 389–422 (2002).