Predicting Individual Psychotic Experiences on a Continuum Using Machine Learning

Jeremy A Taylor[1], Kit Melissa Larsen[1,2],
Ilvana Dzafic[1,2], Marta I Garrido[1,2,3,4]

[1]Queensland Brain Institute, The University of Queensland, Australia
[2]ARC Centre for Integrative Brain Function
[3]School of Mathematics and Physics, The University of Queensland, Australia
[4]Centre for Advanced Imaging, The University of Queensland, Australia

Running Title: Predicting Psychotic Experiences on a Continuum

Correspondence:

Jeremy Taylor
Queensland Brain Institute, The University of Queensland
Building 79, Research Road, St Lucia 4072, Brisbane, Australia
jeremy.taylor@uq.edu.au

Word Counts:

Abstract = 174 words
Full Text = 3962 words

Funding: This work was supported by the Australian Research Council Centre of Excellence for Integrative Brain Function (ARC Centre Grant CE140100007), a University of Queensland Fellowship (2016000071) and a Foundation Research Excellence Award (2016001844) to MIG.

The authors declare no competing financial interests.

Abstract

Previous studies of psychosis using machine learning methods have primarily been concerned with binary classification of patients and healthy controls. The aim of this study was to use electroencephalographic (EEG) data and pattern recognition to predict individual psychotic experiences on a continuum between these two extremes in otherwise healthy people.

From responses evoked by an auditory oddball paradigm, behavioural measures, neural activity, and effective connectivity were extracted as potential contributing features. Optimal performance was achieved using spatiotemporal maps of neural activity in response to frequent sounds, with late-P50 and early-N100 time windows contributing most to higher schizotypy scores. Effective connectivity estimates, in particular top-down frontotemporal connections, were also predictive of psychotic symptoms.

As a proof-of-concept, these findings demonstrate that individual psychotic experiences in healthy people can be predicted from EEG data alone, whilst also supporting the idea of altered sensory responses and the dysconnection hypothesis in schizophrenia, as well as the notion that psychosis may exist on a continuum.

Keywords

regression, schizotypy, psychosis, mismatch negativity, effective connectivity, EEG

1. Introduction

Schizotypy, or psychotic-like experiences in the healthy population[1], is considered a potential phenotype for schizophrenia, particularly as those with schizotypal personality traits are at heightened higher risk of developing more severe schizophrenia spectrum disorders[2]. Consequently, delineating the brain mechanisms underlying the continuum of psychosis in the healthy population could potentially help in understanding the disease trajectory.

Multivariate machine learning regression[3,4,5] is a pertinent means for modelling psychosis on a continuum and making predictions at the individual level. Machine learning algorithms can automatically detect patterns in data and train computational models to make such predictions about new test data, based on prior observations from the training data[6]. In the field of psychiatry, previous machine learning studies using neuroimaging data have primarily been concerned with the binary classification of patient groups and healthy controls. Stephan et al[7] recently argued against this dichotomy and questioned the utility of models capable of reproducing discrete, predefined categories. In order to model the complete spectrum of possible disease states, it is equally important to include those with symptoms between these two extremes, such as psychotic experiences in otherwise healthy individuals. For example, Jeong et al[8] classified individuals into high (top 10% of Schizotypal Personality Questionnaire[9] scores) and low schizotypy groups (defined by the sample mean $\pm$ 0.5 *SD*) with 98% accuracy using event-related potentials from an emotional perception paradigm. However, thus far it is unclear whether the degree of psychosis can be predicted from neuroimaging data when treating the severity of schizotypal experiences as a continuous variable[4,10].

The aim of this study was to predict a range of psychotic experiences on a continuum within a healthy population using machine learning regression methods on electrical brain activity, recorded non-invasively during a sensory learning task[11]. We applied two approaches: 1) a pipeline incorporating sequential feature selection from a range of behavioural measures, event-related potential components, as well as

effective connectivity between brain regions, enabling us to improve model interpretability and better determine the importance of individual features; 2) applying a spatiotemporal approach to event-related potentials as to determine which spatial locations and time points contribute most to predictions. In so doing, we hoped to provide a proof-of-concept with potential clinical application, and evidence to support that psychosis may exist on a continuum, even in the healthy population.

2. Methods

2.1. Participants

A group of 31 functioning individuals with no prior history of psychiatric or neurological disorders (age 19 to 38 years, mean = 24.65 years, $SD = 4.85$, 14 males) participated in this study. All participants provided written informed consent and were paid for their participation. The study received approval by the University of Queensland Human Research Ethics committee.

To obtain individual schizotypy scores, all participants completed the Prodromal Questionnaire[12] (PQ), which comprises of 92 items assessing four subscales; positive, negative, disorganised and general symptoms. Responses recorded the frequency of psychotic experiences ('never', 'once or twice per month', 'once per week', 'few times per week', 'daily'), and for those which had occurred, whether the experience was distressing.

2.2. Experimental design

EEG was recorded whilst participants listened to an auditory oddball paradigm in which streams of tones were presented at 500ms intervals, with frequency of 500Hz, occasionally with differing duration (20% of trials). The duration assigned to 'standard' and 'deviant' stimuli alternated between 'short' (50ms) and 'long' (100ms) tones at varying regularities within 'stable' and 'volatile' blocks. At the end of each block, participants reported the probability of the more frequent tones heard and level of confidence in their estimate, both on a scale of 0 to 100%. For more detail regarding the task, refer to Dzafic et al[11].

2.3. Data collection and pre-processing

2.3.1. *EEG*

Continuous EEG data was recorded using a 64 channel Biosemi ActiveTwo system (Amsterdam, Netherlands) with electrode placement according to the international 10-20 standard[13] at a sampling rate of 1024Hz. Offline signal processing was performed using SPM12 (www.fil.ion.ucl.ac.uk/spm)[14]. Data were referenced to the scalp average, downsampled to 200Hz, and bandpass filtered with a 0.5Hz high-pass Butterworth filter. Eye blinks were detected using the vertical electrooculogram channel and corrected using the Berg method[15]. Events were epoched using a −100 to +400ms peri-stimulus time window. Trials were thresholded for artefacts at 100μV, robustly averaged[16], low-pass filtered at 40 Hz, and baseline corrected between −100 and 0ms.

2.3.2. *Feature selection approach*

To facilitate feature selection, a subset of 20 variables was extracted from the dataset; event-related potential (ERP) components, behavioural measures, and effective connectivity between brain regions. The definition of these features was motivated by previous literature showing group difference between people with schizophrenia and healthy controls.

A total of six ERP components for the standard, deviant and Mismatch Negativity (MMN; deviant − standard) responses under stable and volatile conditions were obtained from the frontocentral channel (Fz), averaging across the 150-250ms time window, i.e. the typical scalp location and latency associated with the prediction error responses observed in prior studies using oddball paradigms[17,18].

Effective brain connectivity associated with the effect of environment stability was estimated using Dynamic Causal Modelling (DCM) of a well-defined functional brain architecture[19, 20], incorporating a three-level hierarchical brain model which underlies the generation of prediction error in auditory oddball paradigms. This

model included: bilateral primary auditory cortices (A1; MNI coordinates: left [-42, -22, 7] and right [46, -14, 8]; the cortical input sources), bilateral superior temporal gyri (STG; left [-61, -32, 8] and right [59, -25, 8]), and bilateral inferior frontal gyri (IFG; left [-46, 20, 8] and right [46, 20, 8]). Nine model architectures with differing intrinsic, backward, and forward connections were defined. Bayesian model averaging was performed to estimate the strength of effective connectivity for each of the ten connections (referred to here as the DCM parameters) across all models (resulting from a weighted average of each connectivity parameter by the likelihood of the model under which that parameter was estimated). Note that the parameter estimation process was done separately for each participant without model selection, such that each training set remains completely independent from the testing dataset. A similar approach, referred to a generative embedding has been used before[21], although in that paper, model optimization was done using the pooled data (both training and testing). For more detail regarding DCM fitting and parameter estimation, refer to Dzafic et al[11].

2.3.3. *Spatiotemporal approach*

To facilitate spatiotemporal analyses, the ERPs for each participant, response and condition were converted to NIfTI images. At each time point, data were spatially interpolated between channels to form a two-dimensional scalp map, represented by a 32 × 32 matrix. These spatial images were then concatenated in temporal order, resulting in a three-dimensional spatiotemporal volume. Images were smoothed using a Gaussian filter with full width at half maximum (FWHM) of 12mm × 12mm × 20ms. In contrast with the feature selection approach, we did not make any further assumptions about channel locations or time windows.

2.4. Modelling

In all approaches described below, the target variables were the individual PQ scores, weighted by the frequency of occurrence and log transformed to fulfil normality assumptions.

2.4.1. *Feature selection approach*

Regression modelling was performed on three sets of features extracted from behavioural, ERP and DCM data using the Huber regressor[22,23] with an additional sequential feature selection process[24], reducing the dimensionality of the data from the complete set of M features, X_M, to a subset of m features, where $m \leq M$. Our pipeline, as illustrated in Figure 1, was implemented in the Scikit-learn[25] library (version 0.19.1) for Python (version 3.6.1; www.python.org) with sequential feature selection using the mlxtend library (version 0.10.0; rasbt.github.io/mlxtend).

All features were rescaled from their respective maximum and minimum values to range from zero to one. Stratified $K = 5$-fold cross-validation was performed on the whole data set (Figure 1, shown in red) with $R = 10$ repetitions, resampling the data whilst ensuring that the distribution of scores was sampled uniformly between folds. We therefore obtain 50 sub-models (K folds × R repetitions, highlighted in green), one for each subset, all of which contribute to a collective ensemble model[26].

Sequential forward selection was performed with an inner $J = 5$-fold cross-validation loop[27] (orange) on the training set from the main cross-validation loop (blue) to obtain the features which minimised mean-square error. We represent the order in which features are selected for a given fold, k, and repetition, r, as a binary $M \times M$ matrix, denoted here as $\mathbf{\Omega}_{k,r}$, with individual features as rows, the feature count or algorithm iterations as columns, selected features as ones and unselected features as zeros. We enforce the criterion that all sub-models must share the same number of features and level of complexity, however the differences in training data between folds and repetitions will result in different features being selected for each

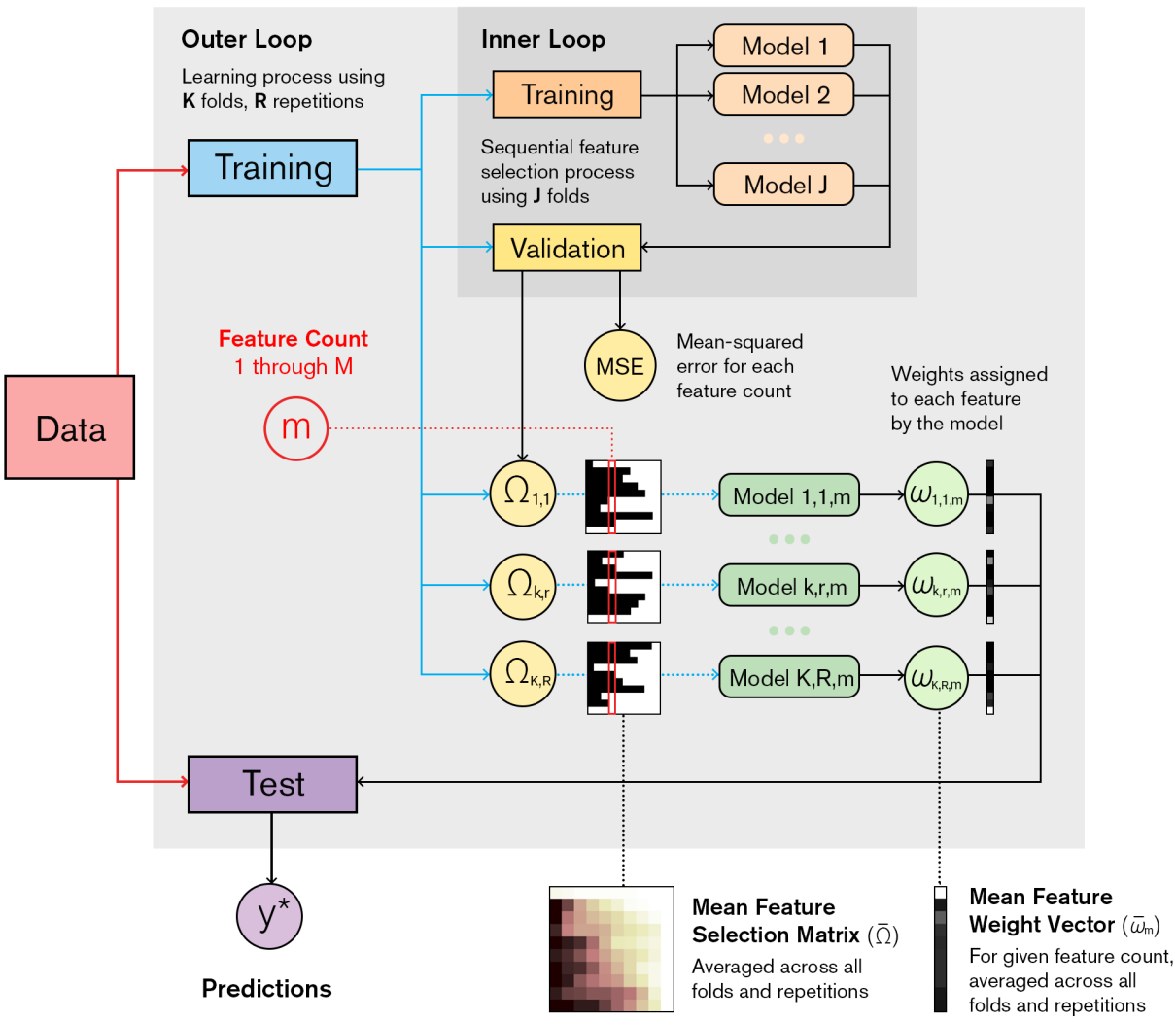

Figure 1. Schematic of machine learning pipeline. In the outer cross-validation loop, the collective dataset is partitioned into K folds, of which one is used for testing (purple) and the remaining set of K–1 folds for training (blue) a sub-model. This partitioning is repeated R times with differences in sampling. Feature selection is performed with an inner cross-validation loop, which partitions the training set from the outer loop into J folds, which form secondary training (orange) and validation sets (yellow). This process returns a binary matrix, $\Omega_{k,r}$, indicating the order in which features are selected. The optimal feature count, m, is that which achieves minimal mean-squared error across all folds and repetitions. There are a total $K \times R$ models (green), each of which are based on the mth column of $\Omega_{k,r}$. The training process assigns weights, $\omega_{k,r,m}$, to each feature and are used to compute predictions, y^*, of independent samples in the test set. The mean feature selection matrix, $\overline{\mathbf{\Omega}}$, and mean feature weight vector, $\overline{\boldsymbol{\omega}}_m$, i.e. the average across all K folds and R repetitions, are used to intuit the feature importances.

sub-model. The optimal number of features, denoted m, is that which achieved minimal mean-squared error on the validation set (yellow). From the $\mathbf{\Omega}_{k,r}$ matrices, we extract a feature selection mask from the mth column, determining which features form the basis for each sub-model. Through the training process, each of these features are assigned weights, denoted $\omega_{k,r,m}$. Each model then uses these weightings to make predictions, denoted y^*, of the test data (purple), which is strictly independent of both the optimisation and training processes.

Using this pipeline, we specified four models; one based on each subset of the data, categorised as behavioural, ERP and DCM measures, as well as a combined model using all three measures. The p-values for each model was computed through 1000 random permutations of the target PQ scores, with statistically significant results ($p \leq 0.05$) suggesting that the algorithm had truly learnt some pattern within the data[28] above chance, subject to Šidák correction[29] for multiple comparisons.

To summarise the frequency at which a given feature is selected, we average across all feature selection matrices ($\overline{\mathbf{\Omega}}$), sorting features according to mean selection rate (i.e. the average across all columns) and in particular, inspect the optima from column m. To assess the weightings assigned to each feature upon selection, we also average across all weight vectors for the feature count m ($\overline{\omega}_m$).

2.4.2. *Spatiotemporal feature sets*

Regression modelling was performed using the kernel ridge regression algorithm[30], as implemented within the Pattern Recognition for Neuroimaging Toolbox (PRoNTo, version 2.0; www.mlnl.cs.ucl.ac.uk/pronto)[31]. Ridge regression minimises the sum of squared residual errors whilst also penalising the size of coefficients through regularisation.

A total of four models were specified; one for each response (i.e. the standard, deviant and MMN difference waves), and a combined model concatenating all three images. For each model, we obtained mean feature weight maps, again averaging

across all folds and repetitions. We compute an element-wise multiplication between the mean weight map and the grand mean activity associated with the model of interest, referred to henceforth as the 'feature importance' map (for illustration, refer to Supplementary Material A).

2.4.3. *Evaluation of performance*

Mean square error (MSE), as well as the linear dependence between predictions and target variables, as measured using Pearson's correlation coefficient (R) were used to evaluate model performance. Both measures were used to examine model performance and identify any potential biases.

3. Results

The overall performance of all models is summarised in Table 1. Below, we describe the findings from the two approaches used.

Table 1. Summary of model performance.

Method	Features	Correlation	p-value uncorrected	p-value Sidak corrected	MSE	p-value uncorrected	p-value Sidak corrected
Feature Extraction and Selection	DCM	0.43	0.012	0.047	1.267	0.012	0.047
	ERP	0.15	0.210	0.610	1.488	0.210	0.610
	Behavioural	-0.02	0.528	0.950	1.514	0.528	0.950
	Combined	0.10	0.330	0.798	1.806	0.322	0.789
Spatiotemporal Images	Stable Standard	0.58	0.001	0.008	0.963	0.001	0.008
	Stable Deviant	0.33	0.002	0.016	1.301	0.005	0.039
	Stable MMN	0.49	0.001	0.008	1.095	0.001	0.008
	Stable Combined	0.46	0.001	0.008	1.128	0.002	0.016
	Volatile Standard	0.33	0.007	0.055	1.298	0.004	0.032
	Volatile Deviant	0.22	0.017	0.128	1.410	0.104	0.585
	Volatile MMN	0.01	0.259	0.909	1.622	0.987	1.000
	Volatile Combined	0.27	0.010	0.077	1.351	0.013	0.099

Of all models using feature extraction and selection procedures, the model based on DCM connectivity estimates yielded best performance. All models based on spatiotemporal ERPs were statistically significant, with the standard response proving more accurate than the deviant and MMN. p-values were computed via 1000 permutations with Šidák correction for multiple comparisons.

3.3.1. Feature selection approach

The model containing only connectivity parameters outperformed both ERP components and behavioural metrics, with higher correlation ($R = 0.43$) and lower mean-squared error (1.267, $p = 0.047$).

To assess which features are driving model performance, the mean feature selection matrices and weight vectors are shown below in Figure 2. In the connectivity model, the right IFG → right STG parameter was chosen earliest in the sequence, with greatest frequency (0.98) and was also assigned higher weighting (5.10) in comparison with all other parameters. Although right A1 → right A1 (0.64),

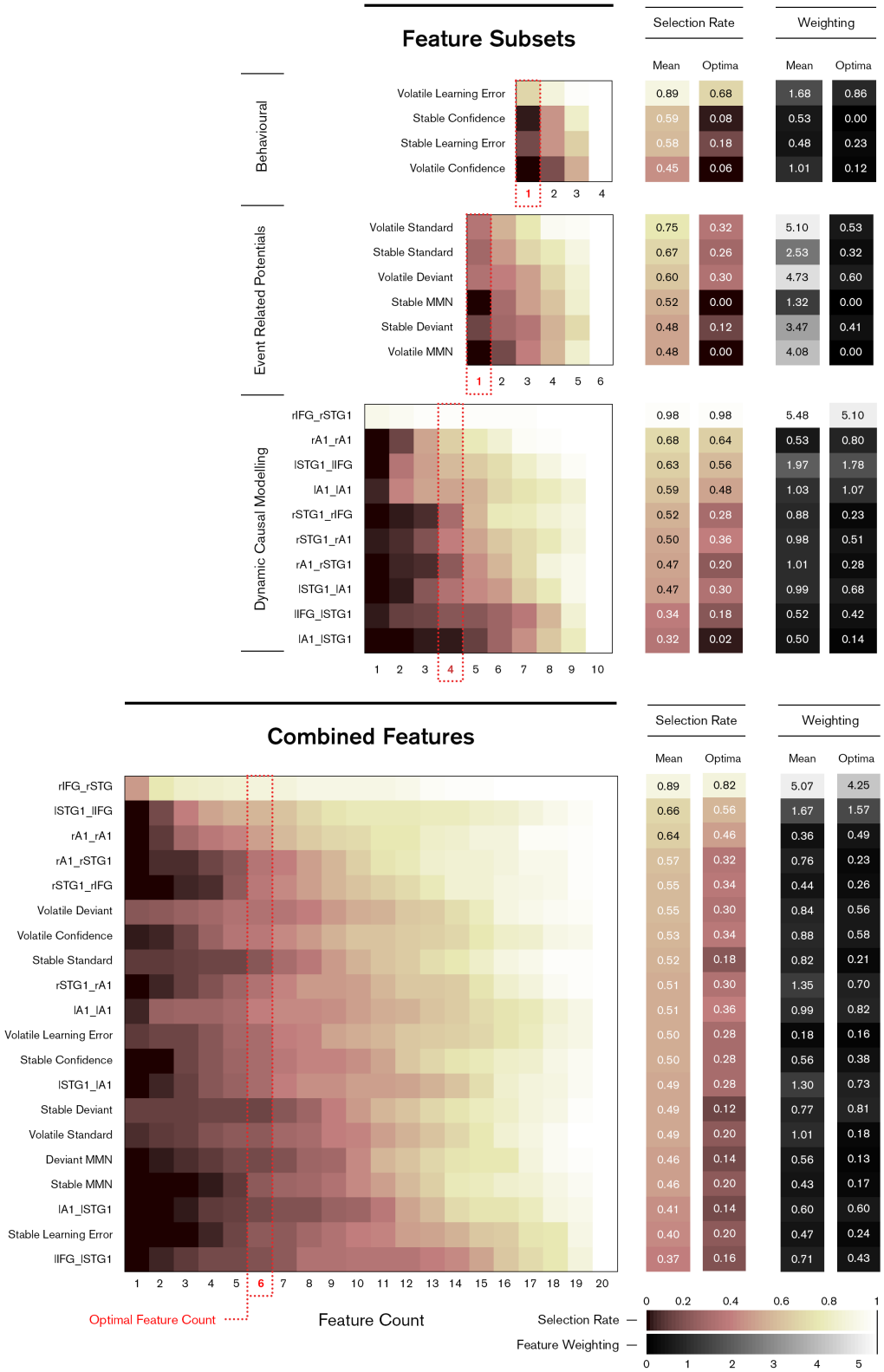

Figure 2. Feature selection matrices and weight vectors for each feature set, indicating the frequency at which each feature was selected and the contributions to model predictions upon selection. Matrices are computed as the average across all K folds and R repetitions. The mean selection rate and weightings are the average across all feature counts, i.e. 1 through M. The optimal feature count, m, for each model is highlighted with red dotted line. Feature selection rates are indicated using warm colormap and feature weightings are shown in greyscale.

frequency for the remaining features, the weights were markedly lower by comparison (0.80-1.78) and the level of variability amongst folds and repetitions remains inconclusive. These four features were also selected with highest frequency in the model using a combined feature set.

3.3.2. Spatiotemporal approach

Overall, the spatiotemporal approach yielded greater performance in comparison with the feature selection procedure (Table 1), of which the stable standard response demonstrated superior predictive performance ($R = 0.58$, MSE = 0.963, $p = 0.008$) in comparison with the deviant, MMN and combined feature sets ($R = 0.33$-0.49, MSE = 1.095-1.301, $p \leq 0.016$). Equivalent models based on the volatile condition demonstrated markedly lower performance ($R = 0.01$-0.33, MSE = 1.298-1.622) and were non-significant.

Observed and predicted PQ scores are shown in Figure 3 for the best performing models in the two approaches. Although the stable standard spatiotemporal model (Figure 3B) has demonstrably lower error on the mid to high-range band of the spectrum, we can observe that despite fewer examples at the lower end of the spectrum, the DCM feature selection model (Figure 3A) is able to capture more of this information.

The distribution of all signed feature importances within the volume is shown in Figure 4C, with positive importances contributing to a higher PQ and negative importance contributing to lower PQ. To intuit which features are driving the model predictions, we threshold the top 5% of both positive and negative importances for the purposes of visualisation in Figures 4A and 4B. Note however, that although the majority of features have zero or low importance, all voxels in the image contribute towards the model predictions. The voxels contributing most to lower scores (Figure 4A) peak at fronto-central channels at approximately 50ms and between 150 and 275ms, whereas those contributing most to high scores (Figure 4B) peak between approximately 35 and 125ms over the frontal and left posterior channels.

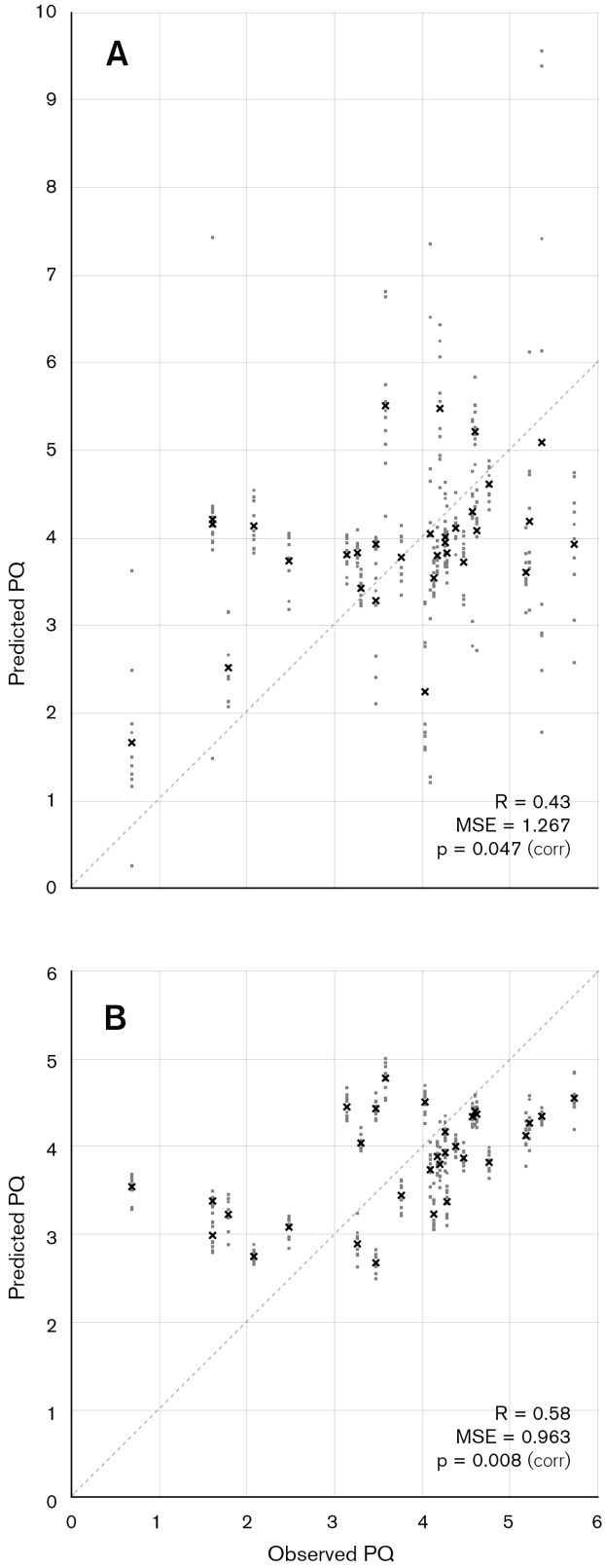

Figure 3. Plot of model predictions and true observed PQ scores. (A) DCM connectivity feature set, using Huber regression and forward selection, optimised to four features; (B) Spatiotemporal stable standard ERP feature set, using kernel ridge regression and regularisation. Predictions from each repetition are shown as grey dots and overall predictions as black crosses. All scores are normalised via log transform.

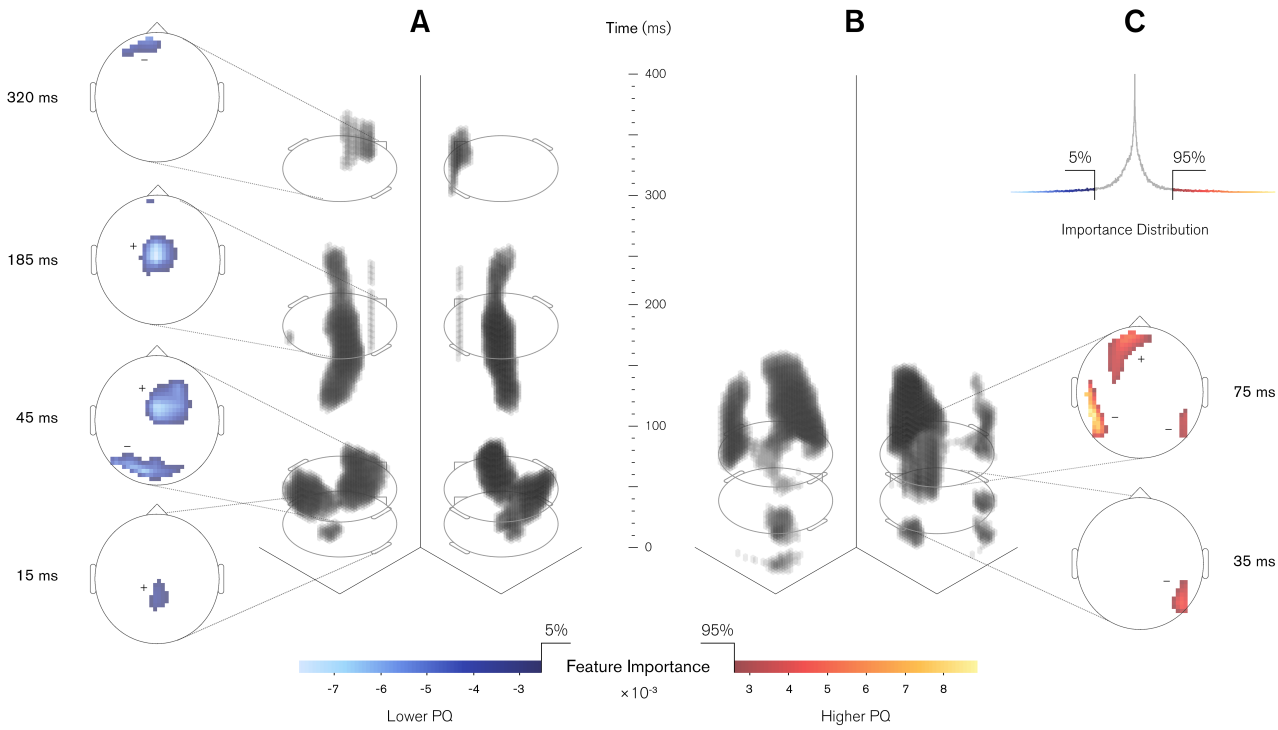

Figure 4. Spatiotemporal features of greatest importance to schizotypy predictions in stable standard response model. (A-B) 3D volumes represent clusters of greatest feature importance for low (cool) and high (warm) predictions of PQ, thresholded at top 5% of negative and positive importances respectively, with spatial dimensions on x-y plane and time domain along vertical z axis, displayed from dual angles. 2D scalp maps refer to peaks within main clusters, with $\pm$ labels indicating polarity of the original ERP signal; (C) Distribution of all feature importances in the feature set.

4. Discussion

In this study, we demonstrate that psychotic experiences in the healthy population can be predicted from EEG data using machine learning. Specifically, we show that the dynamics of brain activity captured by EEG and its underlying effective connectivity, are predictive of Prodromal Questionnaire scores at the individual level, with top-down frontotemporal connectivity and early ERP components having greatest contribution.

Using a sequential feature selection pipeline, we observed that of all models based on behavioural measures, ERP components, effective connectivity (DCM), and combined feature sets, the connectivity parameters alone achieved the greatest performance. Within both the DCM and combined models, the right IFG to right STG connection was the main feature driving the prediction of PQ. Note that the optimal feature count for our DCM-based model was four (including primarily the left STG → left IFG, as well as the right A1 and left A1 intrinsic connections), which suggests that this right top-down frontotemporal connection holds greater predictive value when combined with other features. In our spatiotemporal models, we found that the brain responses to standard, predictable sounds, were more predictive of PQ scores than the deviant and differential MMN responses (standard ERP components were also selected before deviants and MMN in our feature selection routine). Within the corresponding feature importance maps, voxels contributing most to higher predicted scores lie in the late-P50 and early-N100 time windows, in keeping with known alterations in schizophrenia[32] and genetically high risk groups[33].

Top-down frontotemporal connections and intrinsic right A1 connections have been previously identified by Dima et al[34] as disrupted in patients with schizophrenia. In addition, Larsen et al[33] also observed that the top-down frontotemporal was weaker in 22q11.2 deletion carriers, one of the largest known genetic risk factors for schizophrenia. Finally, this finding is aligned with the dysconnectivity hypothesis for schizophrenia[35], which postulates that aberrant functional connectivity is particularly found between the frontal and temporal regions.

Failure to suppress the P50 and N100 have consistently been reported in the schizophrenia literature, not only in chronic schizophrenia, but also in at-risk groups[36] and relatives of patients[37], linking these deficits as potential endophenotypes for schizophrenia[38]. Parallels can also be drawn with alterations in sensory gating, pre-pulse inhibition[39], and the failure to habituate in response to stimulus repetition in schizophrenia. Such findings are in keeping with the predictive coding theory which considers these early effects as a failure to encode the patterns associated with the standard sound[40]. As a result, all stimuli, both standard and deviant sounds, are rendered surprising. Consequently, aberrant predictive models may be formed and what should be predictable sensory stimuli may appear unusually salient, leading to the formation of unusual percepts and beliefs.

The machine learning methods employed here go a step beyond correlation analysis, in that they can predict schizotypal personality traits of an individual based solely on their EEG profile. This is closely aligned with the concept of precision medicine, which focuses on subtypes, shared pathological mechanisms and tailoring personalised interventions for the individual[41,42], as opposed to a reduction into heterogenous diagnostic categories. We envisage that these neuroimaging approaches, which simultaneously provide accurate predictions and highlight underlying biological mechanisms, can contribute to precision psychiatry[43] if integrated with other instruments as part of a wholistic strategy[44], augmenting the role of the clinical practitioner.

Technical challenges associated with this type of modelling stem from the fact that in typical neuroimaging applications, the number of available features in high-dimensional data informing model predictions vastly outweighs the number of samples[45]. As such, fitting a continuous function to a high-dimensional, yet sparsely populated space which also generalises to new samples becomes difficult. In order to control for any feature selection bias and overfitting model parameters[46], we applied two different methodologies. Firstly, we extracted sets of features from multiple domains *a priori*, then using sequential feature selection, we were able to vastly reduce the number of features and pinpoint those which contributed most to the model. We then utilised an existing and widely used machine learning method in

neuroimaging known as kernel learning[47,31], which instead reduces the risk of overfitting by computing pairwise similarities between whole brain images and further constrains the size of feature weights through regularisation. With this approach, we obtained a spatial representation of the predictive function, indicating which brain regions were assigned the largest weights in predicting the outcome variable[48,49], however direct interpretation of these weight maps is not as straightforward[50,51]. Although well suited for spatially distributed features, this does not incorporate features from other domains, such as behavioural and connectivity measures.

Although our findings are based on a relatively small sample, this is in keeping with recent studies using similar methods to predict continuous measures of personality traits[49,52], whilst also yielding comparable performance to other studies with larger cohorts[48,53,54].

Most cross-validation schema commonly used to assess model generalisability have an inherent bias and variance in their error estimation. Throughout our analysis, we have been conscious of limiting potential biases in our sampling and thus estimated model performance using a stratified 5-fold cross-validation scheme with 10 repetitions. Varoquaux recently demonstrated that commonly used schema in neuroimaging studies may report an inflated level of performance[55,56], namely leave-one-out cross-validation[57]. Although it is commonplace in neuroimaging to compute a single cross-validation split, the broader field of machine learning typically perform repeated cross-validation schema with varied sampling, averaging the results across replicates[58], which is demonstrably more robust to random variations in the training and test partitioning[59]. By employing these methods here, our reported performance metrics are thus conservative estimates of model generalisability.

In summary, we have demonstrated the validity of predicting schizotypal personality traits via machine learning regression techniques on an individual, rather than group basis. Moreover, we provide a proof-of-concept with potential applications in precision psychiatry, aligned with the dysconnection hypothesis of schizophrenia and the notion that psychosis exists on a continuum, even in the healthy general population.

Supplementary Material

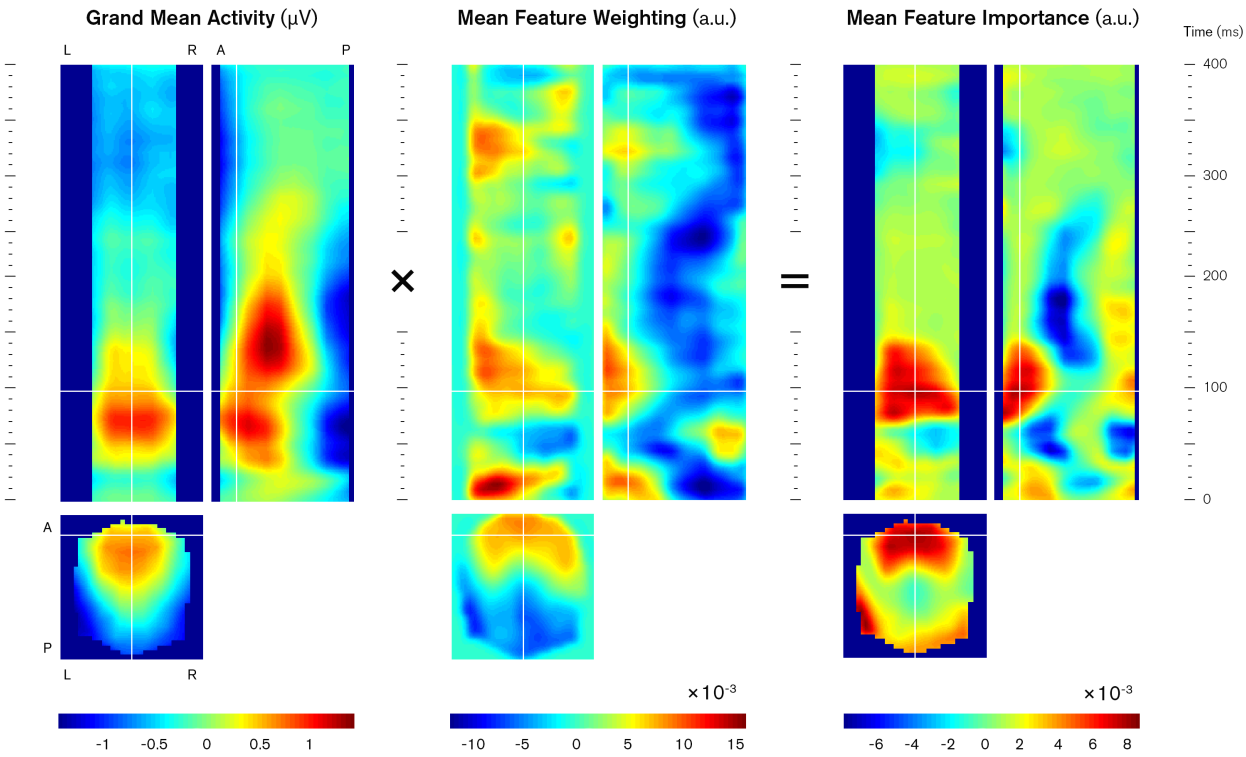

Figure A. Exemplar orthogonal views of spatiotemporal feature importance calculation (right) as the element-wise multiplication of the grand mean activity (left) and feature weights (centre). White lines indicate the intersection of spatiotemporal planes through a point of high importance in predicting high PQ scores (red), occurring frontally at approximately 100ms post-stimulus. Scalp maps orientations are annotated left (L), right (R), anterior (A) and posterior (P).

References

1. Verdoux H, van Os J. Psychotic symptoms in non-clinical populations and the continuum of psychosis. *Schizophrenia Research* 2002;54(1):59-65.

2. Wang Y, Shi H-s, Liu W-h, et al. Trajectories of schizotypy and their emotional and social functioning: An 18-month follow-up study. *Schizophrenia Research* 2018;193:384-390.

3. Cohen JR, Asarnow RF, Sabb FW, Bilder RM, Bookheimer SY, Knowlton BJ, Poldrack RA. Decoding Continuous Variables from Neuroimaging Data: Basic and Clinical Applications. *Frontiers in Neuroscience* 2011;5:75.

4. Madsen KH, Krohne LG, Cai X-l, Wang Y, Chan RCK. Perspectives on Machine Learning for Classification of Schizotypy Using fMRI Data. *Schizophrenia Bulletin* 2018:sby026.

5. Janssen RJ, Mourão-Miranda J, Schnack HG. Making Individual Prognoses in Psychiatry Using Neuroimaging and Machine Learning. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging* 2018.

6. Alpaydin E. *Introduction to Machine Learning.* 3rd ed. Cambridge, Massechusetts: The MIT Press; 2014.

7. Stephan KE, Schlagenhauf F, Huys QJM, et al. Computational neuroimaging strategies for single patient predictions. *NeuroImage* 2017;145:180-199.

8. Jeong JW, Wendimagegn TW, Chang E, Chun Y, Park JH, Kim HJ, Kim HT. Classifying Schizotypy Using an Audiovisual Emotion Perception Test and Scalp Electroencephalography. *Frontiers in Human Neuroscience* 2017;11:450.

9. Raine A. The SPQ: A Scale for the Assessment of Schizotypal Personality Based on DSM-III-R Criteria. *Schizophrenia Bulletin* 1991;17(4):555-564.

10. Taylor JA, Matthews N, Michie PT, Rosa MJ, Garrido MI. Auditory prediction errors as individual biomarkers of schizophrenia. *NeuroImage: Clinical* 2017;15:264-273.

11. Dzafic I, Randeniya R, Garrido MI. Reduced top-down connectivity as an underlying mechanism for psychotic experiences in healthy people. *bioRxiv* 2018.

12. Loewy RL, Bearden CE, Johnson JK, Raine A, Cannon TD. The prodromal questionnaire (PQ): Preliminary validation of a self-report screening measure for prodromal and psychotic syndromes. *Schizophrenia Research* 2005;79(1):117-125.

13. Oostenveld R, Praamstra P. The five percent electrode system for high-resolution EEG and ERP measurements. *Clinical Neurophysiology* 2001;112(4): 713-719.

14. Litvak V, Mattout J, Kiebel S, et al. EEG and MEG Data Analysis in SPM8. *Computational Intelligence and Neuroscience* 2011;2011:852961.

15. Berg P, Scherg M. A multiple source approach to the correction of eye artifacts. *Electroencephalography and Clinical Neurophysiology* 1994;90(3):229-241.

16. Wager TD, Keller MC, Lacey SC, Jonides J. Increased sensitivity in neuroimaging analyses using robust regression. *NeuroImage* 2005;26(1):99-113.

17. Näätänen R. The role of attention in auditory information processing as revealed by event-related potentials and other brain measures of cognitive function. *Behavioral and Brain Sciences* 1990;13(2):201-233.

18. Näätänen R, Pakarinen S, Rinne T, Takegata R. The mismatch negativity (MMN): towards the optimal paradigm. *Clinical Neurophysiology* 2004;115(1): 140-144.

19. Opitz B, Rinne T, Mecklinger A, von Cramon DY, Schroger E. Differential contribution of frontal and temporal cortices to auditory change detection: fMRI and ERP results. *Neuroimage* Jan 2002;15(1):167-174.

20. Garrido MI, Kilner JM, Kiebel SJ, Stephan KE, Friston KJ. Dynamic causal modelling of evoked potentials: a reproducibility study. *Neuroimage* Jul 01 2007;36(3):571-580.

21. Brodersen KH, Schofield TM, Leff AP, Ong CS, Lomakina EI, Buhmann JM, Stephan KE. Generative Embedding for Model-Based Classification of fMRI Data. *PLOS Computational Biology* 2011;7(6):e1002079.

22. Huber P. *Robust Statistics*. New York: Wiley; 1981.

23. Holland PW, Welsch RE. Robust regression using iteratively reweighted least-squares. *Communications in Statistics - Theory and Methods* 1977;6(9):813-827.

24. Guyon I, Elisseeff A. An Introduction to Variable and Feature Selection. *Journal of Machine Learning Research* 2003;3:1157-1182.

25. Pedregosa F, Varoquaux G, Gramfort A, et al. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research* 2011;12:2825-2830.

26. Hastie T, Friedman J, Tibshirani R. The Elements of Statistical Learning: Data Mining, Inference and Prediction. 1st ed. New York, NY: Springer; 2001.

27. Cawley GC, Talbot NLC. On Over-fitting in Model Selection and Subsequent Selection Bias in Performance Evaluation. *Journal of Machine Learning Research* 2010;11:2079-2107.

28. Golland P, Fischl B. Permutation Tests for Classification: Towards Statistical Significance in Image-Based Studies. In: Taylor C, Noble JA, eds. *Information Processing in Medical Imaging: 18th International Conference, IPMI 2003, Ambleside, UK, July 20-25, 2003. Proceedings*. Berlin, Heidelberg: Springer Berlin Heidelberg; 2003:330-341.

29. Šidák Z. Rectangular Confidence Regions for the Means of Multivariate Normal Distributions. *Journal of the American Statistical Association* 1967;62(318): 626-633.

30. Hoerl AE, Kennard RW. Ridge Regression: Biased Estimation of Nonorthogonal Problems. *Technometrics* 1970;12(1):55-67.

31. Schrouff J, Rosa MJ, Rondina JM, et al. PRoNTo: Pattern Recognition for Neuroimaging Toolbox. *Neuroinformatics* 2013;11(3):319-337.

32. Rosburg T, Boutros NN, Ford JM. Reduced auditory evoked potential component N100 in schizophrenia — A critical review. *Psychiatry Research* 2008;161(3):259-274.

33. Larsen KM, Mørup M, Birknow MR, et al. Altered auditory processing and effective connectivity in 22q11.2 deletion syndrome. *Schizophrenia Research*.

34. Dima D, Frangou S, Burge L, Braeutigam S, James AC. Abnormal intrinsic and extrinsic connectivity within the magnetic mismatch negativity brain network in schizophrenia: A preliminary study. *Schizophrenia Research*;135(1):23-27.

35. Stephan KE, Friston KJ, Frith CD. Dysconnection in schizophrenia: from abnormal synaptic plasticity to failures of self-monitoring. *Schizophr Bull* May 2009;35(3):509-527.

36. Brockhaus-Dumke A, Schultze-Lutter F, Mueller R, Tendolkar I, Bechdolf A, Pukrop R, Klosterkoetter J, Ruhrmann S. Sensory Gating in Schizophrenia: P50

and N100 Gating in Antipsychotic-Free Subjects at Risk, First-Episode, and Chronic Patients. *Biological Psychiatry* 2008;64(5):376-384.

37. Turetsky BI, Greenwood TA, Olincy A, et al. Abnormal Auditory N100 Amplitude: A Heritable Endophenotype in First-Degree Relatives of Schizophrenia Probands. *Biological Psychiatry* 2008;64(12):1051-1059.

38. Earls HA, Curran T, Mittal V. A Meta-analytic Review of Auditory Event-Related Potential Components as Endophenotypes for Schizophrenia: Perspectives From First-Degree Relatives. *Schizophrenia Bulletin* 2016;42(6): 1504-1516.

39. Javitt DC, Freedman R. Sensory Processing Dysfunction in the Personal Experience and Neuronal Machinery of Schizophrenia. *The American journal of psychiatry* 2015;172(1):17-31.

40. Adams R, Stephan K, Brown H, Frith C, Friston K. The Computational Anatomy of Psychosis. *Frontiers in Psychiatry* 2013;4(47).

41. Cholerton B, Larson EB, Quinn JF, et al. Precision Medicine: Clarity for the Complexity of Dementia. *The American Journal of Pathology* 2016;186(3): 500-506.

42. Espay AJ, Brundin P, Lang AE. Precision medicine for disease modification in Parkinson disease. *Nature Reviews Neurology* 2017;13:119.

43. Friston KJ, Redish AD, Gordon JA. Computational Nosology and Precision Psychiatry. *Computational Psychiatry* 2017;1:2-23.

44. McMillan CT. MRI biomarkers — a precision medicine tool in neurology? *Nature Reviews Neurology* 2016;12:323.

45. Bellman RE. Adaptive Control Processes: A Guided Tour: Princeton University Press; 1961.

46. Arbabshirani MR, Plis S, Sui J, Calhoun VD. Single subject prediction of brain disorders in neuroimaging: Promises and pitfalls. *NeuroImage* 2017;145, Part B: 137-165.

47. Schölkopf B, Smola AJ. *Learning with Kernels.* Cambridge, Massachusetts: The MIT Press; 2000.

48. Portugal LCL, Rosa MJ, Rao A, et al. Can Emotional and Behavioral Dysregulation in Youth Be Decoded from Functional Neuroimaging? *PLOS ONE* 2016;11(1):e0117603.

49. Fernandes Jr O, Portugal LCL, Alves RdCS, et al. Decoding negative affect personality trait from patterns of brain activation to threat stimuli. *Neuroimage* 2017;145(Pt B):337-345.

50. Haufe S, Meinecke F, Görgen K, Dähne S, Haynes J-D, Blankertz B, Bießmann F. On the interpretation of weight vectors of linear models in multivariate neuroimaging. *NeuroImage* 2014;87:96-110.

51. Schrouff J, Monteiro JM, Portugal L, Rosa MJ, Phillips C, Mourão-Miranda J. Embedding Anatomical or Functional Knowledge in Whole-Brain Multiple Kernel Learning Models. *Neuroinformatics* 2018;16(1):117-143.

52. Poole VN, Robinson ME, Singleton O, DeGutis J, Milberg WP, McGlinchey RE, Salat DH, Esterman M. Intrinsic functional connectivity predicts individual differences in distractibility. *Neuropsychologia* 2016;86:176-182.

53. Moradi E, Hallikainen I, Hänninen T, Tohka J, Alzheimer's Disease Neuroimaging I. Rey's Auditory Verbal Learning Test scores can be predicted from whole brain MRI in Alzheimer's disease. *NeuroImage: Clinical* 2017;13:415-427.

54. Wang Y, Iyengar V, Hu J, Kho D, Falconer E, Docherty JP, Yuen GY. Predicting Future High-Cost Schizophrenia Patients Using High-Dimensional Administrative Data. *Frontiers in Psychiatry* 2017;8(114).

55. Varoquaux G, Raamana PR, Engemann DA, Hoyos-Idrobo A, Schwartz Y, Thirion B. Assessing and tuning brain decoders: Cross-validation, caveats, and guidelines. *NeuroImage* 2017;145:166-179.

56. Little MA, Varoquaux G, Saeb S, Lonini L, Jayaraman A, Mohr DC, Kording KP. Using and understanding cross-validation strategies. Perspectives on Saeb et al. *GigaScience* 2017;6(5):1-6.

57. Veronese E, Castellani U, Peruzzo D, Bellani M, Brambilla P. Machine Learning Approaches: From Theory to Application in Schizophrenia. *Computational and Mathematical Methods in Medicine* 2013;2013:867924.

58. Witten IH, Frank E. Credibility: Evaluating What's Been Learned. In: Gray J, ed. *Data Mining: Practical Machine Learning Tools and Techniques*. 2nd ed. San Francisco, CA: Morgan Kaufmann Publishers; 2005:143-185.

59. Rodriguez JD, Perez A, Lozano JA. Sensitivity Analysis of k-Fold Cross Validation in Prediction Error Estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 2010;32(3):569-575.