



RESEARCH ARTICLE

Predicting individual improvement in schizophrenia symptom severity at 1-year follow-up: Comparison of connectomic, structural, and clinical predictors

Akhil Kottaram^{1,2}  | Leigh A. Johnston^{1,3} | Ye Tian^{2,4} | Eleni P. Ganella^{2,4,5} | Liliana Laskaris^{2,4,6} | Luca Cocchi⁷  | Patrick McGorry^{8,9} | Christos Pantelis^{2,4,5,6,10,11} | Ramamohanarao Kotagiri¹² | Vanessa Cropley^{2,4,13} | Andrew Zalesky^{1,2}

¹Department of Biomedical Engineering, The University of Melbourne, Melbourne, Victoria, Australia

²Melbourne Neuropsychiatry Centre, The University of Melbourne, Melbourne, Victoria, Australia

³Melbourne Brain Centre Imaging Unit, The University of Melbourne, Melbourne, Victoria, Australia

⁴Department of Psychiatry, The University of Melbourne, Melbourne, Victoria, Australia

⁵Cooperative Research Centre for Mental Health, Carlton, Victoria, Australia

⁶Centre for Neural Engineering, Department of Electrical and Electronic Engineering, The University of Melbourne, Melbourne, Victoria, Australia

⁷Clinical Brain Networks Group, QIMR Berghofer Medical Research Institute, Brisbane, Queensland, Australia

⁸Orygen, Parkville, Victoria, Australia

⁹Centre for Youth Mental Health, The University of Melbourne, Parkville, Victoria, Australia

¹⁰North Western Mental Health, Melbourne Health, Parkville, Victoria, Australia

¹¹Florey Institute for Neurosciences and Mental Health, Parkville, Victoria, Australia

¹²Department of Computing and Information Systems, The University of Melbourne, Melbourne, Victoria, Australia

¹³Centre for Mental Health, Faculty of Health, Arts and Design, School of Health Sciences, Swinburne University, Hawthorn, Victoria, Australia

Correspondence

Akhil Kottaram, Department of Biomedical Engineering, The University of Melbourne, Level 1, Building 261, 203 Bouverie Street, Melbourne, VIC 3010, Australia
Email: akottaram@unimelb.edu.au

Funding information

NHMRC, Grant/Award Numbers: APP1138711, APP1099082, APP1136649, 1105825, 628386, 628880; University of Melbourne Early Career Researcher, Grant/Award Number: 601253; Australian National Health and Medical Research Council (NHMRC), Grant/Award Number: 1065742

Abstract

In a machine learning setting, this study aims to compare the prognostic utility of connectomic, brain structural, and clinical/demographic predictors of individual change in symptom severity in individuals with schizophrenia. Symptom severity at baseline and 1-year follow-up was assessed in 30 individuals with a schizophrenia-spectrum disorder using the Brief Psychiatric Rating Scale. Structural and functional neuroimaging was acquired in all individuals at baseline. Machine learning classifiers were trained to predict whether individuals improved or worsened with respect to positive, negative, and overall symptom severity. Classifiers were trained using various combinations of predictors, including regional cortical thickness and gray matter volume, static and dynamic resting-state connectivity, and/or baseline clinical and demographic variables. Relative change in overall symptom severity between baseline and 1-year follow-up varied markedly among individuals (interquartile range: 55%). Dynamic resting-state connectivity measured within the default-mode network was

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2020 The Authors. *Human Brain Mapping* published by Wiley Periodicals, Inc.

the most accurate single predictor of change in positive (accuracy: 87%), negative (83%), and overall symptom severity (77%) at follow-up. Incorporating predictors based on regional cortical thickness, gray matter volume, and baseline clinical variables did not markedly improve prediction accuracy and the prognostic utility of these predictors in isolation was moderate (<70%). Worsening negative symptoms at 1-year follow-up were predicted by hyper-connectivity and hypo-dynamism within the default-mode network at baseline assessment, while hypo-connectivity and hyper-dynamism predicted worsening positive symptoms. Given the modest sample size investigated, we recommend giving precedence to the relative ranking of the predictors investigated in this study, rather than the prediction accuracy estimates.

KEYWORDS

default mode network, dynamic functional connectivity, outcome prediction, schizophrenia, symptoms

1 | INTRODUCTION

While many individuals with schizophrenia can achieve varying degrees of remission, the risk of relapse, recurrence and worsening of symptoms remains significant, particularly after treatment discontinuation (Emsley, Chiliza, Asmal, & Harvey, 2013). Relapse may indicate disease progression (Lieberman et al., 1993), and in many cases, is accompanied by brain changes (Croypley & Pantelis, 2014) or serious consequences, including extreme despair and increased risk of suicide (Pompili et al., 2007). Predicting relapse events and illness course is therefore critical to enable early and targeted initiation of appropriate treatments to individuals with a high risk of relapse. Moreover, individuals predicted to follow a favorable illness course may be spared unnecessary treatment and associated side-effects (Arana, 2000; Bruijnzeel, Suryadevara, & Tandon, 2014; Cha & McIntyre, 2012; Fusar-Poli et al., 2013; Vita, De Peri, Deste, & Sacchetti, 2012).

Predicting individual illness course, particularly the likelihood of specific relapse events, is very challenging. Different variables have been identified as predictors of outcome in the past, including: (a) demographics such as age of onset and sex (Eggers & Bunk, 1997; Fleischhaker et al., 2005; Hafner, Maurer, Löffler, & Riecher-Rössler, 1993; Kydd & Werry, 1982) and family history of psychiatric illness (Fleischhaker et al., 2005); (b) cognitive measures such as IQ (Werry & McClellan, 1992; Werry, McClellan, & Chard, 1991) and processing speed (Milev, Ho, Arndt, & Andreasen, 2005; Puig et al., 2014; Renschmidt & Theisen, 2012); (c) baseline clinical symptoms (Maziade et al., 1996; Meng et al., 2006; Milev et al., 2005); (d) biological markers such as baseline antioxidant status (Fraguas et al., 2012; Martínez-Cengotitabengoa et al., 2014) and neuroimaging markers (Greenstein, Wolfe, Gochman, Rapoport, & Gogtay, 2008; Khodayari-Rostamabad, Hasey, MacCrimmon, Reilly, & de Bruin, 2010; Koutsouleris et al., 2018). The most accurate predictors of relapse and worsening of symptoms include nonadherence with medication (Robinson et al., 1999), poor premorbid adjustment, and substance abuse (Alvarez-Jimenez et al., 2012), yet these predictors are

unreliable in some cases, potentially difficult to ascertain and at most explain a three to fourfold increase in relapse risk. While biomarker discovery in schizophrenia remains a key research area (Singh & Rose, 2009), which specific types of variables are most useful in predicting outcome remains unclear (Díaz-Caneja et al., 2015; Menezes & Milovan, 2000). Accurate, objective and reliable predictors of long-term illness outcome are therefore needed to enable individualized prognostication and treatment.

Functional and structural neuroimaging is a promising candidate for such a predictor that has received significant attention. Several recent studies demonstrate the prognostic utility of training machine learning algorithms to recognize features derived from an individual's neuroimaging data that predict transition to psychosis (Koutsouleris et al., 2009, 2012, 2015; Ramyeed et al., 2016), illness course (Kambeitz-Ilankovic et al., 2016; Koutsouleris et al., 2018; Mourao-Miranda et al., 2012), treatment response (Khodayari-Rostamabad et al., 2010; Koutsouleris, Wobrock, et al., 2018), relapse (Nieuwenhuis et al., 2017), and resilience (de Wit et al., 2016). These studies have sought to compute individual prognostications at follow-ups ranging between 1 and 7 years, primarily using anatomical features derived from structural neuroimaging, such as regional cortical thickness estimates or measures of gray matter volume. Prediction accuracies typically range between 60 and 90% (Janssen, Mourão-Miranda, & Schnack, 2018).

More recently, the prognostic capability of resting-state functional connectivity inferred from functional MRI has emerged in the context of predicting antipsychotic treatment response in individuals experiencing a first episode of psychosis (Cao et al., 2018; Sarpal et al., 2016). However, it remains unclear whether resting-state connectivity enables improved prognostication in a machine learning setting compared to classic neuroimaging measures indexing cortical structure and morphology, such as cortical thickness and gray matter volume, and/or demographic and clinical variables. Furthermore, the dynamic properties of resting-state connectivity particularly warrant investigation as candidate predictors of illness outcome. When

compared to classic time-averaged resting-state functional connectivity approaches, dynamic measures of connectivity yield enhanced sensitivity to disease-related effects (Damaraju et al., 2014) and improved diagnostic accuracy (Kottaram et al., 2018).

Schizophrenia is associated with widespread reductions in resting-state connectivity strength (Bluhm et al., 2007; Fornito, Zalesky, & Breakspear, 2015; Fox & Greicius, 2010; Ganella et al., 2017; Liu et al., 2008; Salvador et al., 2007; Zalesky, Fornito, & Bullmore, 2010), which are associated with the severity of specific symptom dimensions (Hare et al., 2018; Repovs, Csernansky, & Barch, 2011; Rotarska-Jagiela et al., 2010; Venkataraman, Whitford, Westin, Golland, & Kubicki, 2012). In particular, the default-mode network, involved in interoceptive functions (Andrews-Hanna, Smallwood, & Spreng, 2014; Buckner, Andrews-Hanna, & Schacter, 2008; Raichle et al., 2001), has been consistently shown to exhibit aberrant connectivity in schizophrenia (Bluhm et al., 2007; Hu et al., 2017; Jafri, Pearlson, Stevens, & Calhoun, 2008; Liemburg et al., 2012; Liu et al., 2012; Shim et al., 2010; Skudlarski et al., 2010; Whitfield-Gabrieli et al., 2009; Zhou et al., 2007). Assessing the temporal variations in functional connectivity revealed periods of both hypo and hyper-connectivity with the default-mode network (Damaraju et al., 2014; Kottaram et al., 2018) and reduced dynamism in patients (Kottaram et al., 2019; Miller et al., 2016).

In the present study, we aim to predict individual illness outcomes (i.e., worsening or improvement in symptoms) at 1-year follow-up in a cohort of individuals with schizophrenia. We train a simple (linear) machine learning algorithm to predict individual outcomes based on baseline measures of: (a) static and dynamic resting-state connectivity (functional neuroimaging), (b) cortical thickness and gray matter volume (structural neuroimaging), and/or (c) clinical and demographic variables. We aim to determine which of these measures, or combinations thereof, provide the most accurate prediction of illness outcome. We hypothesized that dynamic measures of resting-state connectivity within the default-mode network would provide the most accurate and objective predictors of outcome. Our hypothesis is motivated by recent studies that suggest aberrant switching and connectivity dynamics within the default-mode network are core pathophysiological features of schizophrenia (Damaraju et al., 2014; Du et al., 2016; Kottaram et al., 2018, 2019; Miller et al., 2016; Rashid et al., 2016). This emerging literature led us to specifically target the default-mode network as a predictor, particularly given that: (a) the default-mode network shows the highest levels of activation during rest (Buckner et al., 2008; Raichle et al., 2001); (b) while wide-spread connectivity disruptions have been reported in schizophrenia, the most consistent and robust abnormalities have been reported in this network (Bluhm et al., 2007; Hu et al., 2017; Jafri et al., 2008; Liemburg et al., 2012; Liu et al., 2012; Shim et al., 2010; Skudlarski et al., 2010; Whitfield-Gabrieli et al., 2009; Zhou et al., 2007); (c) from a methodological perspective, the default-mode network can be more easily and robustly detected compared to other networks (Raichle & Snyder, 2007). This is the first study to evaluate the prognostic capability of dynamic measures in resting-state connectivity. Given our modest sample size in the context of machine learning ($N = 30$), we consider this study

preliminary and emphasize that our primary goal is to compare and rank the predictive power of various neuroimaging and clinical variables. This study can inform the choice of predictors (i.e., static vis-a-vis dynamic, functional vis-a-vis structural neuroimaging) investigated in future studies, which should be conducted in larger cohorts that are preferably acquired across multiple sites.

2 | METHODS

2.1 | Participants

Individuals with schizophrenia ($N = 30$; age = 27.3 ± 8.6 years; 10 females) were recruited through Orygen Youth Health clinical sites in Victoria, Australia. All participants had a schizophrenia-spectrum disorder. More specifically, 11 individuals had an established diagnosis of schizophrenia based on DSM-IV, while the remaining 19 individuals had recently experienced a first psychotic episode but were without an established diagnosis. All individuals were prescribed antipsychotic medications at the time of recruitment, image acquisition and follow-up assessment. Exclusion criteria included a history of neurological disease or significant brain injury, a documented history of developmental delay or intellectual disability and presence of any contraindications to MRI. All individuals provided written informed consent prior to participation. The study was approved by the Melbourne Health Human Research Ethics Committee (MHREC ID 2012.069).

Baseline and follow-up assessments of symptom severity, behavior and cognitive ability were completed for all individuals. Baseline assessments were performed within 1 week of image acquisition. Follow-up assessments were performed after an interval of approximately 1 year (372 ± 12.7 days). Positive and negative symptoms were assessed using the Brief Psychiatric Rating Scale (BPRS; Andersen et al., 1989) and general cognitive ability was assessed using the Wechsler Abbreviated Scale of Intelligence (WASI; Wechsler, 1955). Additional assessments were administered but not analyzed in this study. Table 1 shows summary statistics at baseline and follow-up. While group-averaged ratings of symptom severity did not significantly differ between baseline and follow-up (Table 1), marked heterogeneity was evident across individuals (Figure 1).

2.2 | Image acquisition and preprocessing

Functional and structural magnetic resonance imaging (MRI) was performed in all individuals. Acquisition details are provided in Supporting Information S1. In brief, T1-weighted structural (TR = 1980 ms, TE = 4.3 ms and voxel resolution of $0.98 \text{ mm} \times 0.98 \text{ mm} \times 1.0 \text{ mm}$) and resting-state (TR = 2 s, TE = 40 ms and voxel dimensions = $3.3 \text{ mm} \times 3.3 \text{ mm} \times 3 \text{ mm}$, 234 frames over ~ 8 min) images were acquired on a 3T Siemens scanner.

Preprocessing was primarily performed using FSL 5.0.9 (FMRIB Software Library, www.fmrib.ox.ac.uk). Each individual's T1-weighted image was skull-stripped and spatially normalized via nonlinear

TABLE 1 Demographic, behavioral, and clinical characteristics

	Baseline	Follow-up (1 year)	Baseline vs. follow-up comparison
Sex (male/ female)	20/10		
Age (years)	27.3 ± 8.6	28.31 ± 8.6	
Illness duration (years)	3 ± 4.6		
IQ (WASI)	96 ± 19.2		
Education (years)	12.4 ± 3.9		
BPRS positive	8.9 ± 5.1	8.2 ± 4.5	$p > .1$
BPRS negative	5.2 ± 2.3	5.8 ± 2.9	$p > .1$
BPRS total	42.1 ± 11.6	44.1 ± 12.6	$p > .1$
Chlorpromazine equivalent dosage (mg/day)	412.9 ± 237.3	411.0 ± 213.5	$p > .1$

Note: BPRS, Brief Psychiatric Rating Scale; IQ, Intelligence Quotient; WASI, Wechsler Abbreviated Scale of Intelligence.

registration to the Montreal Neurological Institute (MNI) 152 template. Resting-state images were slice-time corrected, realigned to the mean functional image to correct for motion, coregistered to the corresponding T1-weighted scan via rigid-body registration and then warped to the normalized structural image using the previously computed warp transformation. Images were sampled at 2 mm isotropic in MNI space. Images were then spatially smoothed using a Gaussian kernel of full width at half maximum of 4 mm. The smoothed images were further processed with ICA-AROMA (Pruim et al., 2015), an automated tool for motion artifact removal, which has been shown to improve both sensitivity and specificity of connectivity analyses (Parkes, Fulcher, Yücel, & Fornito, 2018; Pruim, Mennes, Buitelaar, & Beckmann, 2015). Signals from white matter and ventricles were then regressed from the voxel time courses and the resulting residuals were used for all subsequent analyses. Finally, any linear trends were removed from each voxel time course and band pass filtering (0.01–0.1 Hz) was performed to suppress the effects of physiological noise (Cordes et al., 2001). Framewise displacement (Power, Barnes, Snyder, Schlaggar, & Petersen, 2012) was computed to quantify the extent of intrascan head movement (0.11 ± 0.04 mm).

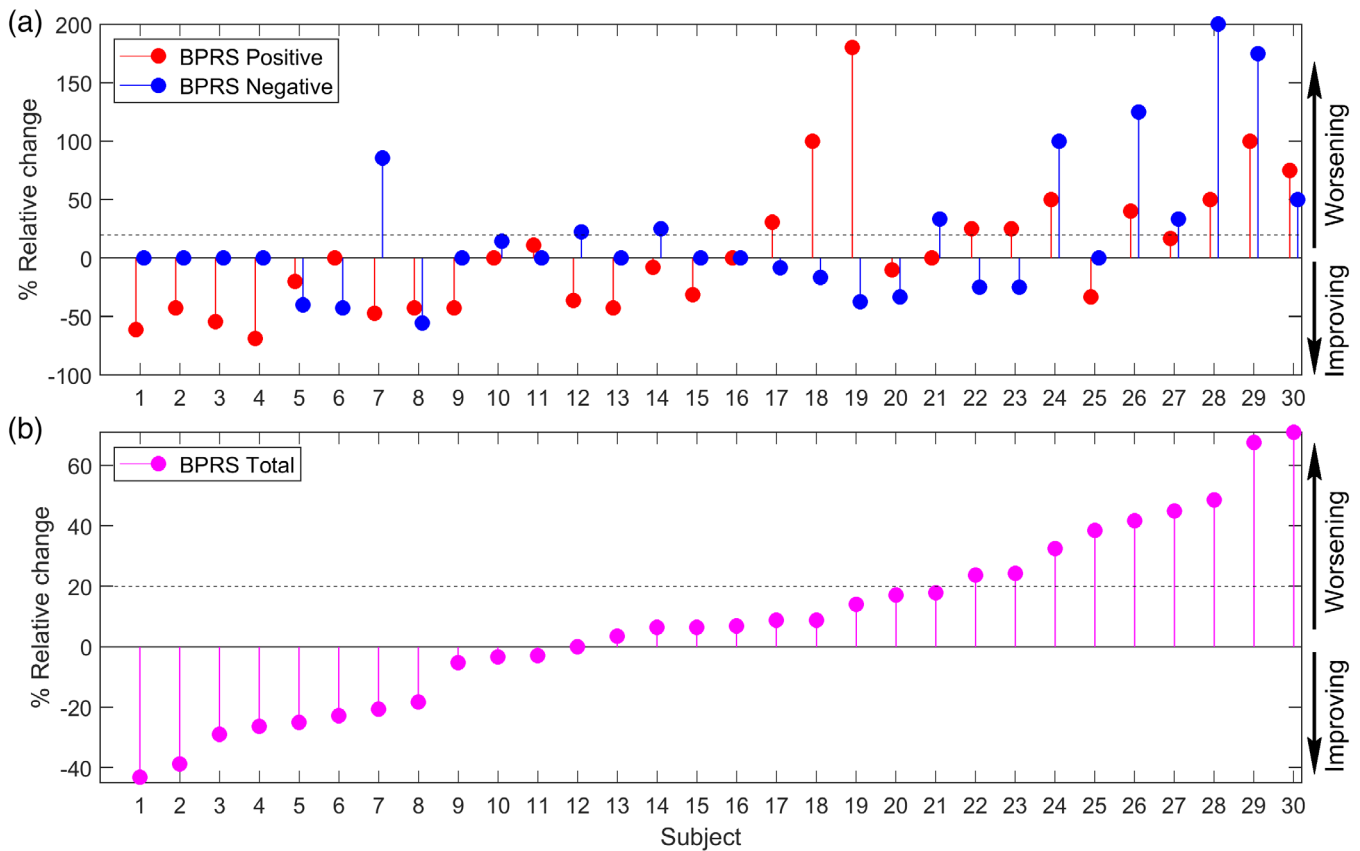


FIGURE 1 Relative change in symptom severity from baseline to 1-year follow-up. (a) Relative change in aggregate scores of positive (red, Δ^+) and negative (blue, Δ^-) symptom severity (b) Relative change in overall symptom severity (Δ^Σ). Individuals ordered from most improved (leftmost) to most worsened (rightmost) in overall symptom severity. Symptom severity assessed with the Brief Psychiatry Rating Scale (BPRS). Binary classifiers were trained to predict individuals that worsened ($\Delta^n > T$) or improved ($\Delta^n < T$) with respect to symptom severity $n \in \{+, -, \Sigma\}$, where cut-off thresholds of $T = 0\%$ (Figure 3) and $T = 20\%$ (Figure S6) were considered. Black dotted lines in Panels a and b correspond to 20% threshold

2.3 | Operationalization of clinical outcome

According to a recently proposed scheme (Dazzi, Shafer, & Lauriola, 2016), aggregate scores indexing the severity of positive (+), negative (-), and overall (Σ) symptoms were computed by summing across relevant BPRS sub-items (positive: 7 BPRS sub-items, negative: 4, total: all sub-items). Further details about sub-item allocation are provided in Supplementary Table 1. For each individual, clinical outcome was operationalized as the relative change in symptom severity between baseline and one-year follow-up. More specifically, if x_t^n denotes the value of aggregate score $n \in \{+, -, \Sigma\}$ at baseline ($t = 0$) and follow-up ($t = 1$), we defined relative change as follows,

$$\Delta^n = \frac{x_1^n - x_0^n}{x_0^n} \times 100\%, \quad n \in \{+, -, \Sigma\}.$$

We classified individuals as either *improving* or *worsening* with respect to Δ^+ , Δ^- , and Δ^Σ . In particular, *worsening of symptoms* was operationalized as $\Delta^n > T$. Individuals not satisfying this criterion were deemed to improve (or remain stable). This operationalization allowed individuals to improve (worsen) with respect to positive symptoms

but worsen (improve) with respect to negative symptoms. While dichotomizing the outcome measure can reduce statistical power, our rationale for doing so is to provide a clear and interpretable measure of outcome that is clinically relevant. Categorizing continuous measure can aid in making sense of unfamiliar measurement scales and treatment effects of uncertain implication (Lewis, 2004).

In this study, the cut-off thresholds of $T = 0\%$ (main analysis) and $T = 20\%$ (supplementary analyses) were investigated. Figure 1 shows the relative change in Δ^+ , Δ^- , and Δ^Σ for each individual.

2.4 | Resting-state functional connectivity

Resting-state functional connectivity was computed between 18 cortical regions comprising the default mode network (DMN; Figure 2a), as defined elsewhere (Andrews-Hanna, Reidler, Sepulcre, Poulin, & Buckner, 2010; Dodell-Feder, DeLisi, & Hooker, 2014; Du et al., 2016; Kucyi & Davis, 2014). Refer to Supporting Information (Section S1, Table S2) for MNI coordinates. For each individual, the processed fMRI data were spatially averaged across the voxels comprising each of the 18 regions, yielding a representative time course

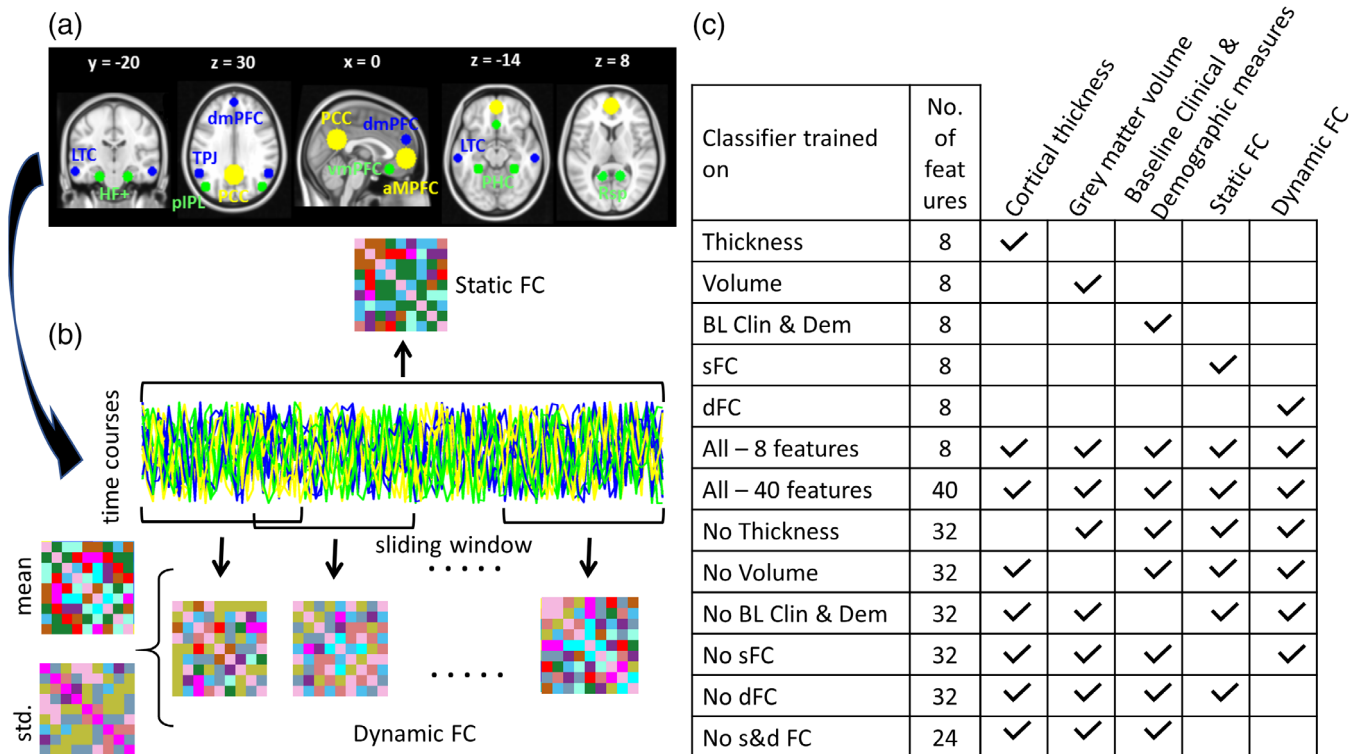


FIGURE 2 Schematic of functional connectivity estimation and feature combinations evaluated for individual prediction of change in symptom severity (a) Cortical regions comprising the default-mode network. Functional connectivity was estimated between all pairs of regions. (b) Schematic of static and dynamic functional connectivity estimation. For each pair of regions, the Pearson correlation coefficient was computed across all time (static) or within overlapping windows (dynamic). (c) List of feature combinations evaluated for individual prediction of change in symptom severity using binary machine learning classifiers. Ticks indicate the presence of a particular class of feature. Feature selection was used to determine the top-8 features within each feature class. Classifiers were trained on the top-8 features from individual feature classes (eight features; Rows 1–5), combination of top-8 features from each of the five feature classes (40 features in total, Row 7) as well as the top-8 features across all feature classes (8 features in total; Row 6). In addition, combinations of the top-8 features from different classes were considered, as indicated (Rows 8–13). dFC, dynamic FC; FC, functional connectivity; sFC, static FC

for each region. A schematic of static and dynamic connectivity estimation is shown in Figure 2b and described below.

2.4.1 | Static connectivity

Static resting-state functional connectivity was computed for each of the $(18 \times 17) / 2 = 153$ unique pairs of DMN regions. To this end, the Pearson correlation coefficient was computed between the representative time courses of each regional pair (Zalesky, Fornito, & Bullmore, 2012). This was repeated independently for each individual.

2.4.2 | Dynamic functional connectivity

Dynamic functional connectivity was computed for each of the 153 unique pairs of DMN regions. Based on our previous study (Kottaram et al., 2018), time was first partitioned into overlapping and contiguous windows of length 20 s, using an overlap of one TR (2 s). The Pearson correlation coefficient was used to compute functional connectivity within each window, yielding a time-resolved (dynamic) measure of connectivity for each regional pair. To reduce the dimensionality of the resulting data, the mean and standard deviation was computed across the set of windows for each pair of regions (Kottaram et al., 2018). Therefore, each individual's connectivity dynamics were characterized in terms of 153 mean values and 153 standard deviations.

2.5 | Gray matter volume and cortical thickness estimation

Gray matter volume and cortical thickness were estimated regionally for each individual using T1-weighted structural images and established software. The Desikan–Killiany atlas (Desikan et al., 2006) was used to define 35 bilateral cortical regions (70 in total). FreeSurfer (version 5.3.0, <https://surfer.nmr.mgh.harvard.edu/>) was then used to estimate cortical thickness and gray matter volume for each region (Dale, Fischl, & Sereno, 1999; Fischl & Dale, 2000; Fischl et al., 2002; Fischl, Sereno, & Dale, 1999). This involved delineating a three-dimensional cortical surface model for each individual, which was then inflated and registered to the Desikan–Killiany brain atlas. All gray/white and gray/pial boundaries were manually examined and corrected by a trained rater to increase the accuracy of the cortical estimates.

2.6 | Baseline clinical and demographic variables

Clinical and demographic measures assessed at baseline were investigated as predictors of change in symptom severity at follow-up. A total of eight baseline variables were considered: aggregate scores indexing the severity of positive, negative, and overall symptoms

(Section 2.3), illness duration, age of onset of psychosis, sex (binary), intelligence quotient (IQ; WASI), and medication (chlorpromazine equivalent dosage, mg/day).

2.7 | Classification

Binary machine learning classifiers were trained to predict individual improvement ($\Delta^n < T$) or worsening ($\Delta^n > T$) with respect to aggregate scores of positive, negative, and overall symptoms, $n \in \{+, -, \Sigma\}$. Under a leave-one-out cross-validation (LOOCV) framework, separate classifiers were trained to predict improvement/worsening in each of the three aggregate scores. Further, classifiers were trained and evaluated using different combinations of features to determine the relative predictive power among different classes of features. We considered five feature classes: (a) cortical thickness (70 regions), (b) gray matter volume (70 regions), (c) baseline clinical and demographic variables (eight variables), (d) static connectivity (153 connections), and (e) dynamic connectivity (153 mean values +153 standard deviations). Feature selection was used to determine the top- K most predictive features within each of the five feature classes. A two-sample t -test was used to perform feature selection within each feature class using only the training data. This involved assessing the between-class separability of individual features, ranking all features based on the t -statistic magnitude and then choosing the top- K features. Among all feature classes, the baseline clinical and demographic variables contained the least number of features ($n = 8$); hence the number of features chosen was set to $K = 8$ in the main analysis. Since our aim was to compare the predictive power of different feature classes, no further analysis was performed to optimize the value of K ; however, all analyses were repeated with $K = 20$ to investigate the impact of variation in the total number of features used. Importantly, in each iteration of the LOOCV, features were selected exclusively based on the training sample ($N = 29$), leaving the testing sample ($N = 1$) aside, thereby avoiding cross-contamination of data. Using a relatively small set of features (i.e., $K = 8$) minimized the risk of overfitting, although feature selection per se does not eliminate this risk.

To predict the improvement/worsening of each aggregate measure, 13 different classifiers were trained using different combinations of features, as listed in Figure 2c. A separate classifier was trained and evaluated for each of the five feature classes. Furthermore, classifiers were trained on the top- K features across all feature classes (K features in total) as well as the top- K features from each of the five feature classes ($5 \times K$ features in total). In addition, to quantify the detriment of excluding a particular feature class, classifiers were trained on the top- K features from 4 of the 5 feature classes ($4 \times K$ features in total). Finally, a classifier was trained using the top- K features from all feature classes except the connectivity-based (static and dynamic) features ($3 \times K$ features in total).

Given the modest sample size of the cohort studied, linear discriminant analysis (LDA) was employed; LDA models the distribution of features separately for each response class and minimizes overfitting, thus being particularly suited to small datasets (Izenman, 2013).

Classification results can vary with the choice of machine learning algorithm and for this reason, ancillary analyses were then performed using other linear (support vector machine, SVM) and nonlinear (decision trees) classifiers. For each LOOCV analysis, established measures of classification performance were computed, including prediction accuracy, specificity, sensitivity, precision, area under the curve (AUC), and the *F*-measure were calculated.

The feature selection described above is a univariate approach, as individual features are ranked independent of other features. We have also performed the classification experiments using a different feature selection method based on principal component analysis (PCA), which is further described in Supporting Information (Section S3).

3 | RESULTS

3.1 | Predicting individual change in symptom severity

Relative change in overall symptom severity between baseline and 1-year follow-up varied markedly among individuals (interquartile range: 55%; Figure 1), consistent with the heterogeneity in outcomes

that is characteristic of schizophrenia. Some individuals worsened with respect to positive symptoms but showed an improvement in negative symptoms, and vice versa. Individuals were classified as either: (a) worsening or (b) improving (or remaining stable) with respect to positive (improving: $N = 18$), negative (improving: $N = 19$), and overall symptom severity (improving: $N = 12$) at 1-year follow-up. Classifiers were then trained to predict these individual outcomes using different combinations of features (predictors), with the aim of determining which of functional neuroimaging (static and dynamic resting-state connectivity), structural neuroimaging (regional gray matter volume and cortical thickness) or baseline demographic and clinical variables, or combinations thereof, provide the most accurate prognostic utility.

Figure 3 shows prediction accuracies for the range of feature combinations indicated in Figure 2c. Combining features across all feature classes (i.e., classifiers labeled “All-8 features” in Figure 3a) yielded the highest accuracies in predicting change in overall symptoms. Classification accuracies consistently exceeded 80% when features were combined across all feature classes. Remarkably, classifiers trained using *only* features associated with dynamic resting-state connectivity yielded accuracies that were comparable to combining features across *all* feature classes. In contrast, the other single-class classifiers trained using only one feature class yielded poorer

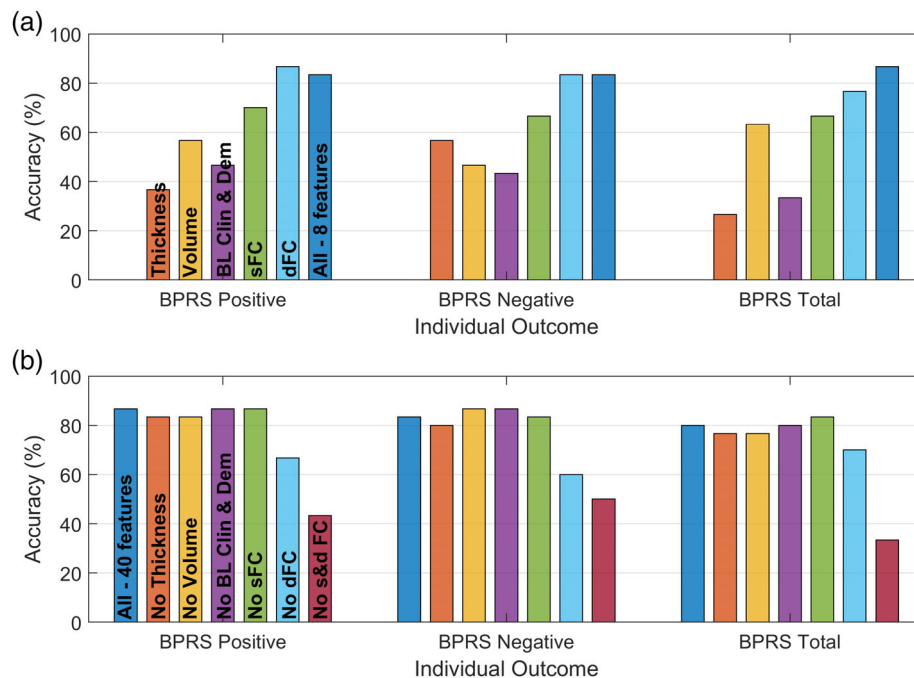


FIGURE 3 Accuracy of predicting improvement in symptom severity at 1-year follow-up using connectivity, structural, and/or clinical features. Individuals were classified as either worsening or improving/stabilizing based on relative change in positive, negative and overall symptom severity. Binary classifiers (linear discriminant analysis) were then trained to predict individual outcome using different combinations of features. (a) Prediction accuracies for classifiers trained using features from only *one* of the following five feature classes: cortical thickness, gray matter volume, baseline (BL) clinical, and demographic variables, static resting-state functional connectivity (sFC) and dynamic resting-state functional connectivity (dFC). A separate classifier was also trained using the top-8 features selected from each of the above five feature classes as well as the combination of all of them (All-8 features). (b) Prediction accuracies for classifiers trained using 5 (All-40 features), 4 (No Thickness, No Volume, No BL Clin & Dem, No sFC, and No dFC), and 3 (No s and d FC) of the five feature classes. BPRS, Brief Psychiatric Rating Scale; FC, functional connectivity

accuracies (<70%, Figure 3a) and often did not yield predictions that exceeded chance level.

The predictive utility of dynamic resting-state connectivity computed within the default-mode network is further evidenced by the substantial reduction in prediction accuracies that were apparent when this feature class was excluded from the classifier combining all feature classes (Figure 3b). Excluding dynamic resting-state connectivity resulted in a 10–20% reduction in accuracy, while excluding both static and dynamic connectivity features resulted in the most substantial accuracy reduction (>30% reduction). In contrast, excluding features related to cortical thickness, gray matter volume or baseline demographic and clinical variables did not affect prediction accuracies by more than 5% (Figure 3b). Table 2 shows additional measures that quantify the accuracy of the classifiers trained using dynamic resting-state connectivity.

Other performance measures evaluated, such as classification sensitivity (Figure S1), specificity (Figure S2), and AUC (Figure S3) accord with the above findings, suggesting that our classification models were not biased to predicting either improving or worsening outcomes alone. Selecting the top-20 (instead of top-8) features within each feature class did not markedly alter prediction accuracies (Figure S4). Compared to LDA (Figure 3), prediction accuracies decreased when using decision trees (Figure S5) or linear support vector machines (Figure S6). Despite this decrease, the relative ranking of features remained consistent, with dynamic resting-state connectivity consistently providing the most accurate prognostic utility. Further, increasing the cut-off threshold to $T = 20\%$ (instead of $T = 0\%$) resulted in a marginal reduction (<5%) in the accuracy of predicting change in overall symptom severity, although prediction of change in positive and negative symptoms was unaffected. Also, the analysis with PCA-based feature selection yielded similar results (Figure S8), consistent with the main analysis reported in Figure 3.

Finally, the variation in prediction accuracy in the case of holding out larger fractions of data for testing was evaluated. While all analyses so far employed a LOOCV approach, the held-out sample size was systematically increased from 1 to 6 and the accuracies are reported in Figure S8. Only classifiers trained on dFC were considered in this supplementary analysis, given that this was the best performing feature class. While the accuracies reduced in general as the test

sample size increased (thus yielding a smaller training sample), the accuracies remained >70%, despite holding out six samples for testing.

3.2 | Characteristics of connections predicting outcome

Having found that dynamic resting-state connectivity computed within an individual's default-mode network yielded the most accurate prognostic utility, we next aimed to determine the characteristics and locations of the specific connections that were most informative with respect to prediction. We assumed a cut-off threshold of $T = 0\%$ in this section to operationalize improving and worsening subgroups of individuals with respect to relative change in symptom severity. Given that feature selection varied between iterations of the cross-validation process, for this analysis, feature selection was performed on the full sample without any individual prediction. Instead, the top-20% of connections associated with the largest effect sizes (two-sample t -test) were identified with respect to the null hypothesis of equality between the improving and worsening subgroups. In particular, a t -test was computed independently for each pair of regions (i.e., each connection); such that a positive t -statistic indicated that the worsening group showed increased connectivity at baseline, relative to the improving group.

Figure 4 shows the location of connections with the largest effect sizes (top-20%), where positive effect sizes (warm colors) indicate increased connectivity strength (left column) and increased connectivity dynamics (standard deviation, right column) in the worsening group at baseline, relative to the improving group. In contrast, negative effect sizes (cool colors) indicate increased connectivity strength and increased connectivity dynamics in the improving group at baseline. It can be seen that worsening positive symptoms are primarily predicted by decreased connectivity strength but increased temporal dynamics among a distributed sub-network spanning the entire DMN (Figure 4, uppermost row). Notably, connection strengths between ventral medial prefrontal cortex (vmPFC) and bilateral temporal poles, and between left temporal pole and right temporal parietal junction (TPJ) were lower in the worsening group. In addition, connectivity of left parahippocampal cortex (PHC), left temporal pole and left retrosplenial cortex to most other regions of DMN were reduced. In contrast, worsening negative symptoms are predominantly predicted by increased strength but decreased dynamics within a separate DMN sub-network (Figure 4, center row). Connectivity of anterior medial prefrontal cortex (amPFC) to posterior cingulate cortex (PCC) and vmPFC to PHC was greater in the worsening group. In addition, left lateral temporal cortex (LTC) was strongly connected to other regions in the left hemisphere. The network that predicted worsening of negative symptoms appears to be largely localized to the left hemisphere, while the network predicting worsening of positive symptoms comprised many bilateral regions.

Figure 4 suggests a marked contrast between predictors of change in positive versus negative symptom severity. In particular,

TABLE 2 Performance of predicting improvement in symptom severity at 1-year follow-up for classifiers trained using dynamic resting-state connectivity

Performance measure	Positive symptoms	Negative symptoms	BPRS total
Accuracy (%)	86.67	83.33	76.67
Sensitivity (%)	75	81.82	72.22
Specificity (%)	94.44	84.21	83.33
Precision (%)	90	75	86.67
AUC	0.85	0.83	0.78
F-measure (%)	81.82	78.26	78.79

Note: AUC—area under the receiver operating characteristic (ROC) curve.

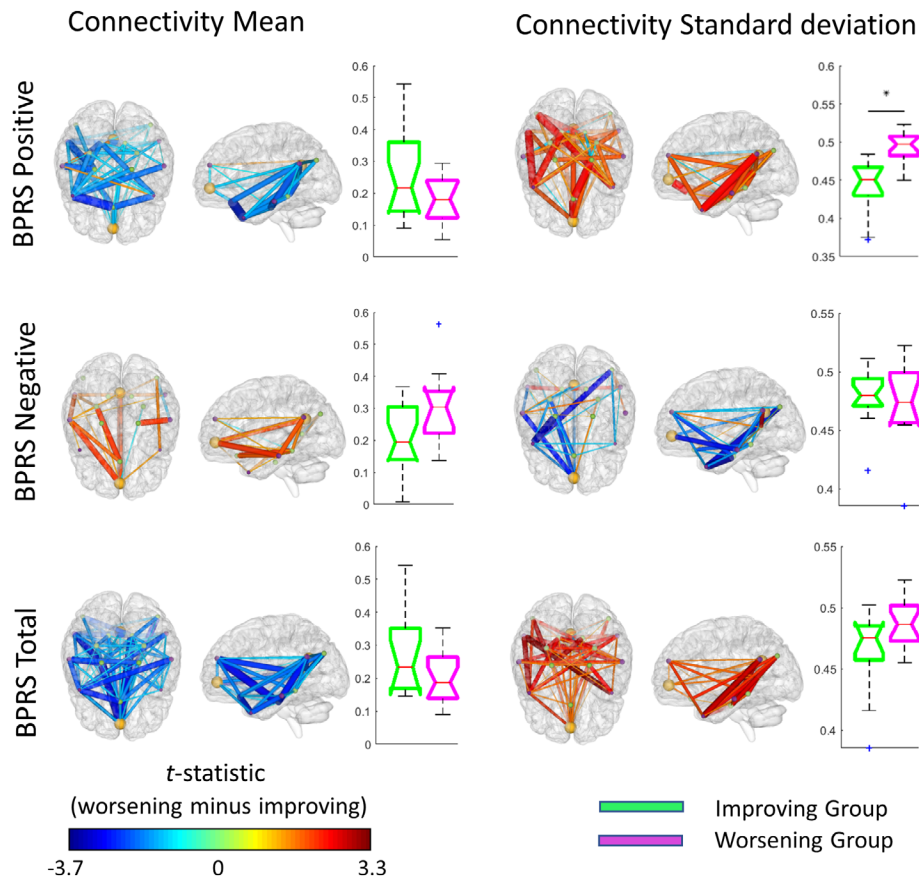


FIGURE 4 Connections most informative in predicting improvement in symptom severity at 1-year follow-up. Individuals were classified as either improving or worsening with respect to positive, negative, and overall symptom severity. Two-sample t -tests were performed to identify which connections differed most in mean connectivity strength (left column) and connectivity dynamics (standard deviation, right column) between the improving and worsening subgroups. The top-20% of connections according to t -statistic magnitude are shown. Note that t -statistic magnitude is proportional to effect size. Each sphere represents a region within the default mode network (depicted in Figure 2a). The color and thickness of each connection represents the t -statistic, as per the color bar. Boxplots represent the distribution of mean connection strength (left column) or standard deviation of connection strengths (right column), averaged over all regions within the default-mode network. “**” indicates significant between-group differences ($p < .05$). The top-20% connections should not be interpreted as significant but are rather visualized here to identify those connection strengths which lead to high classification accuracies

worsening positive symptoms (compared to improving positive symptoms) were predicted by hypo-connectivity and hyper-dynamism in baseline resting-state connectivity, while worsening negative symptoms (compared to improving negative symptoms) were predicted by hyper-connectivity and hypo-dynamism. We hypothesize that hyper-connectivity within the DMN predicted worsening negative symptoms because increased engagement of this network has been associated with increased internal, self-related processing (Molnar-Szakacs & Uddin, 2013; Qin & Northoff, 2011), which is characteristic of rumination and other internalizing negative symptoms of schizophrenia (see Section 4).

In ancillary analyses, statistical inference was performed using the network-based statistic (NBS) (Zalesky et al., 2010) to correct for multiple comparisons (i.e., 153 independent t -tests). This enabled formal identification of the locations of connections that predicted outcome with respect to a family-wise error corrected p -value criterion ($p < .05$), rather than simply investigating the top-20% as above. The primary t -statistic threshold for the NBS was chosen as 2 to ensure

that any significant effects exceeded an effects size (Cohen's d) of approximately 0.3, which is deemed a medium effect, and a total of 5,000 permutations were generated to estimate p -values. This was repeated for changes in positive, negative and overall symptom severity. For overall symptom severity, the NBS identified subnetworks of connections that significantly differentiated the worsening and improving groups with respect to connectivity strength ($p < .05$) and dynamics ($p < .05$; Figure S7). These subnetworks were highly consistent with the top-20% of connections (Figure 4; bottommost row). However, no significant effects were found with respect to positive and negative symptoms ($p > .05$).

4 | DISCUSSION

Mounting evidence suggests that neuroimaging coupled with machine learning can provide utility in guiding long-term prognostication in schizophrenia, and neuropsychiatry more generally; see Janssen and

colleagues (Janssen et al., 2018) for a comprehensive review. However, it remains unclear: (a) whether neuroimaging can provide superior prognostic utility compared to clinical variables such as baseline symptom ratings; and (b) which neuroimaging modalities, or combinations thereof, provide the most accurate prognostic utility. To address these knowledge gaps, the present study aimed to determine which of three broad classes of candidate predictors would yield the most accurate prognostic utility in predicting relative change in schizophrenia symptom severity at 1-year follow-up.

We found that the dynamics of resting-state functional connectivity within the default-mode network (DMN) provided the most accurate prediction of whether an individual worsened or improved with respect to positive, negative and overall symptom severity at 1-year follow-up. Remarkably, while previous studies have reported prognostic utility of gray matter volume and cortical thickness measures (Greenstein et al., 2008; Khodayari-Rostamabad et al., 2010; Koutsouleris, Wobrock, et al., 2018), prediction accuracies showed minimal improvement when incorporating structural neuroimaging predictors (regional cortical thickness and gray matter volume), static resting-state connectivity and/or baseline demographic and clinical variables. The top-8 selected features from each feature group for predicting worsening/improvement of positive, negative and overall symptoms are listed in Tables S3–S5. In particular, for the combination of all feature groups, the top-8 selected features are mostly those from dFC. This explains why we did not observe any improvement in accuracy for the “All-8 features.” However, note that even the classifier trained on the combination of top-8 features from different feature groups (“All-40 features”) did not provide any significant improvement in classification accuracies (Figure 3b). Similarly removing volume, thickness, and baseline clinical and demographic variables did not affect the classifier performance. In the absence of dynamic resting-state connectivity information, these predictors alone yielded moderate-to-poor prognostic utility (<60% accuracy) that often did not exceed chance-level predictions. A notable exception was static resting-state connectivity, which yielded modest prediction accuracies (60–70%). The difference in predictive prognostic powers of sFC and dFC measures supports the notion that static versus dynamic connectivity might be capturing different aspects of physiology (Liegeois, Laumann, Snyder, Zhou, & Yeo, 2017). While regional measures of brain structure did not substantially improve prediction performance here, further investigation is needed in larger cohorts to establish whether change in specific symptom dimensions and cognition is best predicted by structure and function.

Our best performing machine classifiers were able to predict whether an individual improved or worsened in terms of symptom severity with accuracies ranging between 75 and 90%. However, these accuracies must be interpreted with caution given the modest sample size ($N = 30$), homogeneity of the sample (i.e., all individuals were recruited from the same clinical service) and general limitations of cross validation, particularly the risk of unreliable estimates with the LOOCV procedure used here (Varoquaux, 2018; Varoquaux et al., 2017). We emphasize that the main goal of the present study was not to provide precise estimates of prediction accuracy, but rather to rank the relative

predictive power of various connectomic, structural and clinical predictors of individual outcome. If prediction accuracies were overestimated due to the limited sample size, any potential overestimation was most likely uniform across all feature classes, and thus the relative ranking of features (predictors) would remain unchanged.

While this is the first study to utilize connectivity dynamics to predict outcome, it is important to remark that several recent studies have investigated the prognostic utility of static resting-state functional connectivity. Most of these studies have considered prediction of response to antipsychotic medication (Cadena et al., 2018; Cadena, White, Kraguljac, Reid, & Lahti, 2018; Cao et al., 2018; Kraguljac et al., 2016; Sarpal et al., 2015). For example, using task-based fMRI data and a seed-based analysis, Cadena and colleagues (Cadena, White, Kraguljac, Reid, Jindal, et al., 2018; Cadena, White, Kraguljac, Reid, & Lahti, 2018) found that connectivity between the anterior cingulate cortex and putamen was associated with treatment response over a 6-week duration. Another study using resting-state data and seed-based analysis found that the connectivity of striatal regions to other brain regions was correlated with changes in symptoms over a 12-week period (Sarpal et al., 2015). In addition, Cao and colleagues (Cao et al., 2018) reported that striatal connectivity was a predictor of response to antipsychotic medication over a period of 10 weeks in a medication-naïve cohort.

We considered a significantly longer follow-up period (1 year) than these recent studies; and furthermore, all individuals in our study were prescribed antipsychotic medication at baseline and follow-up. Hence, our predictions were with respect to relatively long-term change in symptom severity, as opposed to short-term response to pharmacological treatment. It may be argued that accurately predicting treatment response is a more beneficial clinical capability than predicting long-term change in symptom severity. However, it is important to remark that long-term outcome is an important prognostic consideration given that schizophrenia can often become a chronic and lifelong disorder and such predictive information may orient treatment and follow-up strategies, limiting relapses.

We found certain differences in dynamic connectivity features that predicted worsening positive versus negative symptoms (Figure 4, Tables S3 and S4). Individuals with worsening positive symptoms were predicted by hypo-connectivity and hyper-dynamism within the DMN at baseline, compared to individuals with improving positive symptoms. The converse pattern was evident for negative symptoms (i.e., hyper-connectivity and hypo-dynamism at baseline predicted worsening negative symptoms). While the reasons for this dissociation is unclear, it is consistent with the observation that several individuals improved with respect to one symptom dimension but worsened with respect to others (Figure 1). One hypothesis is that positive and negative symptom dimensions tap distinct and possibly mutually exclusive subsystems of the DMN, and thus predictors of change in positive and negative symptom severity circumscribe distinct loci. This notion is in line with the observation that the subnetwork predicting change in positive symptom severity (Figure 4, topmost row) shows little overlap with the connections predicting negative symptoms (center row). Worsening of negative symptoms was predicted by increased connectivity of a left-lateralized

sub-network with direct involvement of hub regions such as PCC and amPFC, while for positive symptoms, a hypoconnectivity within a bilateral network without the participation of the hub nodes predicted worsening. However, the between-group comparison of connection strengths using the NBS did not reveal any significant differences between the improving and worsening groups.

We hypothesize that this distinction between the predictors of worsening positive versus negative symptoms may relate to the putative introspective and extrospective modes of the DMN. The DMN toggles between an introspective self-referential mode and an extrospective mode that remains alert to changes in the external environment (Fransson, 2005, 2006), suggesting a functional segregation within DMN regions (Fransson, 2005; Hearne, Cocchi, Zalesky, & Mattingley, 2015; Uddin, Clare Kelly, Biswal, Xavier Castellanos, & Milham, 2009). Hyperactivity in prefrontal regions characterizes internalization and self-reflective thinking (Fransson, 2005; Ingvar, 1979), which is frequently interrupted and shifted toward an externalized state vigilant to external stimuli (Fransson, 2005; Laufs et al., 2003). Therefore, frontal DMN hyper-connectivity may predict worsening of negative symptoms due to increased internalization and fewer shifts to the externalized mode, which are characteristic of the internalizing negative symptoms of schizophrenia. Disengagement among DMN regions (i.e., hypo-connectivity) on the other hand leads to externalization (Farina et al., 2018), which may be necessary to engender delusions, hallucinations and other positive symptoms.

While we recommend that the dynamics of resting-state functional connectivity should be given priority when selecting neuroimaging predictors of schizophrenia outcome in future studies, the precise biological and pathological correlates of connectivity dynamics are unknown. In this study, connectivity dynamics were summarized using the standard deviation computed across multiple temporal windows. Therefore, larger temporal fluctuations in functional connectivity strength were indexed by larger standard deviations. Connectivity dynamics could potentially index the rate at which the DMN cycles between putative internalizing and externalizing modes, where an abnormal switching rate may place individuals at greater risk of poor outcome. The hypo-dynamism associated with worsening of negative symptoms (i.e., low standard deviation) suggests that individuals do not (or are unable to) freely switch between externalizing and internalizing states, consistent with internalization and rumination.

It is worth noting that worsening and improvement were defined as relative deviations from baseline scores in the present study. While this is a commonly followed approach (Cao et al., 2018; Leucht, Davis, Engel, Kane, & Wagenpfeil, 2007; Schennach, Riedel, Musil, & Möller, 2012), this definition does not take into account the actual symptom values. For example, as per this definition, an individual with a low score at baseline and a slightly higher score at follow-up will be considered as worsening, even though the actual score at follow-up might still be relatively low. Similarly, a slight improvement in symptoms when the actual scores are high does not imply a good outcome. Note that we performed a supplementary analysis where a threshold of $T = 20\%$ was considered, such that an individual will be considered as worsening only if the percentage increase in follow-up symptoms is

>20%. This analysis yielded similar results (Figure S7), consistent with the main analysis ($T = 0\%$) reported in Figure 3. However, we considered only percentage changes in symptoms in this study and no absolute changes were considered. Future studies are warranted to assess if the neuroimaging-based classifiers are predictive of symptom severity (above or below a clinically meaningful cut-off) or only sensitive to relative change in symptoms.

Several limitations should be considered. First, our sample size is modest for a machine learning study. It is worth mentioning that acquiring longitudinal neuroimaging data with long follow-up periods is difficult in psychiatric cohorts. While we followed a cross-validation framework which reduces the risk of over-fitting, it does not eliminate the risk completely, particularly when all data are acquired from the same site. Intersite cross-validation in which one acquisition site is used for training and another for testing would have further mitigated this risk. Given this, the reported prediction accuracies must be interpreted cautiously, particularly with respect to the high variability of cross-validation estimates when applied to typically sized neuroimaging datasets (Schnack & Kahn, 2016; Varoquaux et al., 2017). For this reason, we used a simple linear classifier to minimize the risk of over-fitting and principally focused on ranking the feature classes investigated, as opposed to interpreting the prediction accuracies in an absolute sense. We also used a relatively small set of predictors to further reduce the risk of overfitting as well as domain-specific knowledge to select the regions and networks considered. Second, the feature class that was found to yield the most accurate predictions was also the class with the greatest number of features (153 mean values + 153 standard deviations), and thus the risk of overfitting may have been increased for this class. However, feature selection was performed to choose the same number of features across different classes. Moreover, prediction accuracy did not markedly improve when combining features from the five feature classes investigated, suggesting that the influence of overfitting on accuracy was possibly minimal; otherwise the classifier combining features across all classes would have yielded the highest accuracy. Third, our feature class with clinical and behavioral measures was limited to the information that was collected during data acquisition, which does not include all clinical/behavioral variables. There could be other variables with better predictive power that were not collected and included in this study. For example, incorporating structural connectivity inferred from diffusion-weighted MRI into the prediction model could potentially enhance accuracy by explaining unique variance that is not captured by connectivity dynamics. Fourth, all individuals were prescribed anti-psychotic medication and medication represents a potential confounding effect on structural and functional neuroimaging. The adherence to medication and certain environmental factors such as substance use during the period between baseline and follow-up examinations affect the outcomes, but these cannot be accounted in a predictive model based on the baseline measures alone. Also, no information regarding adherence to medication was collected during the follow-up assessment; hence we cannot comment on the extent to which the current findings relate to treatment response. Fifth, all participants were recruited from the same clinical service, resulting in

a geographically and ethnically homogeneous cohort. While this homogeneity is beneficial in alleviating confounds inherent to interindividual variation, it is unclear whether the classifiers trained will generalize to other cohorts (Hahn, Ebner-Priemer, & Meyer-Lindenberg, 2019). Finally, the use of sliding windows to estimate connectivity dynamics has been criticized (Hindriks et al., 2016; Laumann et al., 2016), although we show that this methodology can yield predictors that outperform cortical thickness and gray matter volume.

5 | CONCLUSIONS

Functional neuroimaging provides improved prognostic utility in predicting long-term changes in schizophrenia symptom severity, compared to predictors derived from structural neuroimaging and baseline clinical variables. While our results might not necessarily generalize to other cohorts due to the modest sample size, our study can inform and constrain the choice of candidate predictors investigated in future machine learning studies that aim to predict illness outcome using larger cohorts. In particular, we recommend that dynamic fluctuations in resting-state functional connectivity, particularly within the default-mode network, are given prime consideration as candidate predictors in future studies.

ACKNOWLEDGMENTS

We gratefully acknowledge all participants for making this study possible. In addition, we thank Steven Tahtalian, Maria Di Biase, and Cassandra Wannan for assisting with data collection and Drs. Phassouliotis and Toni Merritt for study coordination. We are thankful for the high-performance computing facility (Spartan) provided by the University of Melbourne, Australia (project ID: punim0021). This study was supported by an Australian National Health and Medical Research Council (NHMRC) (grant number 1065742) and a University of Melbourne Early Career Researcher grant awarded to Vanessa Cropley (grant number 601253). Vanessa Cropley was supported by an NHMRC Early Career Fellowship (grant number 628880). Christos Pantelis was supported by NHMRC Senior Principal Research Fellowships (grant numbers 628386 and 1105825). Andrew Zalesky was supported by an NHMRC Senior Research Fellowship B (grant number APP1136649). Luca Cocchi was supported by two Project Grants (grant numbers APP1099082 and APP1138711) from the NHMRC.

CONFLICT OF INTEREST

The authors do not have any conflict of interest to declare.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study were sourced from the Melbourne Neuropsychiatry Centre (<https://www.mncresearch.org/>; Level 3, Alan Gilbert Building, 161 Barry Street, Carlton, Victoria, 3053, Australia). Restrictions apply to the availability of these data, which were used under license for this study. Requests to access data should be directed in writing to Dr Vanessa Cropley (vcropley@unimelb.edu.au) and will be subject to relevant ethical approvals.

ORCID

Akhil Kottaram  <https://orcid.org/0000-0003-2446-2704>

Luca Cocchi  <https://orcid.org/0000-0003-3651-2676>

REFERENCES

- Alvarez-Jimenez, M., Priede, A., Hetrick, S. E., Bendall, S., Killackey, E., Parker, A. G., ... Gleeson, J. F. (2012). Risk factors for relapse following treatment for first episode psychosis: A systematic review and meta-analysis of longitudinal studies. *Schizophrenia Research, 139*(1–3), 116–128. <https://doi.org/10.1016/j.schres.2012.05.007>
- Andersen, J., Larsen, J. K., Schultz, V., Nielsen, B. M., Kørner, A., Behnke, K., ... Bech, P. (1989). The brief psychiatric rating scale. *Psychopathology, 22*(2–3), 168–176. <https://doi.org/10.1159/000284591>
- Andrews-Hanna, J. R., Reidler, J. S., Sepulcre, J., Poulin, R., & Buckner, R. L. (2010). Functional-anatomic fractionation of the brain's default network. *Neuron, 65*(4), 550–562. <https://doi.org/10.1016/j.neuron.2010.02.005>
- Andrews-Hanna, J. R., Smallwood, J., & Spreng, R. N. (2014). The default network and self-generated thought: Component processes, dynamic control, and clinical relevance. *Annals of the New York Academy of Sciences, 1316*(1), 29–52. <https://doi.org/10.1111/nyas.12360>
- Arana, G. W. (2000). An overview of side effects caused by typical antipsychotics. *The Journal of Clinical Psychiatry, 61*(Suppl 8), 5–11; discussion 12–13.
- Bluhm, R. L., Miller, J., Lanius, R. A., Osuch, E. A., Boksman, K., Neufeld, R. W. J., ... Williamson, P. (2007). Spontaneous low-frequency fluctuations in the BOLD signal in schizophrenic patients: Anomalies in the default network. *Schizophrenia Bulletin, 33*(4), 1004–1012. <https://doi.org/10.1093/schbul/sbm052>
- Brujnzeel, D., Suryadevara, U., & Tandon, R. (2014). Antipsychotic treatment of schizophrenia: An update. *Asian Journal of Psychiatry, 11*, 3–7. <https://doi.org/10.1016/j.ajp.2014.08.002>
- Buckner, R. L., Andrews-Hanna, J. R., & Schacter, D. L. (2008). The brain's default network. *Annals of the New York Academy of Sciences, 1124*(1), 1–38. <https://doi.org/10.1196/annals.1440.011>
- Cadena, E. J., White, D. M., Kraguljac, N. V., Reid, M. A., Jindal, R., Pixley, R. M., & Lahti, A. C. (2018). Cognitive control network dysconnectivity and response to antipsychotic treatment in schizophrenia. *Schizophrenia Research, 204*, 262–270. <https://doi.org/10.1016/j.schres.2018.07.045>
- Cadena, E. J., White, D. M., Kraguljac, N. V., Reid, M. A., & Lahti, A. C. (2018). Evaluation of fronto-striatal networks during cognitive control in unmedicated patients with schizophrenia and the effect of antipsychotic medication. *NPJ Schizophrenia, 4*(1), 8. <https://doi.org/10.1038/s41537-018-0051-y>
- Cao, B., Cho, R. Y., Chen, D., Xiu, M., Wang, L., Soares, J. C., & Zhang, X. Y. (2018). Treatment response prediction and individualized identification of first-episode drug-naïve schizophrenia using brain functional connectivity. *Molecular Psychiatry, 25*, 906–913. <https://doi.org/10.1038/s41380-018-0106-5>
- Cha, D. S., & McIntyre, R. S. (2012). Treatment-emergent adverse events associated with atypical antipsychotics. *Expert Opinion on Pharmacotherapy, 13*(11), 1587–1598. <https://doi.org/10.1517/14656566.2012.656590>
- Cordes, D., Haughton, V. M., Arfanakis, K., Carew, J. D., Turski, P. A., Moritz, C. H., ... Meyerand, M. E. (2001). Frequencies contributing to functional connectivity in the cerebral cortex in “resting-state” data. *AJNR. American Journal of Neuroradiology, 22*(7), 1326–1333.
- Cropley, V. L., & Pantelis, C. (2014). Using longitudinal imaging to map the ‘relapse signature’ of schizophrenia and other psychoses. *Epidemiology and Psychiatric Sciences, 23*(3), 219–225. <https://doi.org/10.1017/S2045796014000341>
- Dale, A. M., Fischl, B., & Sereno, M. I. (1999). Cortical surface-based analysis. *NeuroImage, 9*(2), 179–194. <https://doi.org/10.1006/nimg.1998.0395>

- Damaraju, E., Allen, E. A., Belger, A., Ford, J. M., McEwen, S., Mathalon, D. H., ... Calhoun, V. D. (2014). Dynamic functional connectivity analysis reveals transient states of dysconnectivity in schizophrenia. *NeuroImage: Clinical*, 5, 298–308. <https://doi.org/10.1016/j.nicl.2014.07.003>
- Dazzi, F., Shafer, A., & Lauriola, M. (2016). Meta-analysis of the brief psychiatric rating scale – expanded (BPRS-E) structure and arguments for a new version. *Journal of Psychiatric Research*, 81, 140–151. <https://doi.org/10.1016/j.jpsychires.2016.07.001>
- de Wit, S., Wierenga, L. M., Oranje, B., Ziermans, T. B., Schothorst, P. F., van Engeland, H., ... Durston, S. (2016). Brain development in adolescents at ultra-high risk for psychosis: Longitudinal changes related to resilience. *NeuroImage: Clinical*, 12, 542–549. <https://doi.org/10.1016/j.nicl.2016.08.013>
- Desikan, R. S., Ségonne, F., Fischl, B., Quinn, B. T., Dickerson, B. C., Blacker, D., ... Killiany, R. J. (2006). An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *NeuroImage*, 31(3), 968–980. <https://doi.org/10.1016/j.neuroimage.2006.01.021>
- Díaz-Caneja, C. M., Pina-Camacho, L., Rodríguez-Quiroga, A., Fraguas, D., Parellada, M., & Arango, C. (2015). Predictors of outcome in early-onset psychosis: A systematic review. *NPJ Schizophrenia*, 1(1), 14005. <https://doi.org/10.1038/npschz.2014.5>
- Dodell-Feder, D., DeLisi, L. E., & Hooker, C. I. (2014). The relationship between default mode network connectivity and social functioning in individuals at familial high-risk for schizophrenia. *Schizophrenia Research*, 156(1), 87–95. <https://doi.org/10.1016/j.schres.2014.03.031>
- Du, Y., Pearson, G. D., Yu, Q., He, H., Lin, D., Sui, J., ... Calhoun, V. D. (2016). Interaction among subsystems within default mode network diminished in schizophrenia patients: A dynamic connectivity approach. *Schizophrenia Research*, 170(1), 55–65. <https://doi.org/10.1016/j.schres.2015.11.021>
- Eggers, C., & Bunk, D. (1997). The long-term course of childhood-onset schizophrenia: A 42-year followup. *Schizophrenia Bulletin*, 23(1), 105–117. <https://doi.org/10.1093/schbul/23.1.105>
- Emsley, R., Chiliza, B., Asmal, L., & Harvey, B. H. (2013). The nature of relapse in schizophrenia. *BMC Psychiatry*, 13, 50. <https://doi.org/10.1186/1471-244X-13-50>
- Farina, B., Della Marca, G., Maestoso, G., Amoroso, N., Valenti, E. M., Carbone, G. A., ... Imperatori, C. (2018). The association among default mode network functional connectivity, mentalization, and psychopathology in a nonclinical sample: An eLORETA study. *Psychopathology*, 51(1), 16–23. <https://doi.org/10.1159/000485517>
- Fischl, B., & Dale, A. M. (2000). Measuring the thickness of the human cerebral cortex from magnetic resonance images. *Proceedings of the National Academy of Sciences*, 97(20), 11050–11055. <https://doi.org/10.1073/pnas.200033797>
- Fischl, B., Salat, D. H., Busa, E., Albert, M., Dieterich, M., Haselgrove, C., ... Dale, A. M. (2002). Whole brain segmentation: Automated labeling of neuroanatomical structures in the human brain. *Neuron*, 33(3), 341–355.
- Fischl, B., Sereno, M. I., & Dale, A. M. (1999). Cortical surface-based analysis. *NeuroImage*, 9(2), 195–207. <https://doi.org/10.1006/nimg.1998.0396>
- Fleischhaker, C., Schulz, E., Tepper, K., Martin, M., Hennighausen, K., & Remschmidt, H. (2005). Long-term course of adolescent schizophrenia. *Schizophrenia Bulletin*, 31(3), 769–780. <https://doi.org/10.1093/schbul/sbi014>
- Fornito, A., Zalesky, A., & Breakspear, M. (2015). The connectomics of brain disorders. *Nature Reviews Neuroscience*, 16(3), 159–172. <https://doi.org/10.1038/nrn3901>
- Fox, M. D., & Greicius, M. (2010). Clinical applications of resting state functional connectivity. *Frontiers in Systems Neuroscience*, 4, 19. <https://doi.org/10.3389/fnsys.2010.00019>
- Fraguas, D., Gonzalez-Pinto, A., Micó, J. A., Reig, S., Parellada, M., Martínez-Cengotitabengoa, M., ... Arango, C. (2012). Decreased glutathione levels predict loss of brain volume in children and adolescents with first-episode psychosis in a two-year longitudinal study. *Schizophrenia Research*, 137(1–3), 58–65. <https://doi.org/10.1016/j.schres.2012.01.040>
- Fransson, P. (2005). Spontaneous low-frequency BOLD signal fluctuations: An fMRI investigation of the resting-state default mode of brain function hypothesis. *Human Brain Mapping*, 26(1), 15–29. <https://doi.org/10.1002/hbm.20113>
- Fransson, P. (2006). How default is the default mode of brain function?: Further evidence from intrinsic BOLD signal fluctuations. *Neuropsychologia*, 44(14), 2836–2845. <https://doi.org/10.1016/J.NEUROPSYCHOLOGIA.2006.06.017>
- Fusar-Poli, P., Smieskova, R., Kempton, M. J., Ho, B. C., Andreasen, N. C., & Borgwardt, S. (2013). Progressive brain changes in schizophrenia related to antipsychotic treatment? A meta-analysis of longitudinal MRI studies. *Neuroscience & Biobehavioral Reviews*, 37(8), 1680–1691. <https://doi.org/10.1016/J.NEUROPSYCHOLOGIA.2013.06.001>
- Ganella, E. P., Bartholomeusz, C. F., Seguin, C., Whittle, S., Bousman, C., Phassouliotis, C., ... Zalesky, A. (2017). Functional brain networks in treatment-resistant schizophrenia. *Schizophrenia Research*, 184, 73–81. <https://doi.org/10.1016/j.schres.2016.12.008>
- Greenstein, D. K., Wolfe, S., Gochman, P., Rapoport, J. L., & Gogtay, N. (2008). Remission status and cortical thickness in childhood-onset schizophrenia. *Journal of the American Academy of Child and Adolescent Psychiatry*, 47(10), 1133–1140. <https://doi.org/10.1097/CHI.0b013e3181825b0c>
- Hafner, H., Maurer, K., Löffler, W., & Riecher-Rössler, A. (1993). The influence of age and sex on the onset of early course of schizophrenia. *British Journal of Psychiatry*, 162, 80–86. <https://doi.org/10.1192/bjp.162.1.80>
- Hahn, T., Ebner-Priemer, U., & Meyer-Lindenberg, A. (2019). Transparent artificial intelligence – A conceptual framework for evaluating AI-based clinical decision support systems. *SSRN Electronic Journal*, 2018, 3303123. <https://doi.org/10.2139/ssrn.3303123>
- Hare, S. M., Ford, J. M., Mathalon, D. H., Damaraju, E., Bustillo, J., Belger, A., ... Turner, J. A. (2018). Salience-default mode functional network connectivity linked to positive and negative symptoms of schizophrenia. *Schizophrenia Bulletin*, 45, 892–901. <https://doi.org/10.1093/schbul/sby112>
- Hearne, L., Cocchi, L., Zalesky, A., & Mattingley, J. B. (2015). Interactions between default mode and control networks as a function of increasing cognitive reasoning complexity. *Human Brain Mapping*, 36(7), 2719–2731. <https://doi.org/10.1002/hbm.22802>
- Hindriks, R., Adhikari, M. H., Murayama, Y., Ganzetti, M., Mantini, D., Logothetis, N. K., & Deco, G. (2016). Can sliding-window correlations reveal dynamic functional connectivity in resting-state fMRI? *NeuroImage*, 127, 242–256. <https://doi.org/10.1016/j.neuroimage.2015.11.055>
- Hu, M.-L., Zong, X.-F., Mann, J. J., Zheng, J.-J., Liao, Y.-H., Li, Z.-C., ... Tang, J.-S. (2017). A review of the functional and anatomical default mode network in schizophrenia. *Neuroscience Bulletin*, 33(1), 73–84. <https://doi.org/10.1007/s12264-016-0090-1>
- Ingvar, D. H. (1979). “Hyperfrontal” distribution of the cerebral grey matter flow in resting wakefulness; on the functional anatomy of the conscious state. *Acta Neurologica Scandinavica*, 60(1), 12–25.
- Izenman, A. J. (2013). *Modern multivariate statistical techniques*. New York, NY: Springer. https://doi.org/10.1007/978-0-387-78189-1_8
- Jafri, M. J., Pearson, G. D., Stevens, M., & Calhoun, V. D. (2008). A method for functional network connectivity among spatially independent resting-state components in schizophrenia. *NeuroImage*, 39(4), 1666–1681. <https://doi.org/10.1016/j.neuroimage.2007.11.001>
- Janssen, R. J., Mourão-Miranda, J., & Schnack, H. G. (2018). Making individual prognoses in psychiatry using neuroimaging and machine learning. *Biological Psychiatry. Cognitive Neuroscience and Neuroimaging*, 3(9), 798–808. <https://doi.org/10.1016/j.bpsc.2018.04.004>

- Kambeitz-Illankovic, L., Meisenzahl, E. M., Cabral, C., von Saldern, S., Kambeitz, J., Falkai, P., ... Koutsouleris, N. (2016). Prediction of outcome in the psychosis prodrome using neuroanatomical pattern classification. *Schizophrenia Research*, 173(3), 159–165. <https://doi.org/10.1016/j.schres.2015.03.005>
- Khodayari-Rostamabad, A., Hasey, G. M., MacCrimmon, D. J., Reilly, J. P., & de Bruin, H. (2010). A pilot study to determine whether machine learning methodologies using pre-treatment electroencephalography can predict the symptomatic response to clozapine therapy. *Clinical Neurophysiology*, 121(12), 1998–2006. <https://doi.org/10.1016/j.clinph.2010.05.009>
- Kottaram, A., Johnston, L., Ganella, E., Pantelis, C., Kotagiri, R., & Zalesky, A. (2018). Spatio-temporal dynamics of resting-state brain networks improve single-subject prediction of schizophrenia diagnosis. *Human Brain Mapping*, 39, 3663–3681. <https://doi.org/10.1002/hbm.24202>
- Kottaram, A., Johnston, L. A., Cocchi, L., Ganella, E. P., Everall, I., Pantelis, C., ... Zalesky, A. (2019). Brain network dynamics in schizophrenia: Reduced dynamism of the default mode network. *Human Brain Mapping*, 40, 2212–2228. <https://doi.org/10.1002/hbm.24519>
- Koutsouleris, N., Borgwardt, S., Meisenzahl, E. M., Bottlender, R., Möller, H.-J., & Riecher-Rössler, A. (2012). Disease prediction in the at-risk mental state for psychosis using neuroanatomical biomarkers: Results from the FePsy study. *Schizophrenia Bulletin*, 38(6), 1234–1246. <https://doi.org/10.1093/schbul/sbr145>
- Koutsouleris, N., Kambeitz-Illankovic, L., Ruhmann, S., Rosen, M., Ruef, A., Dwyer, D. B., ... Borgwardt, S. (2018). Prediction models of functional outcomes for individuals in the clinical high-risk state for psychosis or with recent-onset depression. *JAMA Psychiatry*, 75(11), 1156–1172. <https://doi.org/10.1001/jamapsychiatry.2018.2165>
- Koutsouleris, N., Meisenzahl, E. M., Davatzikos, C., Bottlender, R., Frodl, T., Scheuerecker, J., ... Gaser, C. (2009). Use of neuroanatomical pattern classification to identify subjects in at-risk mental states of psychosis and predict disease transition. *Archives of General Psychiatry*, 66(7), 700–712. <https://doi.org/10.1001/archgenpsychiatry.2009.62>
- Koutsouleris, N., Riecher-Rössler, A., Meisenzahl, E. M., Smieskova, R., Studerus, E., Kambeitz-Illankovic, L., ... Borgwardt, S. (2015). Detecting the psychosis prodrome across high-risk populations using neuroanatomical biomarkers. *Schizophrenia Bulletin*, 41(2), 471–482. <https://doi.org/10.1093/schbul/sbu078>
- Koutsouleris, N., Wobrock, T., Guse, B., Langguth, B., Landgrebe, M., Eichhammer, P., ... Hasan, A. (2018). Predicting response to repetitive transcranial magnetic stimulation in patients with schizophrenia using structural magnetic resonance imaging: A multisite machine learning analysis. *Schizophrenia Bulletin*, 44(5), 1021–1034. <https://doi.org/10.1093/schbul/sbx114>
- Kraguljac, N. V., White, D. M., Hadley, N., Hadley, J. A., ver Hoef, L., Davis, E., & Lahti, A. C. (2016). Aberrant hippocampal connectivity in unmedicated patients with schizophrenia and effects of antipsychotic medication: A longitudinal resting state functional MRI study. *Schizophrenia Bulletin*, 42(4), 1046–1055. <https://doi.org/10.1093/schbul/sbv228>
- Kucyi, A., & Davis, K. D. (2014). Dynamic functional connectivity of the default mode network tracks daydreaming. *NeuroImage*, 100, 471–480. <https://doi.org/10.1016/j.neuroimage.2014.06.044>
- Kydd, R. R., & Werry, J. S. (1982). Schizophrenia in children under 16 years. *Journal of Autism and Developmental Disorders*, 12(4), 343–357. <https://doi.org/10.1007/BF01538322>
- Laufs, H., Krakow, K., Sterzer, P., Eger, E., Beyerle, A., Salek-Haddadi, A., & Kleinschmidt, A. (2003). Electroencephalographic signatures of attentional and cognitive default modes in spontaneous brain activity fluctuations at rest. *Proceedings of the National Academy of Sciences*, 100(19), 11053–11058. <https://doi.org/10.1073/pnas.1831638100>
- Laumann, T. O., Snyder, A. Z., Mitra, A., Gordon, E. M., Gratton, C., Adeyemo, B., ... Petersen, S. E. (2016). On the stability of BOLD fMRI correlations. *Cerebral Cortex*, 45(28), 100–118. <https://doi.org/10.1093/cercor/bhw265>
- Leucht, S., Davis, J. M., Engel, R. R., Kane, J. M., & Wagenpfeil, S. (2007). Defining “response” in antipsychotic drug trials: Recommendations for the use of scale-derived cutoffs. *Neuropsychopharmacology*, 32(9), 1903–1910. <https://doi.org/10.1038/sj.npp.1301325>
- Lewis, J. A. (2004). In defence of the dichotomy. *Pharmaceutical Statistics*, 3(2), 77–79. <https://doi.org/10.1002/pst.107>
- Lieberman, J., Jody, D., Geisler, S., Alvir, J., Loebel, A., Szymanski, S., ... Borenstein, M. (1993). Time course and biologic correlates of treatment response in first-episode schizophrenia. *Archives of General Psychiatry*, 50(5), 369–376.
- Liegeois, R., Laumann, T. O., Snyder, A. Z., Zhou, H. J., & Yeo, B. T. T. (2017). Interpreting temporal fluctuations in resting-state functional connectivity MRI. *NeuroImage*, 163, 437–455.
- Liemburg, E. J., van der Meer, L., Swart, M., Curcic-Blake, B., Bruggeman, R., Kneegtering, H., & Aleman, A. (2012). Reduced connectivity in the self-processing network of schizophrenia patients with poor insight. *PLoS One*, 7(8), e42707. <https://doi.org/10.1371/journal.pone.0042707>
- Liu, H., Kaneko, Y., Ouyang, X., Li, L., Hao, Y., Chen, E. Y. H., ... Liu, Z. (2012). Schizophrenic patients and their unaffected siblings share increased resting-state connectivity in the task-negative network but not its anticorrelated task-positive network. *Schizophrenia Bulletin*, 38(2), 285–294. <https://doi.org/10.1093/schbul/sbq074>
- Liu, Y., Liang, M., Zhou, Y., He, Y., Hao, Y., Song, M., ... Jiang, T. (2008). Disrupted small-world networks in schizophrenia. *Brain*, 131(4), 945–961. <https://doi.org/10.1093/brain/awn018>
- Martínez-Cengotitabengoa, M., Micó, J. A., Arango, C., Castro-Fornieles, J., Graell, M., Payá, B., ... González-Pinto, A. (2014). Basal low antioxidant capacity correlates with cognitive deficits in early onset psychosis. A 2-year follow-up study. *Schizophrenia Research*, 156(1), 23–29. <https://doi.org/10.1016/j.schres.2014.03.025>
- Maziade, M., Bouchard, S., Gingras, N., Charron, L., Cardinal, A., Roy, M. A., ... Martínez, M. (1996). Long-term stability of diagnosis and symptom dimensions in a systematic sample of patients with onset of schizophrenia in childhood and early adolescence. II: Positive/negative distinction and childhood predictors of adult outcome. *British Journal of Psychiatry*, 169(3), 371–378. <https://doi.org/10.1192/bjp.169.3.371>
- Menezes, N. M., & Milovan, E. (2000). First-episode psychosis: A comparative review of diagnostic evolution and predictive variables in adolescents versus adults. *The Canadian Journal of Psychiatry*, 45(8), 710–716. <https://doi.org/10.1177/070674370004500803>
- Meng, H., Schimmelfmann, B. G., Mohler, B., Lambert, M., Branik, E., Koch, E., ... Bürgin, D. (2006). Pretreatment social functioning predicts 1-year outcome in early onset psychosis. *Acta Psychiatrica Scandinavica*, 114(4), 249–256. <https://doi.org/10.1111/j.1600-0447.2006.00773.x>
- Milev, P., Ho, B. C., Arndt, S., & Andreasen, N. C. (2005). Predictive values of neurocognition and negative symptoms on functional outcome in schizophrenia: A longitudinal first-episode study with 7-year follow-up. *American Journal of Psychiatry*, 162(3), 495–506. <https://doi.org/10.1176/appi.ajp.162.3.495>
- Miller, R. L., Yaesoubi, M., Turner, J. A., Mathalon, D., Preda, A., Pearlson, G., ... Calhoun, V. D. (2016). Higher dimensional meta-state analysis reveals reduced resting fMRI connectivity dynamism in schizophrenia patients. *PLoS One*, 11(3), e0149849. <https://doi.org/10.1371/journal.pone.0149849>
- Molnar-Szakacs, I., & Uddin, L. Q. (2013). Self-processing and the default mode network: Interactions with the mirror neuron system. *Frontiers in Human Neuroscience*, 7, 571. <https://doi.org/10.3389/fnhum.2013.00571>
- Mourao-Miranda, J., Reinders, A. A. T. S., Rocha-Rego, V., Lappin, J., Rondina, J., Morgan, C., ... Dazzan, P. (2012). Individualized prediction of illness course at the first psychotic episode: A support vector machine MRI study. *Psychological Medicine*, 42(5), 1037–1047. <https://doi.org/10.1017/S0033291711002005>

- Nieuwenhuis, M., Schnack, H. G., van Haren, N. E., Lappin, J., Morgan, C., Reinders, A. A., ... Dazzan, P. (2017). Multi-center MRI prediction models: Predicting sex and illness course in first episode psychosis patients. *NeuroImage*, 145(Pt B), 246–253. <https://doi.org/10.1016/j.neuroimage.2016.07.027>
- Parkes, L., Fulcher, B., Yücel, M., & Fornito, A. (2018). An evaluation of the efficacy, reliability, and sensitivity of motion correction strategies for resting-state functional MRI. *NeuroImage*, 171, 415–436. <https://doi.org/10.1016/j.neuroimage.2017.12.073>
- Pompili, M., Amador, X. F., Girardi, P., Harkavy-Friedman, J., Harrow, M., Kaplan, K., ... Tatarelli, R. (2007). Suicide risk in schizophrenia: Learning from the past to change the future. *Annals of General Psychiatry*, 6, 10. <https://doi.org/10.1186/1744-859X-6-10>
- Power, J. D., Barnes, K. A., Snyder, A. Z., Schlaggar, B. L., & Petersen, S. E. (2012). Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. *NeuroImage*, 59(3), 2142–2154. <https://doi.org/10.1016/j.neuroimage.2011.10.018>
- Pruim, R. H. R., Mennes, M., Buitelaar, J. K., & Beckmann, C. F. (2015). Evaluation of ICA-AROMA and alternative strategies for motion artifact removal in resting state fMRI. *NeuroImage*, 112, 278–287. <https://doi.org/10.1016/J.NEUROIMAGE.2015.02.063>
- Pruim, R. H. R., Mennes, M., van Rooij, D., Llera, A., Buitelaar, J. K., & Beckmann, C. F. (2015). ICA-AROMA: A robust ICA-based strategy for removing motion artifacts from fMRI data. *NeuroImage*, 112, 267–277. <https://doi.org/10.1016/j.neuroimage.2015.02.064>
- Puig, O., Penadés, R., Baeza, I., De La Serna, E., Sánchez-Gistau, V., Bernardo, M., & Castro-Fornieles, J. (2014). Cognitive remediation therapy in adolescents with early-onset schizophrenia: A randomized controlled trial. *Journal of the American Academy of Child and Adolescent Psychiatry*, 53(8), 859–868. <https://doi.org/10.1016/j.jaac.2014.05.012>
- Qin, P., & Northoff, G. (2011). How is our self related to midline regions and the default-mode network? *NeuroImage*, 57(3), 1221–1233. <https://doi.org/10.1016/j.neuroimage.2011.05.028>
- Raichle, M. E., MacLeod, A. M., Snyder, A. Z., Powers, W. J., Gusnard, D. A., & Shulman, G. L. (2001). A default mode of brain function. *Proceedings of the National Academy of Sciences of the United States of America*, 98(2), 676–682. <https://doi.org/10.1073/pnas.98.2.676>
- Raichle, M. E., & Snyder, A. Z. (2007). A default mode of brain function: A brief history of an evolving idea. *NeuroImage*, 37(4), 1083–1090. <https://doi.org/10.1016/J.NEUROIMAGE.2007.02.041>
- Ramyead, A., Studerus, E., Kometer, M., Uttinger, M., Gschwandtner, U., Fuhr, P., & Riecher-Rössler, A. (2016). Prediction of psychosis using neural oscillations and machine learning in neuroleptic-naïve at-risk patients. *The World Journal of Biological Psychiatry*, 17(4), 285–295. <https://doi.org/10.3109/15622975.2015.1083614>
- Rashid, B., Arbabshirani, M. R., Damaraju, E., Cetin, M. S., Miller, R., Pearson, G. D., & Calhoun, V. D. (2016). Classification of schizophrenia and bipolar patients using static and dynamic resting-state fMRI brain connectivity. *NeuroImage*, 134, 645–657. <https://doi.org/10.1016/J.NEUROIMAGE.2016.04.051>
- Remschmidt, H., & Theisen, F. (2012). Early-onset schizophrenia1. *Neuropsychobiology*, 66(1), 63–69. <https://doi.org/10.1159/000338548>
- Repovs, G., Csernansky, J. G., & Barch, D. M. (2011). Brain network connectivity in individuals with schizophrenia and their siblings. *Biological Psychiatry*, 69(10), 967–973. <https://doi.org/10.1016/J.BIOPSYCH.2010.11.009>
- Robinson, D., Woerner, M. G., Alvir, J. M., Bilder, R., Goldman, R., Geisler, S., ... Lieberman, J. A. (1999). Predictors of relapse following response from a first episode of schizophrenia or schizoaffective disorder. *Archives of General Psychiatry*, 56(3), 241–247.
- Rotarska-Jagiela, A., van de Ven, V., Oertel-Knöchel, V., Uhlhaas, P. J., Voegeley, K., & Linden, D. E. J. (2010). Resting-state functional network correlates of psychotic symptoms in schizophrenia. *Schizophrenia Research*, 117(1), 21–30. <https://doi.org/10.1016/J.SCHRES.2010.01.001>
- Salvador, R., Martínez, A., Pomarol-Clotet, E., Sarró, S., Suckling, J., & Bullmore, E. (2007). Frequency based mutual information measures between clusters of brain regions in functional magnetic resonance imaging. *NeuroImage*, 35(1), 83–88. <https://doi.org/10.1016/j.neuroimage.2006.12.001>
- Sarpal, D. K., Argyelan, M., Robinson, D. G., Szeszko, P. R., Karlsgodt, K. H., John, M., ... Malhotra, A. K. (2016). Baseline striatal functional connectivity as a predictor of response to antipsychotic drug treatment. *American Journal of Psychiatry*, 173(1), 69–77. <https://doi.org/10.1176/appi.ajp.2015.14121571>
- Sarpal, D. K., Robinson, D. G., Lencz, T., Argyelan, M., Ikuta, T., Karlsgodt, K., ... Malhotra, A. K. (2015). Antipsychotic treatment and functional connectivity of the striatum in first-episode schizophrenia. *JAMA Psychiatry*, 72(1), 5–13. <https://doi.org/10.1001/jamapsychiatry.2014.1734>
- Schennach, R., Riedel, M., Musil, R., & Möller, H. J. (2012). Treatment response in first-episode schizophrenia. *Clinical Psychopharmacology and Neuroscience*, 10, 78–87. <https://doi.org/10.9758/cpn.2012.10.2.78>
- Schnack, H. G., & Kahn, R. S. (2016). Detecting neuroimaging biomarkers for psychiatric disorders: Sample size matters. *Frontiers in Psychiatry*, 7, 50. <https://doi.org/10.3389/fpsy.2016.00050>
- Shim, G., Oh, J. S., Jung, W., Jang, J., Choi, C.-H., Kim, E., ... Kwon, J. (2010). Altered resting-state connectivity in subjects at ultra-high risk for psychosis: An fMRI study. *Behavioral and Brain Functions*, 6(1), 58. <https://doi.org/10.1186/1744-9081-6-58>
- Singh, I., & Rose, N. (2009, July 9). Biomarkers in psychiatry. *Nature*, 460, 202–207. <https://doi.org/10.1038/460202a>
- Skudlarski, P., Jagannathan, K., Anderson, K., Stevens, M. C., Calhoun, V. D., Skudlarska, B. A., & Pearlson, G. (2010). Brain connectivity is not only lower but different in schizophrenia: A combined anatomical and functional approach. *Biological Psychiatry*, 68(1), 61–69. <https://doi.org/10.1016/j.biopsych.2010.03.035>
- Uddin, L. Q., Clare Kelly, A. M., Biswal, B. B., Xavier Castellanos, F., & Milham, M. P. (2009). Functional connectivity of default mode network components: Correlation, anticorrelation, and causality. *Human Brain Mapping*, 30(2), 625–637. <https://doi.org/10.1002/hbm.20531>
- Varoquaux, G. (2018). Cross-validation failure: Small sample sizes lead to large error bars. *NeuroImage*, 180(Pt A), 68–77. <https://doi.org/10.1016/j.neuroimage.2017.06.061>
- Varoquaux, G., Raamana, P. R., Engemann, D. A., Hoyos-Idrobo, A., Schwartz, Y., & Thirion, B. (2017). Assessing and tuning brain decoders: Cross-validation, caveats, and guidelines. *NeuroImage*, 145, 166–179. <https://doi.org/10.1016/J.NEUROIMAGE.2016.10.038>
- Venkataraman, A., Whitford, T. J., Westin, C.-F., Golland, P., & Kubicki, M. (2012). Whole brain resting state functional connectivity abnormalities in schizophrenia. *Schizophrenia Research*, 139(1–3), 7–12. <https://doi.org/10.1016/J.SCHRES.2012.04.021>
- Vita, A., De Peri, L., Deste, G., & Sacchetti, E. (2012). Progressive loss of cortical gray matter in schizophrenia: A meta-analysis and meta-regression of longitudinal MRI studies. *Translational Psychiatry*, 2(11), e190. <https://doi.org/10.1038/tp.2012.116>
- Wechsler, D. (1955). *Manual for the Wechsler adult intelligence scale*. San Antonio, Texas: Psychological Corp.
- Werry, J. S., & McClellan, J. M. (1992). Predicting outcome in child and adolescent (early onset) schizophrenia and bipolar disorder. *Journal of the American Academy of Child and Adolescent Psychiatry*, 31(1), 147–150. <https://doi.org/10.1097/00004583-199201000-00022>
- Werry, J. S., McClellan, J. M., & Chard, L. (1991). Childhood and adolescent schizophrenic, bipolar, and schizoaffective disorders: A clinical and outcome study. *Journal of the American Academy of Child and Adolescent Psychiatry*, 30(3), 457–465. <https://doi.org/10.1097/00004583-199105000-00017>
- Whitfield-Gabrieli, S., Thermenos, H. W., Milanovic, S., Tsuang, M. T., Faraone, S. V., McCarley, R. W., ... Seidman, L. J. (2009). Hyperactivity and hyperconnectivity of the default network in schizophrenia and in

- first-degree relatives of persons with schizophrenia. *Proceedings of the National Academy of Sciences*, 106(4), 1279–1284. <https://doi.org/10.1073/pnas.0809141106>
- Zalesky, A., Fornito, A., & Bullmore, E. (2012). On the use of correlation as a measure of network connectivity. *NeuroImage*, 60(4), 2096–2106. <https://doi.org/10.1016/j.neuroimage.2012.02.001>
- Zalesky, A., Fornito, A., & Bullmore, E. T. (2010). Network-based statistic: Identifying differences in brain networks. *NeuroImage*, 53(4), 1197–1207. <https://doi.org/10.1016/j.neuroimage.2010.06.041>
- Zhou, Y., Liang, M., Tian, L., Wang, K., Hao, Y., Liu, H., ... Jiang, T. (2007). Functional disintegration in paranoid schizophrenia using resting-state fMRI. *Schizophrenia Research*, 97(1–3), 194–205. <https://doi.org/10.1016/j.schres.2007.05.029>

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Kottaram A, Johnston LA, Tian Y, et al. Predicting individual improvement in schizophrenia symptom severity at 1-year follow-up: Comparison of connectomic, structural, and clinical predictors. *Hum Brain Mapp*. 2020;41: 3342–3357. <https://doi.org/10.1002/hbm.25020>