Functional Coupling and Longitudinal Outcome Prediction in First-Episode Psychosis

Isaac Z. Pope[a,b*], Sidhant Chopra[c,d], Alexander Holmes[e], Shona M. Francey[c,d], Brian O'Donoghue[d,f,g], Vanessa L. Cropley[h], Barnaby Nelson[c,d], Hok Pan Yuen[c,d], Kelly Allott[c,d], Mario Alvarez-Jimenez[c,d], Susy Harrigan[c,d,i], Christos Pantelis[h,j], Stephen J. Wood[c,d,k], Patrick D. McGorry[c,d], Alex Fornito[a,b]

[a]Turner Institute for Brain and Mental Health, School of Psychological Science, Monash University, Clayton, Australia

[b]Monash Biomedical Imaging, Monash University, Clayton, Australia

[c]Orygen, Parkville, Victoria, Australia

[d]Centre for Youth Mental Health, The University of Melbourne, Melbourne, Victoria, Australia

[e]Wellcome Centre for Integrative Neuroimaging, University of Oxford, Oxford, United Kingdom

[f]St. Vincent's University Hospital, Dublin 4, Ireland

[g]School of Medicine, University College Dublin, Dublin 4, Ireland

[h]Melbourne Neuropsychiatry Centre, Department of Psychiatry, University of Melbourne, Melbourne, Victoria, Australia

[i]Centre for Mental Health, Melbourne School of Global and Population Health, The University of Melbourne, Parkville, Victoria, Australia

[j]Western Hospital Sunshine, St. Albans, Victoria, Australia

[k]School of Psychology, University of Birmingham, Edgbaston, United Kingdom

*Corresponding author. Email: isaac.pope@monash.edu

Abstract

Background Clinical outcomes following a first episode of psychosis (FEP) are highly heterogeneous between patients. The identification of prognostic biomarkers would greatly facilitate personalized treatments. Psychosis patients often display brain-wide disruptions of inter-regional functional coupling (FC), with some being linked to symptom severity and remission. FC may thus hold prognostic potential for people experiencing psychosis.

Methods Ninety antipsychotic-naïve FEP patients (51% female, 15-25 years) were randomized to receive either antipsychotic or placebo tablets for 6 months alongside psychosocial interventions. A subset of these patients also completed functional magnetic resonance imaging (55 with usable data at baseline and 37 at 3 months), which was used to evaluate whether baseline FC, or 3-month change in FC, could predict 6- and 12-month changes in symptoms and functioning, quantified using the Brief Psychiatric Rating Scale and the Social and Occupational Functioning Assessment Scale, respectively. We considered three different cross-validated prediction algorithms: (i) connectome-based predictive modelling; (ii) kernel ridge regression; and (iii) multilayer meta-matching.

Results All algorithms showed poor performance in predicting patients' 6- and 12-month changes in symptoms and functioning (all $r_{mean} < 0.3$), and no model achieved significance via permutation testing (all $p > 0.05$).

Conclusions Our findings suggest that brain-wide measures of FC may not be suitable for predicting extended clinical outcomes over a 6- to 12-month period in FEP patients.

Introduction

Clinical outcomes following a first episode of psychosis (FEP) are heterogenous in both symptoms and functioning, ranging from permanent recovery to chronic and severe illness (1–4). While early intervention with pharmacological and psychosocial treatments has been shown to improve clinical outcomes in FEP patients (5–10), the efficacy and tolerability of these treatments vary considerably between individuals (11,12). Moreover, some FEP patients may recover through psychosocial treatment alone (13). Current clinical guidelines for FEP address this heterogeneity via trial-and-error in prescription (14,15), potentially delaying remission and recovery (16). The identification of reliable predictors of patients' clinical outcomes is therefore a necessary step towards the development of personalized treatment strategies.

Across the psychosis spectrum, magnetic resonance imaging (MRI) studies have revealed disruptions in both the structural connectivity between different brain regions (17–19), and the patterns of inter-regional functional coupling (FC; i.e., correlated activity) that these connections support (18,20–28). Some of these FC disruptions are apparent at the onset of psychosis (20–22,27,28) and in high risk individuals (29) and are related to different symptoms (29–31). Most importantly, prospective studies have found associations between patients' baseline FC patterns and their changes in symptoms and functioning following treatment (28,32–36). Despite the promise of this work, the analyses have generally relied on within-sample quantification of associations between FC and outcomes, which may inflate effect size estimates due to over-fitting (37). A thorough assessment of the generalizability of any putative prognostic biomarker therefore requires analytic strategies such as cross-validation (37).

Some studies have reported cross-validated evidence that baseline FC can predict longitudinal symptom changes in FEP patients (38–44), but most only involved short-term outcomes (i.e., up to 16 weeks post-baseline) in patients diagnosed with schizophrenia, potentially impeding the generalizability of their findings to the kinds of transdiagnostic FEP cohorts encountered in real-world clinical care (45). Moreover, a focus on predicting symptom changes may not translate to an impact on functional outcomes, since the two are not always related (46,47). Functional recovery in domains such as social relationships, educational attainment, and occupational functioning is particularly important in this regard, as it may better address patient needs compared to symptom recovery (48). Critically, all prior studies only use a single prediction algorithm, rather than comparing the performance of multiple gold-standard approaches, limiting our ability to establish reliable and reproducible prediction methods for FEP.

In this study, we apply multiple cross-validated prediction algorithms to task-free resting-state functional MRI (fMRI) and clinical outcome data from the Staged Treatment and Acceptability Guidelines in Early Psychosis Study (STAGES), a triple-blind randomized control trial of antipsychotics in previously antipsychotic-naïve FEP patients (13,49). Our aim was to assess whether patients' baseline resting-state FC could predict their changes in symptoms and functioning after 6 and 12 months. We secondarily evaluated whether longitudinal changes in FC between baseline and 3 months (ΔFC) could predict the same outcomes.

Methods and Materials

Design and Participants

This study used clinical and neuroimaging data from STAGES (13,49), which was approved by the Melbourne Health Human Research and Ethics Committee and registered under the Australian New Zealand Clinical Trials Registry (ACTRN12607000608460) in November 2007. Ninety FEP patients (ages 15-25 years) were recruited between 2008 and 2016 at the Early Psychosis Prevention and Intervention Centre at Orygen Youth Health in Melbourne, Australia. Eligibility requirements comprised: (i) meeting criteria for a psychotic disorder in the DSM-IV via a Structured Clinical Interview; (ii) a duration of untreated psychosis under 6 months; (iii) ability to provide informed consent; (iv) low risk of suicidality, self-harm, or hostility; (v) negligible lifetime use of antipsychotics, or present use of mood stabilizers; (vi) stable housing and support; and (vii) no pregnancy. Patients were randomized to receive antipsychotic medication (risperidone or paliperidone) or matched placebo tablets for 6 months, along with psychosocial intervention in the form of cognitive behavioral case management (50). Further details about the trial design, including safety protocols, can be found in the supplement and previous publications (13,49).

Patients' functioning and symptoms were measured at baseline and after 6 and 12 months via the Social and Occupational Functioning Assessment Scale (SOFAS) and the Brief Psychiatric Rating Scale (BPRS), respectively (Figure 1). Patients' total scores on these inventories were preregistered as the primary and secondary outcomes of the larger clinical trial, respectively. Since we have previously demonstrated no differences in these outcomes between the placebo and medicated patients at 6 or 12 months, both treatment groups were combined in predictive models to maximize statistical power and generalizability. Since dichotomizing of patients into 'responders' and 'non-responders' is statistically inefficient in ignoring within-group variability (51), and is not supported by a consensus definition of clinical recovery (3), we defined patients' clinical outcomes as continuous proportional change scores $(y_2 - y_1)/y_1$, where y_1 and y_2 denote their SOFAS or BPRS score at baseline and at a later timepoint (6 or 12 months), respectively. Despite attrition throughout the trial, most patients who discontinued the trial still chose to complete clinical assessments. We included their data in our analysis to maximize statistical power and generalizability.

MRI Acquisition and Processing

A 3T Siemens Trio Tim scanner with a 32-channel head coil was used to acquire T1-weighted anatomical MRI and resting-state fMRI data at the Royal Children's Hospital in Melbourne, Australia. Of the 90 patients in the clinical trial, 58 and 39 were scanned at baseline and after 3 months, respectively. We used the software MRI-QC to automatically compute image quality metrics for each scan (52), then excluded three scans from further analysis due to localized signal dropout or high noise. We also removed one scan due to excessive head motion, classified using standard exclusion criteria (53). The remaining scans were preprocessed via fMRIPrep v1.4.1 (54), then denoised using a pipeline previously shown to mitigate the impacts of noise sources such as head motion and non-neuronal physiological fluctuations (53). For each voxel in a scan, this involved linear detrending, regressing out motion-related signals automatically identified by the software ICA-AROMA (55), regressing out mean signals of grey matter, white matter, and cerebrospinal fluid tissues, and band-pass temporal filtering at 0.008-0.08 Hz (Figure S1). Each brain was

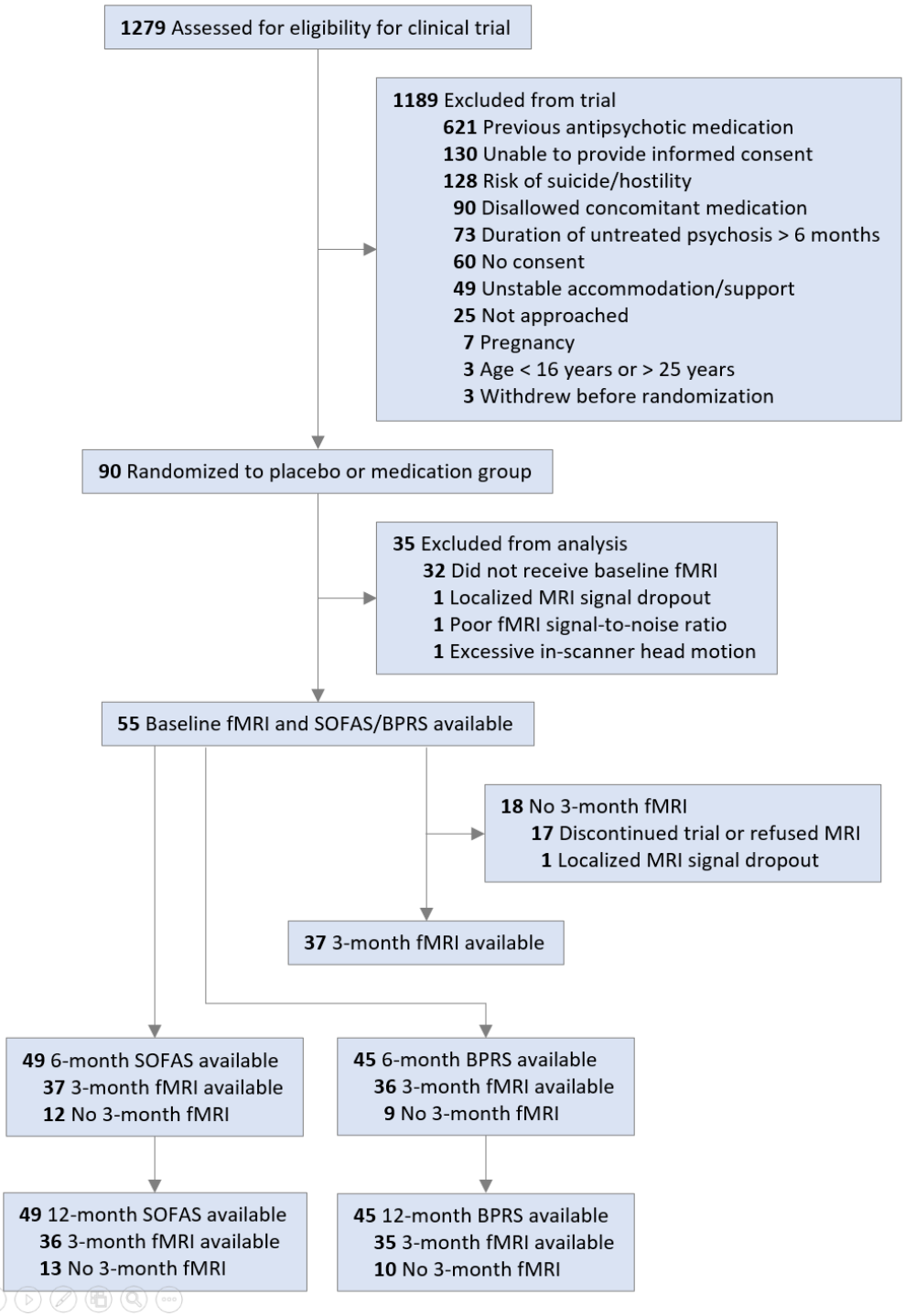


Figure 1: Flow diagram for patients involved in the larger clinical trial and in the present analysis. MRI, magnetic resonance imaging; fMRI, functional MRI; SOFAS, Social and Occupational Functioning Assessment Scale; BPRS, Brief Psychiatric Rating Scale.

parcellated using the 300-region Schaefer cortical atlas (56) and the Melbourne Scale II subcortical atlas (57), producing a mean fMRI signal time series for each region. We then computed Pearson's correlation for every pair of these time series to generate a 328×328 FC matrix for each scan.

Since one prediction algorithm that we used (multilayer meta-matching) has been pre-trained on FC from larger datasets, its use requires that FC input shares the same parcellation as its training data (58). Accordingly, we combined the 400-

region Schaefer cortical atlas with the automated subcortical segmentation in FreeSurfer (59,60) to separately produce the required 419×419 FC matrices.

For ΔFC-based predictions, matrices were instead derived by subtracting each patient's baseline FC matrix from their 3-month FC matrix. This means that each entry in a patient's ΔFC matrix contains their longitudinal change in FC over the first 3 months, for some coupling of brain regions. Further details on acquisition parameters, quality control, and preprocessing are provided in the Supplement.

Predictive Modelling

To evaluate the robustness of our findings, we predicted patient outcomes using three different algorithms: (i) connectome-based predictive modelling (CPM); (ii) kernel ridge regression (KRR); and (iii) multilayer meta-matching. The first two are commonly used in the field. The last has been recently shown to improve the performance of predictive models in small samples (58). All algorithms performed 4-fold cross-validation, where data are first divided into four equally-sized folds. Three folds are used for training a model while the remaining fold is held aside for testing this model via prediction. This procedure is repeated four times, with each fold used for testing once. The following presents details of each prediction algorithm.

Connectome-based Predictive Modelling

CPM is a simple approach that prioritizes direct interpretability of brain regions and networks implicated in accurate predictions (61). We implemented CPM via 16 models, which differed in terms of the: (i) outcome measure (symptoms or functioning); (ii) outcome timepoint (6- or 12-month change); (iii) predictor (baseline FC or ΔFC); and (iv) feature selection (positive or negative). In each model, CPM first performs feature selection in the training set, using Pearson's correlation to identify FC estimates that are either positively ($r > 0, p < 0.05$ uncorrected) or negatively ($r < 0, p < 0.05$ uncorrected) associated with outcomes. For each patient, FC values are then summed across these positive or negative features to create a summary score. A linear model is fitted between these FC summary scores and outcomes within the training set, then used to predict the outcomes of patients in the testing set.

Kernel Ridge Regression

KRR is a classical machine learning method that lacks a feature selection step, instead performing multivariable prediction with regularization, where a hyperparameter is optimized to control the trade-off between training error and testing error (62). We used the resting-state FC-based KRR implementation from Li et al. (2019) (63), which has been shown to predict many behavioral measures with accuracy comparable to more computationally demanding deep neural networks (64). KRR was implemented via eight models, which differed in terms of the: (i) outcome measure (symptoms or functioning); (ii) outcome timepoint (6- or 12-month change); and (iii) predictor (baseline FC or ΔFC). In KRR, each outcome in the testing set is predicted as a weighted mean of the outcomes observed in the training set. These weights include measures of similarity in FC between patients and an l_2-regularization hyperparameter whose value is optimized via an inner loop of 4-fold cross-validation within the training set (further details are provided in the Supplement).

Multilayer Meta-matching

Recent work suggests that prediction models that perform well within a small sample often fail to generalize across datasets (65–69). To boost our prediction accuracies while avoiding overoptimism, we additionally used multilayer meta-matching, a transfer learning framework that transposes prediction models trained on large healthy datasets to a smaller dataset of interest. This approach exploits neural degeneracy, whereby a relatively small set of resting-state FC patterns underlie many phenotypes spanning cognition, demographics, and mental health (70–72). We used the pre-trained model described in Chen et al. (2024), which significantly outperformed classical KRR in samples as small as 10 individuals (58) and significantly predicted cognitive measures within and across multiple psychiatric cohorts (73).

Briefly, this involved using a deep neural network and linear ridge regression to predict 67 different phenotypes in the UK Biobank dataset ($n = 36834$) (74) based on resting-state FC. These techniques were then used to predict another 162 phenotypes in other datasets (Adolescent Brain Cognitive Development, $n = 5985$ (75); Genomics Superstruct Project, $n = 862$ (76); Healthy Brain Network, $n = 930$ (77); Enhanced Nathan Kline Institute Rockland Sample, $n = 896$ (78)). We used this pre-trained model to predict all phenotypes for each STAGES patient, thereby generating 458 proxy variables that replaced FC as the inputs for a final KRR step to predict the clinical outcomes. Since no models have yet been pre-trained using longitudinal changes in FC, we implemented multilayer meta-matching via four models, which differed in terms of: (i) outcome measure (symptoms or functioning); and (ii) outcome timepoint (6- or 12-month change).

Evaluating Prediction Performance and Significance

To further reduce sampling bias, we ran each of the above models across 100 splits, where each split used a different random allocation of patients to the four folds. We then quantified each model's prediction performance as the mean correlation between predicted and observed outcomes across all 100 splits (r_{mean}). We assessed prediction significance via permutation testing, which generates empirical null distributions of r_{mean}. For each model, outcomes were randomly shuffled amongst patients 1000 times, then prediction algorithms were rerun for each of these permutations. 100 splits per permutation were used for CPM null models, but to reduce computational burden we used 50 splits for KRR and 20 splits for multilayer meta-matching. We calculated p-values as the proportion of null r_{mean} values that exceeded the true r_{mean}.

To correct for family-wise error (FWE) arising from the inclusion of multiple predictive models, the inference-based Westfall-Young method was chosen (79). For this, the highest r_{mean} across all 16 null CPM models was retained at each of the 1000 permutations, producing a single FWE-corrected null distribution for the calculation of all 16 p_{FWE}-values. FWE correction was similarly applied across all eight KRR models and across all four multilayer meta-matching models. This ensures that models only achieve significance by surpassing the strongest prediction performances across all null models sharing the same prediction algorithm. We chose this approach to reduce false positives in a more tailored manner compared to alternative methods such as Bonferroni correction, which may have inflated the rate of false negatives (80,81). Statistical significance was assessed at $p_{FWE} < 0.05$.

Results

Clinical Outcomes

Across all patients included in our analysis, 73% had increased functioning (SOFAS) after the 6 months of STAGES treatment (mean proportional change $= 20\%$, $n = 49$; Figure S2) and 69% had increased after 12 months (mean proportional change $= 22\%$, $n = 49$). Similarly, 84% of patients had decreased symptoms (BPRS) after 6 months (mean proportional change $= -25\%$, $n = 45$) and 84% had decreased after 12 months (mean proportional change $= -23\%$, $n = 45$). Patients' baseline SOFAS scores (52.3 [12.1]; mean [SD]; Table S1) were comparable to those of a large epidemiologically representative FEP cohort (82), and baseline BPRS total scores (57.6 [9.4]) indicated marked illness (83). Patients who discontinued prior to 3 months did not receive a second MRI scan, meaning predictive models based on ΔFC used subsamples of 35-37 patients. Clinical outcomes did not significantly differ between patients with 3-month FC data and those without (two-tailed t-test, $p > 0.05$ for all 6- and 12-month SOFAS and BPRS data).

Predictive Modelling

Connectome-based Predictive Modelling

The eight CPM models that used baseline FC to predict outcomes showed limited predictive value (r_{mean} range of -0.27 to 0.17; Table 1 and Figure 2). This was also the case for the eight ΔFC models (r_{mean} range of -0.28 to 0.18). Permutation testing via empirical null distributions revealed that none of the CPM models passed the threshold for statistical significance before or after FWE correction (all $p \geq 0.07$; Table 1 and Figure S3). We found similar results when using alternative feature selection thresholds of $p < 0.05$ and $p < 0.001$ in CPM (all $r_{mean} < 0.3$; Table S2-S3, Figures S6 and S8) or significance (all $p \geq 0.06$; Table S2-S3, Figures S7 and S9).

Kernel Ridge Regression

The four KRR models that used baseline FC to predict outcomes also showed limited predictive value (r_{mean} range of -0.24 to 0.23; Table 2 and Figure 3). This was also the case for the four ΔFC models (r_{mean} range of -0.28 to 0.21). Permutation testing via empirical null distributions revealed that none of the KRR models passed the threshold for statistical significance before or after FWE correction (all $p \geq 0.12$; Table 2 and Figure S4).

Multilayer Meta-matching

The four multilayer meta-matching models that used baseline FC to predict outcomes also demonstrated limited performance (r_{mean} range of -0.16 to 0.24; Table 3 and Figure 4) and none achieved significance before or after FWE correction (all $p \geq 0.09$; Table 3 and Figure S5).

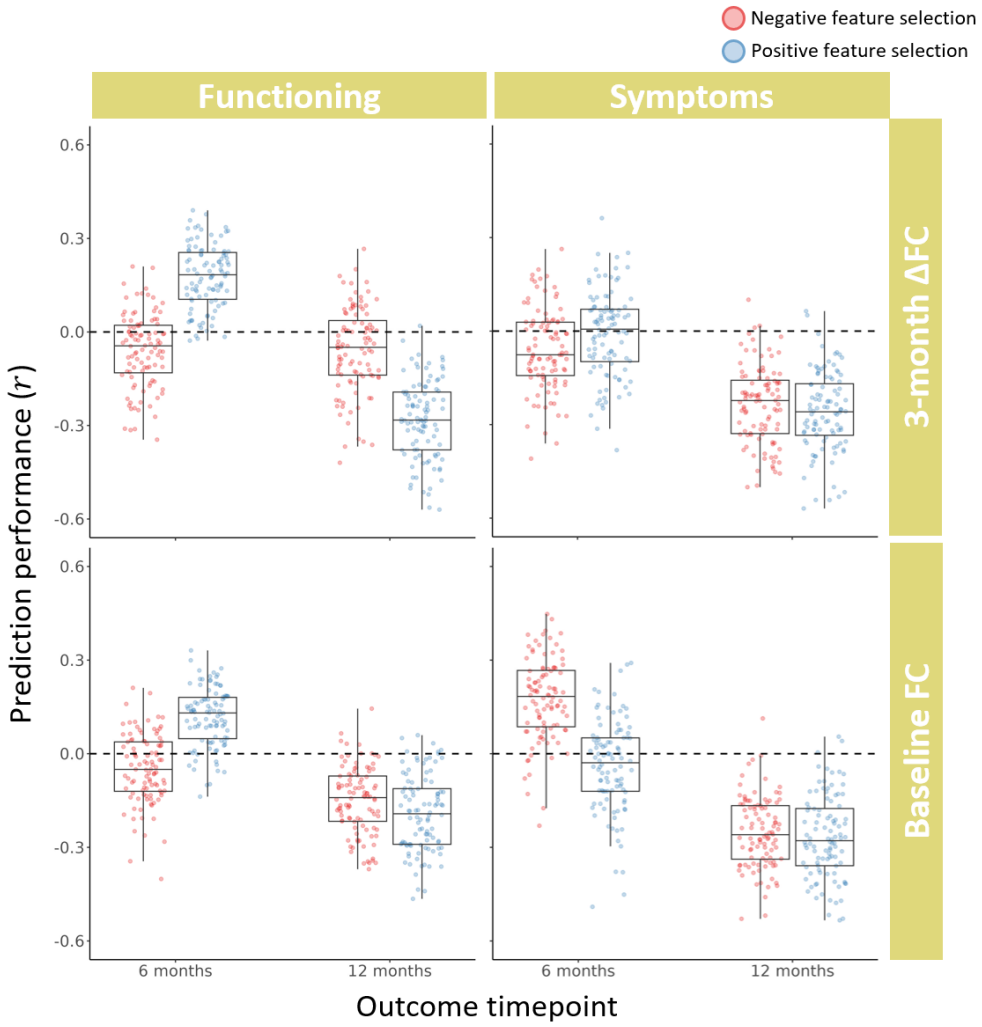


Figure 2: Performance of connectome-based predictive modelling (CPM) for predicting patients' 6- and 12-month changes in symptoms and functioning, using either baseline functional coupling (FC) or 3-month change in FC. Each data point shows the strength of Pearson's correlation between predicted and observed clinical outcomes for a single split of 4-fold cross-validation, with each of the 16 models comprising 100 random splits.

Table 1: Prediction performance and significance for all 16 connectome-based predictive modelling (CPM) models. pos, positive feature model; neg, negative feature model.

		Baseline FC			**3-month ΔFC**		
Functioning		r_{mean}	p	p_{FWE}	r_{mean}	p	p_{FWE}
6-month	*pos*	0.12	0.13	0.84	0.18	0.07	0.62
	neg	-0.05	0.47	1	-0.06	0.38	1
12-month	*pos*	-0.19	0.77	1	-0.28	0.89	1
	neg	-0.15	0.65	1	-0.06	0.45	1
Symptoms		r_{mean}	p	p_{FWE}	r_{mean}	p	p_{FWE}
6-month	*pos*	-0.04	0.40	1	-0.01	0.31	1
	neg	0.17	0.07	0.64	-0.06	0.38	1
12-month	*pos*	-0.27	0.87	1	-0.26	0.81	1
	neg	-0.25	0.86	1	-0.24	0.78	1

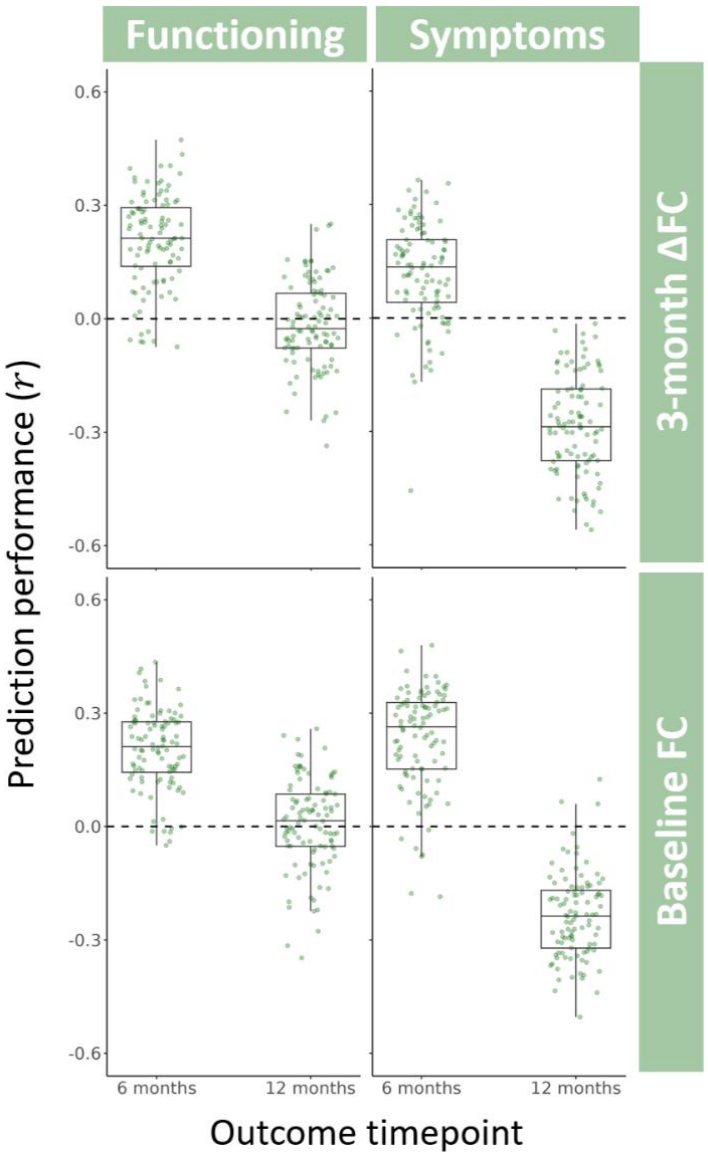


Figure 3: Performance of kernel ridge regression (KRR) for predicting patients' 6- and 12-month changes in symptoms and functioning, using either baseline functional coupling (FC) or 3-month change in FC. Each data point shows the strength of Pearson's correlation between predicted and observed clinical outcomes for a single split of 4-fold cross-validation, with each of the eight models comprising 100 random splits.

Table 2: Prediction performance and significance for all eight kernel ridge regression (KRR) models.

	Baseline FC			**3-month ΔFC**		
Functioning	r_{mean}	p	p_{FWE}	r_{mean}	p	p_{FWE}
6-month	0.21	0.12	0.68	0.21	0.17	0.67
12-month	-0.01	0.50	1	-0.02	0.52	1
Symptoms	r_{mean}	p	p_{FWE}	r_{mean}	p	p_{FWE}
6-month	0.23	0.12	0.59	0.12	0.32	0.91
12-month	-0.24	0.90	1	-0.28	0.92	1

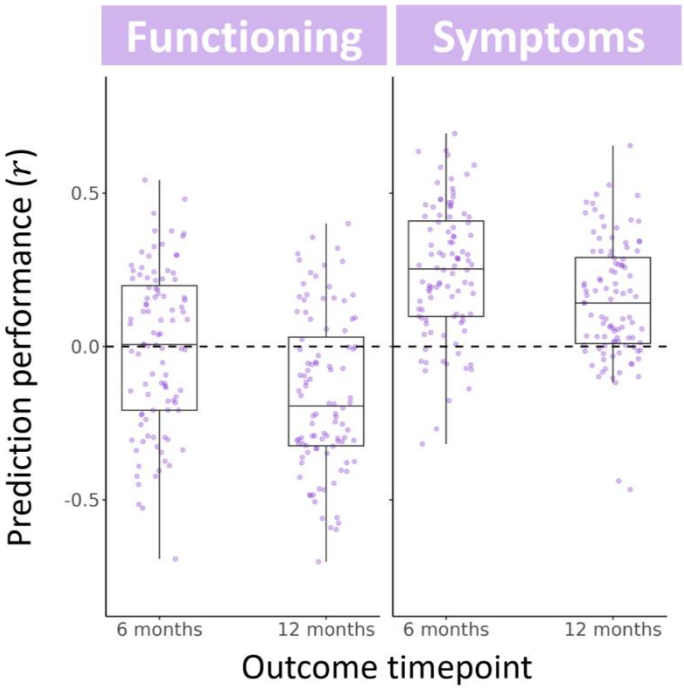


Figure 4: Performance of multilayer meta-matching for predicting patients' 6- and 12-month changes in symptoms and functioning, using baseline functional coupling (FC). Each data point shows the strength of Pearson's correlation between predicted and observed clinical outcomes for a single split of 4-fold cross-validation, with each of the four models comprising 100 random splits.

Table 3: Prediction performance and significance for all four multilayer meta-matching models.

	Baseline FC		
Functioning	r_{mean}	p	p_{FWE}
6-month	-0.01	0.53	0.94
12-month	-0.16	0.83	1
Symptoms	r_{mean}	p	p_{FWE}
6-month	0.24	0.09	0.29
12-month	0.15	0.21	0.58

Discussion

Many fMRI studies have reported disrupted FC across the psychosis spectrum by comparing patients to healthy controls (18,20–28), yet it remains unclear whether FC holds potential as a prognostic biomarker that can predict treatment outcomes in FEP. We leveraged the long-term longitudinal measurement of symptoms and functioning from a clinical trial, along with gold-standard prediction algorithms, to evaluate whether resting-state FC could predict FEP patients' outcomes at the end of the 6-month trial and after 12 months.

All the CPM, KRR, and multilayer meta-matching algorithms showed poor prediction performance (all $r_{mean} < 0.3$) for both outcome measures and timepoints, and no model achieved statistical significance. Our results contrast prior FEP prediction studies, where FC either predicted treatment response compared to non-response (defined via a priori symptom thresholds) with balanced accuracies exceeding 75% (39,40,42,43), or significantly predicted longitudinal changes in symptoms (38,41,42,44). Several differences between the prior and current work may explain this difference. First, our FEP patient sample included the full spectrum of psychotic disorders, whereas most prior studies only included patients diagnosed with schizophrenia. Second, patients in our sample received psychosocial interventions alongside either placebo or antipsychotic tablets, whereas prior studies generally involved more rigid treatment protocols (e.g., all patients received the same antipsychotic dosing with no psychosocial interventions). Third, we considered whole-brain FC, whereas most prior studies defined a priori predictors localized to specific regions or networks. Predictions were consequently driven by different FC patterns across studies, including cerebello-thalamo-cortical (41), cortico-cortical (38,42), and hippocampal-cortical (39) circuits as well as FC seeded in the bilateral anterior cingulate cortex (44) and striatum (40).

When taken with our null findings, this body of work suggests that FC may only predict clinical outcomes within highly selected samples and contexts, and may not be generalizable to the heterogeneous samples and treatment settings encountered in real-world clinics (84,85). It is also possible that FC may only capture a patient's capacity for clinical changes within the early phase of treatment, given that the only other study to predict outcomes beyond 16 weeks also reported the lowest performance ($r = 0.21$) amongst prior studies using continuous outcomes (41).

A separate study of this cohort used CPM with baseline measures of structural connectivity to significantly predict 12-month changes in functioning ($r_{mean} = 0.44$), significantly outperforming predictive models based on patients' baseline clinical, cognitive, and demographic data (19). This result suggests that diffusion MRI-derived estimates of structural connectivity hold greater potential for a prognostic biomarker in FEP than resting-state FC, which has a lower test-retest reliability (86) and correlates with a range of in-scanner internal states (87–89) and self-reported experiences (90,91). However, prognostic trait-like FC patterns may be revealed by estimating FC during movie-watching (92) or across multiple task and rest conditions (93–96) [see Cao et al. (2023) for FEP outcome prediction based on such cross-paradigm FC (38)].

Limitations

Although our sample size of $n = 35\text{-}49$ was similar to those featured in prior studies, it may be underpowered for robust predictive modelling (65–69). We included multilayer meta-matching to compensate for the small sample size, but this approach only improves prediction performance compared to single-sample methods when the phenotype of interest is strongly correlated with at least one phenotype from the large source datasets used in training (58,97). To our knowledge, no prior studies have used multilayer meta-matching for longitudinal prediction, and so patients' clinical outcomes may not have been closely related to any of the 229 phenotypes in the five cross-sectional source datasets. The reliability and generalizability of our finding, that resting-state FC cannot predict FEP patients' clinical outcomes, can therefore only be assessed by performing cross-validation on a larger sample (67,84,85). This remains challenging, however, due to the paucity of open-access FEP neuroimaging datasets with longitudinal outcomes.

The BPRS, which we used to measure symptom severity, demonstrates lower inter-rater reliability, internal consistency, and clinical predictive power than other inventories which require more burdensome administration (98,99). This may have weakened our statistical power for detecting FC-outcome associations and explain the poor prediction performance for symptom changes compared to studies that used other scales (41–43). The construct validity of our symptom outcomes may have been limited by summing scores across the 18 items in the BPRS, as some psychometric literature suggests that sum-scores can obfuscate separate dimensions of psychopathology in highly heterogenous disorders (100–102). In the present study, we chose our predicted outcomes to align with the preregistered clinical trial outcomes of STAGES and minimize FWE correction. To improve clinical relevance, future studies could instead predict changes within subscales demarcated to capture more homogenous symptom dimensions (e.g., depression, anxiety, psychotic symptoms) (103,104), or use factor analytic techniques to maximize effect sizes (101,105).

Heterogeneity in clinical outcomes amongst FEP patients can also be explained by non-neurobiological factors, including baseline symptoms and functioning (4,5,10,85,106–113), age (4,107,114,115), sex (4,10,110,112), schizotypal traits (10,107), inflammatory markers (116,117), education and employment status (109,110,115,118), substance use (109,112,119,120), recent life events (121), previous depressive episodes (110), social environment (110,122), and duration of untreated psychosis (4–6,106–108,113,123–127). If these factors do not covary with neuroimaging measures, their effects will impose an upper limit to the performance and clinical utility of brain-based prognostic biomarkers. It will be important for future work to evaluate how the combination of different biological, demographic, and clinical measures affects prediction of outcomes (128–130). Indeed, given the expense and complexity of fMRI, any promising FC-based prognostic biomarkers should demonstrate predictive capacity beyond that afforded by simpler measures (85,108,131,132).

Conclusions

Our analysis, using multiple cross-validated prediction algorithms, indicates that brain-wide resting-state FC at baseline, or FC change over 3 months, does not significantly predict FEP patients' changes in symptoms or functioning over a 6- to 12-month period after commencing treatment. When taken together with past work, this finding suggests that FC may only hold prognostic utility within narrow clinical and experimental contexts.

Acknowledgements

Janssen-Cilag partially supported the early years of this study with an unrestricted investigator-initiated grant and provided risperidone, paliperidone, and matched placebo for the first 30 participants. The study was then funded by an Australian National Health and Medical Research Project (NHMRC) grant (Grant No. 1064704 [to SMF, BO, BN, HPY, KA, MA-J, SH, SJW, PM, AF]). The funders had no role in study design, data collection, data analysis, data interpretation, writing, approval, or submission of this manuscript. No patients or members of the public were involved in the design, conduct, reporting, interpretation, or dissemination of the study.

This study has been supported by a large number of clinical staff at Orygen Youth Health (Craig Macneil, Kingsley Crisp, Dylan Alexander, Tina Proffitt, Rachel Tindall, Jennifer Hall, Lisa Rumney, Franco Scalzo, Melissa Pane, Linda Kader, Frank Hughes, Clare Shelton, Ryan Kaplan, David Hallford, Bridget Moller, and Rick Fraser) and research assistants (Daniela Cagliarini, Suzanne Wiltink, Janine Ward, Sumudu Mallawaarachichi, Lara Baldwin, and Jessica Graham). The trial took place at the Early Psychosis Prevention and Intervention Centre, which is part of Orygen Youth Health, Melbourne, Australia.

SC is supported by a Graduate Education Fund Scholarship from the American Australian Association. SF, MA-J, and AF reported receiving grants from the NHMRC and Australian Research Council during the conduct of the study. CP reported receiving grants from the Australian NHMRC and from the Lundbeck Foundation and personal fees from Lundbeck Australia Pty Ltd. Advisory Board for talks presented at educational meetings organized by Lundbeck. PM reported receiving grants from the Australian NHMRC, the Colonial Foundation, the National Alliance for Research on Schizophrenia and Depression, the Stanley Foundation, the National Institutes of Health, Wellcome Trust, the Australian and Victorian governments, and Janssen-Cilag (unrestricted investigator-initiated grant) during the conduct of the study; past unrestricted grant funding from Janssen-Cilag, AstraZeneca, Eli Lilly, Novartis, and Pfizer; and honoraria for consultancy and teaching from Janssen-Cilag, Eli Lilly, Pfizer, AstraZeneca, Roche, Bristol Myers Squibb, and Lundbeck. BN was supported by an NHMRC Senior Research Fellowship (Fellowship No. 1137687) and VLC was supported by an NHMRC EL2 Fellowship (Fellowship No. 1177370). BN, KA, and VLC were supported by a University of Melbourne Dame Kate Campbell Fellowship.

This work was supported by the MASSIVE HPC facility (www.massive.org.au) (133). All code used for statistical analyses and generating main results figures can be found at https://github.com/izachp/psychosis-FC-prediction.

Disclosures

Dr Francey reported receiving grants from the Australian National Health & Medical Research Council (NHMRC) during the conduct of the study. Dr Allott reported receiving grants from the Australian National Health and Medical Research Council during the conduct of the study. Dr Pantelis reported receiving grants from the Australian NHMRC and from the Lundbeck Foundation and personal fees from Lundbeck Australia Pty Ltd Advisory Board for talks presented at educational meetings organized by Lundbeck. Dr McGorry reported receiving grants from the Australian NHMRC, the Colonial Foundation, the National Alliance for Research on Schizophrenia and Depression, the Stanley Foundation, the National Institutes of Health, Wellcome Trust, the Australian and Victorian governments, and Janssen-Cilag (unrestricted investigator-initiated grant) during the conduct of the study; past unrestricted grant funding from Janssen-Cilag, AstraZeneca, Eli Lilly, Novartis, and Pfizer; honoraria for consultancy and teaching from Janssen-Cilag, Eli Lilly, Pfizer, AstraZeneca, Roche, Bristol Myers Squibb, and Lundbeck. Dr Fornito received grants from the National Health and Medical Research Council, Australian Research Council, and Sylvia and Charles Viertel Foundation during this project. No other disclosures were reported.

References

1. Edwards J, Maude D, McGorry PD, Harrigan SM, Cocks JT. Prolonged recovery in first-episode psychosis. The British Journal of Psychiatry. 1998 Jun;172(S33):107–16.

2. Lally J, Ajnakina O, Stubbs B, Cullinane M, Murphy KC, Gaughran F, et al. Remission and recovery from first-episode psychosis in adults: systematic review and meta-analysis of long-term outcome studies. The British Journal of Psychiatry. 2017 Dec;211(6):350–8.

3. Åsbø G, Ueland T, Haatveit B, Bjella T, Flaaten CB, Wold KF, et al. The Time is Ripe for a Consensus Definition of Clinical Recovery in First-episode Psychosis: Suggestions Based on a 10-Year Follow-up Study. Schizophrenia Bulletin. 2022 Jul 1;48(4):839–49.

4. Verma S, Subramaniam M, Abdin E, Poon LY, Chong SA. Symptomatic and functional remission in patients with first-episode psychosis. Acta Psychiatrica Scandinavica. 2012;126(4):282–9.

5. Petersen L, Thorup A, Øqhlenschlaeger J, Christensen TØ, Jeppesen P, Krarup G, et al. Predictors of remission and recovery in a first-episode schizophrenia spectrum disorder sample: 2-year follow-up of the OPUS trial. Can J Psychiatry. 2008 Oct;53(10):660–70.

6. Kane JM, Robinson DG, Schooler NR, Mueser KT, Penn DL, Rosenheck RA, et al. Comprehensive Versus Usual Community Care for First-Episode Psychosis: 2-Year Outcomes From the NIMH RAISE Early Treatment Program. AJP. 2016 Apr;173(4):362–72.

7. Correll CU, Galling B, Pawar A, Krivko A, Bonetto C, Ruggeri M, et al. Comparison of Early Intervention Services vs Treatment as Usual for Early-Phase Psychosis: A Systematic Review, Meta-analysis, and Meta-regression. JAMA Psychiatry. 2018 Jun 1;75(6):555–65.

8. Srihari VH, Tek C, Kucukgoncu S, Phutane VH, Breitborde NJK, Pollard J, et al. First-Episode Services for Psychotic Disorders in the U.S. Public Sector: A Pragmatic Randomized Controlled Trial. Psychiatr Serv. 2015 Jul;66(7):705–12.

9. Ruggeri M, Bonetto C, Lasalvia A, Fioritti A, de Girolamo G, Santonastaso P, et al. Feasibility and Effectiveness of a Multi-Element Psychosocial Intervention for First-Episode Psychosis: Results From the Cluster-Randomized Controlled GET UP PIANO Trial in a Catchment Area of 10 Million Inhabitants. Schizophrenia Bulletin. 2015 Sep 1;41(5):1192–203.

10. Chang WC, Kwong VWY, Chan GHK, Jim OTT, Lau ESK, Hui CLM, et al. Prediction of functional remission in first-episode psychosis: 12-month follow-up of the randomized-controlled trial on extended early intervention in Hong Kong. Schizophrenia Research. 2016 May 1;173(1):79–83.

11. Solmi M, Murru A, Pacchiarotti I, Undurraga J, Veronese N, Fornaro M, et al. Safety, tolerability, and risks associated with first- and second-generation antipsychotics: a state-of-the-art clinical review. Therapeutics and Clinical Risk Management. 2017 Jun 29;13:757–77.

12. O'Donoghue B, Piacenza F, Plapp H, Siskind D, Lyne J. Response rates to sequential trials of antipsychotic medications according to algorithms or treatment guidelines in psychotic disorders. A systematic review and meta-analysis. Schizophrenia Research. 2024 Jun 1;268:193–204.

13. Francey SM, O'Donoghue B, Nelson B, Graham J, Baldwin L, Yuen HP, et al. Psychosocial Intervention With or Without Antipsychotic Medication for First-Episode Psychosis: A Randomized Noninferiority Clinical Trial. Schizophrenia Bulletin Open. 2020 Jan 1;1(1):sgaa015.

14. Galletly C, Castle D, Dark F, Humberstone V, Jablensky A, Killackey E, et al. Royal Australian and New Zealand College of Psychiatrists clinical practice guidelines for the management of schizophrenia and related disorders. Aust N Z J Psychiatry. 2016 May 1;50(5):410–72.

15. Kuipers E, Yesufu-Udechuku A, Taylor C, Kendall T. Management of psychosis and schizophrenia in adults: summary of updated NICE guidance. BMJ. 2014 Feb 12;348:g1173.

16. Griffiths SL, Lalousis PA, Wood SJ, Upthegrove R. Heterogeneity in treatment outcomes and incomplete recovery in first episode psychosis: does one size fit all? Transl Psychiatry. 2022 Nov 17;12(1):1–6.

17. van Dellen E, Bohlken MM, Draaisma L, Tewarie PK, van Lutterveld R, Mandl R, et al. Structural Brain Network Disturbances in the Psychosis Spectrum. Schizophrenia Bulletin. 2016 May 1;42(3):782–9.

18. Skudlarski P, Jagannathan K, Anderson K, Stevens MC, Calhoun VD, Skudlarska BA, et al. Brain connectivity is not only lower but different in schizophrenia: A combined anatomical and functional approach. Biol Psychiatry. 2010 Jul 1;68(1):61–9.

19. Chopra S, Levi PT, Holmes A, Orchard ER, Segal A, Francey SM, et al. Brainwide Anatomical Connectivity and Prediction of Longitudinal Outcomes in Antipsychotic-Naïve First-Episode Psychosis. Biological Psychiatry. 2025 Jan 15;97(2):157–66.

20. O'Neill A, Mechelli A, Bhattacharyya S. Dysconnectivity of Large-Scale Functional Networks in Early Psychosis: A Meta-analysis. Schizophrenia Bulletin. 2019 Apr 25;45(3):579–90.

21. Li T, Wang Q, Zhang J, Rolls ET, Yang W, Palaniyappan L, et al. Brain-Wide Analysis of Functional Connectivity in First-Episode and Chronic Stages of Schizophrenia. Schizophrenia Bulletin. 2017 Mar 1;43(2):436–48.

22. Gong Q, Hu X, Pettersson-Yeo W, Xu X, Lui S, Crossley N, et al. Network-Level Dysconnectivity in Drug-Naïve First-Episode Psychosis: Dissociating Transdiagnostic and Diagnosis-Specific Alterations. Neuropsychopharmacol. 2017 Mar;42(4):933–40.

23. van Dellen E, Börner C, Schutte M, van Montfort S, Abramovic L, Boks MP, et al. Functional brain networks in the schizophrenia spectrum and bipolar disorder with psychosis. npj Schizophr. 2020 Sep 2;6(1):1–9.

24. Baker JT, Dillon DG, Patrick LM, Roffman JL, Brady RO, Pizzagalli DA, et al. Functional connectomics of affective and psychotic pathology. Proceedings of the National Academy of Sciences. 2019 Apr 30;116(18):9050–9.

25. Cheng H, Newman S, Goñi J, Kent JS, Howell J, Bolbecker A, et al. Nodal centrality of functional network in the differentiation of schizophrenia. Schizophrenia Research. 2015 Oct 1;168(1):345–52.

26. Liu Y, Liang M, Zhou Y, He Y, Hao Y, Song M, et al. Disrupted small-world networks in schizophrenia. Brain. 2008 Apr 1;131(4):945–61.

27. Jensen KM, Calhoun VD, Fu Z, Yang K, Faria AV, Ishizuka K, et al. A whole-brain neuromark resting-state fMRI analysis of first-episode and early psychosis: Evidence of aberrant cortical-subcortical-cerebellar functional circuitry. NeuroImage: Clinical. 2024 Jan 1;41:103584.

28. Chopra S, Francey SM, O'Donoghue B, Sabaroedin K, Arnatkeviciute A, Cropley V, et al. Functional Connectivity in Antipsychotic-Treated and Antipsychotic-Naive Patients With First-Episode Psychosis and Low Risk of Self-harm or Aggression: A Secondary Analysis of a Randomized Clinical Trial. JAMA Psychiatry. 2021 Sep 1;78(9):994–1004.

29. Del Fabro L, Schmidt A, Fortea L, Delvecchio G, D'Agostino A, Radua J, et al. Functional brain network dysfunctions in subjects at high-risk for psychosis: A meta-analysis of resting-state functional connectivity. Neuroscience & Biobehavioral Reviews. 2021 Sep 1;128:90–101.

30. Ramsay IS, Mueller B, Ma Y, Shen C, Sponheim SR. Thalamocortical connectivity and its relationship with symptoms and cognition across the psychosis continuum. Psychological Medicine. 2023 Sep;53(12):5582–91.

31. Wang D, Li M, Wang M, Schoeppe F, Ren J, Chen H, et al. Individual-specific functional connectivity markers track dimensional and categorical features of psychotic illness. Mol Psychiatry. 2020 Sep;25(9):2119–29.

32. Doucet GE, Moser DA, Luber MJ, Leibu E, Frangou S. Baseline brain structural and functional predictors of clinical outcome in the early course of schizophrenia. Mol Psychiatry. 2020 Apr;25(4):863–72.

33. Zhang Y, Xu L, Hu Y, Wu J, Li C, Wang J, et al. Functional Connectivity Between Sensory-Motor Subnetworks Reflects the Duration of Untreated Psychosis and Predicts Treatment Outcome of First-Episode Drug-Naïve Schizophrenia. Biological Psychiatry: Cognitive Neuroscience and Neuroimaging. 2019 Aug 1;4(8):697–705.

34. Choi SY, Ha M, Choi S, Moon SY, Park S, Kim M, et al. Altered intrinsic cerebellar-cerebral functional connectivity is related to negative symptoms in patients with first-episode psychosis. Schizophrenia Research. 2023 Feb 1;252:56–63.

35. Mehta UM, Ibrahim FA, Sharma MS, Venkatasubramanian G, Thirthalli J, Bharath RD, et al. Resting-state functional connectivity predictors of treatment response in schizophrenia – A systematic review and meta-analysis. Schizophrenia Research. 2021 Nov 1;237:153–65.

36. Anticevic A, Hu X, Xiao Y, Hu J, Li F, Bi F, et al. Early-Course Unmedicated Schizophrenia Patients Exhibit Elevated Prefrontal Connectivity Associated with Longitudinal Change. J Neurosci. 2015 Jan 7;35(1):267–86.

37. Dwyer DB, Falkai P, Koutsouleris N. Machine Learning Approaches for Clinical Psychology and Psychiatry. Annual Review of Clinical Psychology. 2018 May 7;14(Volume 14, 2018):91–118.

38. Cao H, Lencz T, Gallego JA, Rubio JM, John M, Barber AD, et al. A Functional Connectome-Based Neural Signature for Individualized Prediction of Antipsychotic Response in First-Episode Psychosis. AJP. 2023 Nov;180(11):827–35.

39. Blessing EM, Murty VP, Zeng B, Wang J, Davachi L, Goff DC. Anterior Hippocampal–Cortical Functional Connectivity Distinguishes Antipsychotic Naïve First-Episode Psychosis Patients From Controls and May Predict Response to Second-Generation Antipsychotic Treatment. Schizophrenia Bulletin. 2020 Apr 10;46(3):680–9.

40. Sarpal DK, Argyelan M, Robinson DG, Szeszko PR, Karlsgodt KH, John M, et al. Baseline Striatal Functional Connectivity as a Predictor of Response to Antipsychotic Drug Treatment. Am J Psychiatry. 2016 Jan;173(1):69–77.

41. Cao H, Wei X, Hu N, Zhang W, Xiao Y, Zeng J, et al. Cerebello-Thalamo-Cortical Hyperconnectivity Classifies Patients and Predicts Long-Term Treatment Outcome in First-Episode Schizophrenia. Schizophrenia Bulletin. 2022 Mar 1;48(2):505–13.

42. Cao B, Cho RY, Chen D, Xiu M, Wang L, Soares JC, et al. Treatment response prediction and individualized identification of first-episode drug-naïve schizophrenia using brain functional connectivity. Mol Psychiatry. 2020 Apr;25(4):906–13.

43. Liu W, Fang P, Guo F, Qiao Y, Zhu Y, Wang H. Graph-Theory-Based Degree Centrality Combined with Machine Learning Algorithms Can Predict Response to Treatment with Antipsychotic Medications in Patients with First-Episode Schizophrenia. Disease Markers. 2022 Oct 13;2022:e1853002.

44. Li H, Ou Y, Liu F, Chen J, Zhao J, Guo W, et al. Reduced connectivity in anterior cingulate cortex as an early predictor for treatment response in drug-naive, first-episode schizophrenia: A global-brain functional connectivity analysis. Schizophrenia Research. 2020 Jan 1;215:337–43.

45. McGorry PD, Killackey E, Yung A. Early intervention in psychosis: concepts, evidence and future directions. World Psychiatry. 2008 Oct;7(3):148–56.

46. Kern RS, Glynn SM, Horan WP, Marder SR. Psychosocial Treatments to Promote Functional Recovery in Schizophrenia. Schizophrenia Bulletin. 2009 Mar 1;35(2):347–61.

47. Breitborde NJ, Moe AM, Ered A, Ellman LM, Bell EK. Optimizing psychosocial interventions in first-episode psychosis: current perspectives and future directions. Psychology Research and Behavior Management. 2017 Apr 27;10:119–28.

48. Lahera G, Gálvez JL, Sánchez P, Martínez-Roig M, Pérez-Fuster JV, García-Portilla P, et al. Functional recovery in patients with schizophrenia: recommendations from a panel of experts. BMC Psychiatry. 2018 Jun 5;18:176.

49. O'Donoghue B, Francey SM, Nelson B, Ratheesh A, Allott K, Graham J, et al. Staged treatment and acceptability guidelines in early psychosis study (STAGES): A randomized placebo controlled trial of intensive psychosocial treatment plus or minus antipsychotic medication for first-episode psychosis with low-risk of self-harm or aggression. Study protocol and baseline characteristics of participants. Early Intervention in Psychiatry. 2019;13(4):953–60.

50. Orygen Youth Health., Early Psychosis Prevention & Intervention Centre (Vic.). Cognitive-behavioural case management in early psychosis : a handbook. Parkville, Vic: Orygen Youth Health; 2010.

51. Royston P, Altman DG, Sauerbrei W. Dichotomizing continuous predictors in multiple regression: a bad idea. Stat Med. 2006 Jan 15;25(1):127–41.

52. Esteban O, Birman D, Schaer M, Koyejo OO, Poldrack RA, Gorgolewski KJ. MRIQC: Advancing the automatic prediction of image quality in MRI from unseen sites. PLOS ONE. 2017 Sep 25;12(9):e0184661.

53. Parkes L, Fulcher B, Yücel M, Fornito A. An evaluation of the efficacy, reliability, and sensitivity of motion correction strategies for resting-state functional MRI. NeuroImage. 2018 May 1;171:415–36.

54. Esteban O, Markiewicz CJ, Blair RW, Moodie CA, Isik AI, Erramuzpe A, et al. fMRIPrep: a robust preprocessing pipeline for functional MRI. Nat Methods. 2019 Jan;16(1):111–6.

55. Pruim RHR, Mennes M, van Rooij D, Llera A, Buitelaar JK, Beckmann CF. ICA-AROMA: A robust ICA-based strategy for removing motion artifacts from fMRI data. NeuroImage. 2015 May 15;112:267–77.

56. Schaefer A, Kong R, Gordon EM, Laumann TO, Zuo XN, Holmes AJ, et al. Local-Global Parcellation of the Human Cerebral Cortex from Intrinsic Functional Connectivity MRI. Cereb Cortex. 2018 Sep 1;28(9):3095–114.

57. Tian Y, Margulies DS, Breakspear M, Zalesky A. Topographic organization of the human subcortex unveiled with functional connectivity gradients. Nat Neurosci. 2020 Nov;23(11):1421–32.

58. Chen P, An L, Wulan N, Zhang C, Zhang S, Ooi LQR, et al. Multilayer meta-matching: Translating phenotypic prediction models from multiple datasets to small data. Imaging Neuroscience. 2024 Jul 17;2:1–22.

59. Fischl B, Salat DH, Busa E, Albert M, Dieterich M, Haselgrove C, et al. Whole Brain Segmentation: Automated Labeling of Neuroanatomical Structures in the Human Brain. Neuron. 2002 Jan 31;33(3):341–55.

60. Fischl B, Salat DH, van der Kouwe AJW, Makris N, Ségonne F, Quinn BT, et al. Sequence-independent segmentation of magnetic resonance images. NeuroImage. 2004 Jan 1;23:S69–84.

61. Shen X, Finn ES, Scheinost D, Rosenberg MD, Chun MM, Papademetris X, et al. Using connectome-based predictive modeling to predict individual behavior from brain connectivity. Nat Protoc. 2017 Mar;12(3):506–18.

62. Murphy KP. Machine learning: a probabilistic perspective. 4. print. (fixed many typos). Cambridge, Mass.: MIT Press; 2013. 1071 p. (Adaptive computation and machine learning series).

63. Li J, Kong R, Liégeois R, Orban C, Tan Y, Sun N, et al. Global signal regression strengthens association between resting-state functional connectivity and behavior. NeuroImage. 2019 Aug 1;196:126–41.

64. He T, Kong R, Holmes AJ, Nguyen M, Sabuncu MR, Eickhoff SB, et al. Deep neural networks and kernel regression achieve comparable accuracies for functional connectivity prediction of behavior and demographics. NeuroImage. 2020 Feb 1;206:116276.

65. Janssen RJ, Mourão-Miranda J, Schnack HG. Making Individual Prognoses in Psychiatry Using Neuroimaging and Machine Learning. Biological Psychiatry: Cognitive Neuroscience and Neuroimaging. 2018 Sep 1;3(9):798–808.

66. Varoquaux G. Cross-validation failure: Small sample sizes lead to large error bars. NeuroImage. 2018 Oct 15;180:68–77.

67. Marek S, Tervo-Clemmens B, Calabro FJ, Montez DF, Kay BP, Hatoum AS, et al. Reproducible brain-wide association studies require thousands of individuals. Nature. 2022 Mar;603(7902):654–60.

68. Poldrack RA, Huckins G, Varoquaux G. Establishment of Best Practices for Evidence for Prediction A Review. JAMA Psychiatry. 2020 May 1;77(5):534–40.

69. Arbabshirani MR, Plis S, Sui J, Calhoun VD. Single subject prediction of brain disorders in neuroimaging: Promises and pitfalls. Neuroimage. 2017 Jan 15;145(Pt B):137–65.

70. Smith SM, Nichols TE, Vidaurre D, Winkler AM, Behrens TEJ, Glasser MF, et al. A positive-negative mode of population covariation links brain connectivity, demographics and behavior. Nat Neurosci. 2015 Nov;18(11):1565–7.

71. Xia CH, Ma Z, Ciric R, Gu S, Betzel RF, Kaczkurkin AN, et al. Linked dimensions of psychopathology and connectivity in functional brain networks. Nat Commun. 2018 Aug 1;9(1):3003.

72. Kebets V, Holmes AJ, Orban C, Tang S, Li J, Sun N, et al. Somatosensory-Motor Dysconnectivity Spans Multiple Transdiagnostic Dimensions of Psychopathology. Biol Psychiatry. 2019 Nov 15;86(10):779–91.

73. Chopra S, Dhamala E, Lawhead C, Ricard JA, Orchard ER, An L, et al. Generalizable and replicable brain-based predictions of cognitive functioning across common psychiatric illness. Science Advances. 2024 Nov 6;10(45):eadn1862.

74. Bycroft C, Freeman C, Petkova D, Band G, Elliott LT, Sharp K, et al. The UK Biobank resource with deep phenotyping and genomic data. Nature. 2018 Oct;562(7726):203–9.

75. Volkow ND, Koob GF, Croyle RT, Bianchi DW, Gordon JA, Koroshetz WJ, et al. The conception of the ABCD study: From substance use to a broad NIH collaboration. Dev Cogn Neurosci. 2018 Aug;32:4–7.

76. Holmes AJ, Hollinshead MO, O'Keefe TM, Petrov VI, Fariello GR, Wald LL, et al. Brain Genomics Superstruct Project initial data release with structural, functional, and behavioral measures. Sci Data. 2015 Jul 7;2(1):150031.

77. Alexander LM, Escalera J, Ai L, Andreotti C, Febre K, Mangone A, et al. An open resource for transdiagnostic research in pediatric mental health and learning disorders. Sci Data. 2017 Dec 19;4(1):170181.

78. Nooner KB, Colcombe SJ, Tobe RH, Mennes M, Benedict MM, Moreno AL, et al. The NKI-Rockland Sample: A Model for Accelerating the Pace of Discovery Science in Psychiatry. Front Neurosci. 2012;6:152.

79. Thomas GE. Resampling-Based Multiple Testing: Examples and Methods for p-Value Adjustment. Journal of the Royal Statistical Society Series D: The Statistician. 1994 Jun 1;43(2):347–8.

80. Lee S, Lee DK. What is the proper way to apply the multiple comparison test? Korean J Anesthesiol. 2018 Oct;71(5):353–60.

81. Alberton BAV, Nichols TE, Gamba HR, Winkler AM. Multiple testing correction over contrasts for brain imaging. NeuroImage. 2020 Aug 1;216:116760.

82. Henry LP, Amminger GP, Harris MG, Yuen HP, Harrigan SM, Prosser AL, et al. The EPPIC Follow-Up Study of First-Episode Psychosis: Longer-Term Clinical and Functional Outcome 7 Years After Index Admission. J Clin Psychiatry. 2010 Jun 15;71(6):6560.

83. Leucht S, Kane JM, Kissling W, Hamann J, Etschel E, Engel R. Clinical implications of Brief Psychiatric Rating Scale scores. The British Journal of Psychiatry. 2005 Oct;187(4):366–71.

84. Chekroud AM, Hawrilenko M, Loho H, Bondar J, Gueorguieva R, Hasan A, et al. Illusory generalizability of clinical prediction models. Science. 2024 Jan 12;383(6679):164–7.

546 85. Slot MIE, Urquijo Castro MF, Winter - van Rossum I, van Hell HH, Dwyer D, Dazzan P, et al. Multivariable
547 prediction of functional outcome after first-episode psychosis: a crossover validation approach in EUFEST and
548 PSYSCAN. Schizophr. 2024 Oct 7;10(1):1–11.

549 86. Lin Q, Dai Z, Xia M, Han Z, Huang R, Gong G, et al. A connectivity-based test-retest dataset of multi-modal
550 magnetic resonance imaging in young healthy adults. Sci Data. 2015 Oct 27;2(1):150056.

551 87. Tagliazucchi E, Laufs H. Decoding Wakefulness Levels from Typical fMRI Resting-State Data Reveals Reliable
552 Drifts between Wakefulness and Sleep. Neuron. 2014 May 7;82(3):695–708.

553 88. Geerligs L, Rubinov M, Cam-CAN, Henson RN. State and Trait Components of Functional Connectivity:
554 Individual Differences Vary with Mental State. J Neurosci. 2015 Oct 14;35(41):13949–61.

555 89. Raut RV, Snyder AZ, Mitra A, Yellin D, Fujii N, Malach R, et al. Global waves synchronize the brain's
556 functional systems with fluctuating arousal. Science Advances. 2021 Jul 21;7(30):eabf2709.

557 90. Wang HT, Bzdok D, Margulies D, Craddock C, Milham M, Jefferies E, et al. Patterns of thought: Population
558 variation in the associations between large-scale network organisation and self-reported experiences at rest.
559 Neuroimage. 2018 Aug 1;176:518–27.

560 91. Harrison BJ, Pujol J, Ortiz H, Fornito A, Pantelis C, Yücel M. Modulation of brain resting-state networks by sad
561 mood induction. PLoS One. 2008 Mar 19;3(3):e1794.

562 92. Finn ES, Bandettini PA. Movie-watching outperforms rest for functional connectivity-based prediction of
563 behavior. NeuroImage. 2021 Jul 15;235:117963.

564 93. Cao H, Chen OY, McEwen SC, Forsyth JK, Gee DG, Bearden CE, et al. Cross-paradigm connectivity:
565 reliability, stability, and utility. Brain Imaging Behav. 2021 Apr;15(2):614–29.

566 94. Gao S, Greene AS, Constable RT, Scheinost D. Combining Multiple Connectomes Improves Predictive
567 Modeling of Phenotypic Measures. Neuroimage. 2019 Nov 1;201:116038.

568 95. Greene AS, Gao S, Scheinost D, Constable RT. Task-induced brain state manipulation improves prediction of
569 individual traits. Nat Commun. 2018 Jul 18;9(1):2807.

570 96. Ooi LQR, Chen J, Zhang S, Kong R, Tam A, Li J, et al. Comparison of individualized behavioral predictions
571 across anatomical, diffusion and functional connectivity MRI. NeuroImage. 2022 Nov 1;263:119636.

572 97. He T, An L, Chen P, Chen J, Feng J, Bzdok D, et al. Meta-matching as a simple framework to translate
573 phenotypic predictive models from big to small data. Nat Neurosci. 2022 Jun;25(6):795–804.

574 98. Bell M, Milstein R, Beam-Goulet J, Lysaker P, Cicchetti D. The Positive and Negative Syndrome Scale and the
575 Brief Psychiatric Rating Scale: Reliability, Comparability, and Predictive Validity. The Journal of Nervous and
576 Mental Disease. 1992 Nov;180(11):723.

577 99. Opler MGA, Yavorsky C, Daniel DG. Positive and Negative Syndrome Scale (PANSS) Training. Innov Clin
578 Neurosci. 2017 Dec 1;14(11–12):77–81.

579 100. Fried EI, Nesse RM. Depression sum-scores don't add up: why analyzing specific depression symptoms is
580 essential. BMC Medicine. 2015 Apr 6;13(1):72.

581 101. Tiego J, Martin EA, DeYoung CG, Hagan K, Cooper SE, Pasion R, et al. Precision behavioral phenotyping as a
582 strategy for uncovering the biological correlates of psychopathology. Nat Ment Health. 2023 May;1(5):304–15.

583 102. Strauss ME, Smith GT. Construct validity: advances in theory and methodology. Annu Rev Clin Psychol.
584 2009;5:1–25.

585 103. Hochstrasser L, Studerus E, Riecher-Rössler A, Schimmelmann BG, Lambert M, Lang UE, et al. Latent state-
586 trait structure of BPRS subscales in clinical high-risk state and first episode psychosis. Sci Rep. 2022 Apr
587 22;12:6652.

588 104. Shafer A. Meta-analysis of the Brief Psychiatric Rating Scale factor structure. Psychological Assessment.
589 2005;17(3):324–35.

590 105. Tiego J, Thompson K, Arnatkeviciute A, Hawi Z, Finlay A, Sabaroedin K, et al. Dissecting Schizotypy and Its
591 Association With Cognition and Polygenic Risk for Schizophrenia in a Nonclinical Sample. Schizophrenia
592 Bulletin. 2023 Sep 1;49(5):1217–28.

593 106. Soldatos RF, Cearns M, Nielsen MØ, Kollias C, Xenaki LA, Stefanatou P, et al. Prediction of Early Symptom
594 Remission in Two Independent Samples of First-Episode Psychosis Patients Using Machine Learning.
595 Schizophrenia Bulletin. 2022 Jan 1;48(1):122–33.

596 107. Demjaha A, Lappin JM, Stahl D, Patel MX, MacCabe JH, Howes OD, et al. Antipsychotic treatment resistance
597 in first-episode psychosis: prevalence, subtypes and predictors. Psychological Medicine. 2017 Aug;47(11):1981–
598 9.

599 108. Leighton SP, Upthegrove R, Krishnadas R, Benros ME, Broome MR, Gkoutos GV, et al. Development and
600 validation of multivariable prediction models of remission, recovery, and quality of life outcomes in people with
601 first episode psychosis: a machine learning approach. The Lancet Digital Health. 2019 Oct 1;1(6):e261–70.

602 109. Conus P, Cotton SM, Francey SM, O'Donoghue B, Schimmelmann BG, McGorry PD, et al. Predictors of
603 favourable outcome in young people with a first episode psychosis without antipsychotic medication.
604 Schizophrenia Research. 2017 Jul 1;185:130–6.

605 110. Koutsouleris N, Kahn RS, Chekroud AM, Leucht S, Falkai P, Wobrock T, et al. Multisite prediction of 4-week
606 and 52-week treatment outcomes in patients with first-episode psychosis: a machine learning approach. The
607 Lancet Psychiatry. 2016 Oct 1;3(10):935–46.

608 111. Alvarez-Jimenez M, O'Donoghue B, Thompson A, Gleeson JF, Bendall S, Gonzalez-Blanch C, et al. Beyond
609 Clinical Remission in First Episode Psychosis: Thoughts on Antipsychotic Maintenance vs. Guided
610 Discontinuation in the Functional Recovery Era. CNS Drugs. 2016 May 1;30(5):357–68.

611 112. Bergé D, Mané A, Salgado P, Cortizo R, Garnier C, Gomez L, et al. Predictors of Relapse and Functioning in
612 First-Episode Psychosis: A Two-Year Follow-Up Study. Psychiatr Serv. 2016 Feb;67(2):227–33.

613 113. Friis S, Melle I, Johannessen JO, Røssberg JI, Barder HE, Evensen JH, et al. Early Predictors of Ten-Year
614 Course in First-Episode Psychosis. PS. 2016 Apr;67(4):438–43.

615 114. Lally J, Ajnakina O, Forti MD, Trotta A, Demjaha A, Kolliakou A, et al. Two distinct patterns of treatment
616 resistance: clinical predictors of treatment resistance in first-episode schizophrenia spectrum psychoses.
617 Psychological Medicine. 2016 Nov;46(15):3231–40.

618 115. Smart SE, Agbedjro D, Pardiñas AF, Ajnakina O, Alameda L, Andreassen OA, et al. Clinical predictors of
619 antipsychotic treatment resistance: Development and internal validation of a prognostic prediction model by the
620 STRATA-G consortium. Schizophrenia Research. 2022 Dec 1;250:1–9.

621 116. Mondelli V, Ciufolini S, Belvederi Murri M, Bonaccorso S, Di Forti M, Giordano A, et al. Cortisol and
622 Inflammatory Biomarkers Predict Poor Treatment Response in First Episode Psychosis. Schizophrenia Bulletin.
623 2015 Sep 1;41(5):1162–70.

624 117. Nettis MA, Pergola G, Kolliakou A, O'Connor J, Bonaccorso S, David A, et al. Metabolic-inflammatory status
625 as predictor of clinical outcome at 1-year follow-up in patients with first episode psychosis.
626 Psychoneuroendocrinology. 2019 Jan 1;99:145–53.

118. Mélissa B, Sabrina G, Charles-Edouard G, Hind Z, Consortium S, Kingsada P, et al. Clinical characteristics associated with functioning trajectories following admission to a psychiatric institution: A prospective cohort study of individuals with psychosis. Psychiatry Research. 2024 Sep 1;339:116062.

119. Weibell MA, Hegelstad W ten V, Auestad B, Bramness J, Evensen J, Haahr U, et al. The Effect of Substance Use on 10-Year Outcome in First-Episode Psychosis. Schizophr Bull. 2017 Jul;43(4):843–51.

120. Brown E, Bedi G, McGorry P, O'Donoghue B. Rates and Predictors of Relapse in First-Episode Psychosis: An Australian Cohort Study. Schizophrenia Bulletin Open. 2020 Jan 1;1(1):sgaa017.

121. Betz LT, Penzel N, Kambeitz-Ilankovic L, Rosen M, Chisholm K, Stainton A, et al. General psychopathology links burden of recent life events and psychotic symptoms in a network approach. npj Schizophr. 2020 Dec 15;6(1):1–8.

122. Verdolini N, Amoretti S, Mezquida G, Cuesta MJ, Pina-Camacho L, García-Rizo C, et al. The effect of family environment and psychiatric family history on psychosocial functioning in first-episode psychosis at baseline and after 2 years. Eur Neuropsychopharmacol. 2021 Aug;49:54–68.

123. Drake RJ, Husain N, Marshall M, Lewis SW, Tomenson B, Chaudhry IB, et al. Effect of delaying treatment of first-episode psychosis on symptoms and social outcomes: a longitudinal analysis and modelling study. The Lancet Psychiatry. 2020 Jul 1;7(7):602–10.

124. Santesteban-Echarri O, Paino M, Rice S, González-Blanch C, McGorry P, Gleeson J, et al. Predictors of functional recovery in first-episode psychosis: A systematic review and meta-analysis of longitudinal studies. Clinical Psychology Review. 2017 Dec 1;58:59–75.

125. Perkins DO, Gu H, Boteva K, Lieberman JA. Relationship Between Duration of Untreated Psychosis and Outcome in First-Episode Schizophrenia: A Critical Review and Meta-Analysis. AJP. 2005 Oct;162(10):1785–804.

126. Marshall M, Lewis S, Lockwood A, Drake R, Jones P, Croudace T. Association Between Duration of Untreated Psychosis and Outcome in Cohorts of First-Episode Patients: A Systematic Review. Archives of General Psychiatry. 2005 Sep 1;62(9):975–83.

127. O'Keeffe D, Kinsella A, Waddington JL, Clarke M. 20-Year Prospective, Sequential Follow-Up Study of Heterogeneity in Associations of Duration of Untreated Psychosis With Symptoms, Functioning, and Quality of Life Following First-Episode Psychosis. Am J Psychiatry. 2022 Apr;179(4):288–97.

128. Coutts F, Koutsouleris N, McGuire P. Psychotic disorders as a framework for precision psychiatry. Nat Rev Neurol. 2023 Apr;19(4):221–34.

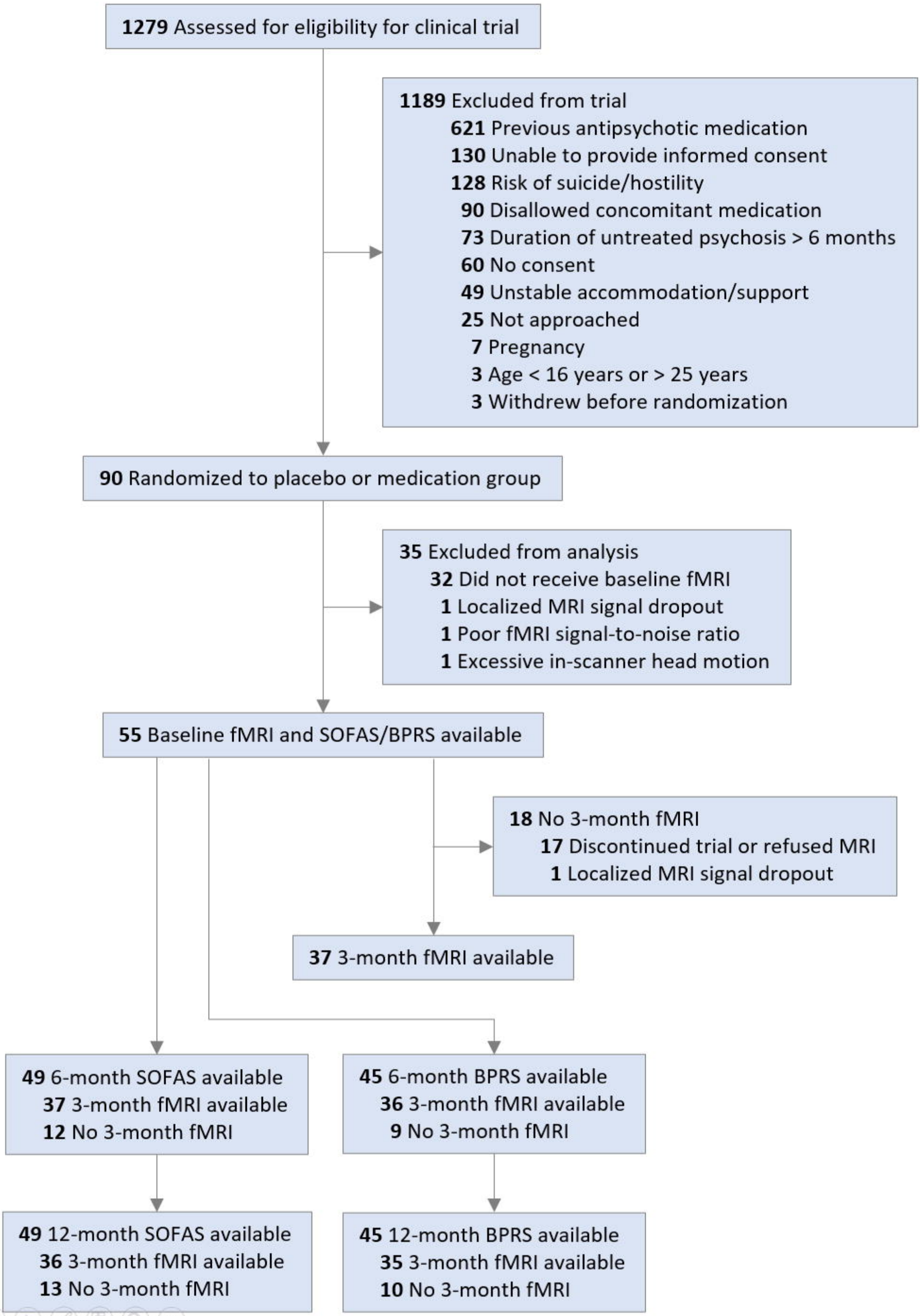
129. Ambrosen KS, Skjerbæk MW, Foldager J, Axelsen MC, Bak N, Arvastson L, et al. A machine-learning framework for robust and reliable prediction of short- and long-term treatment response in initially antipsychotic-naïve schizophrenia patients based on multimodal neuropsychiatric data. Transl Psychiatry. 2020 Aug 10;10(1):1–13.

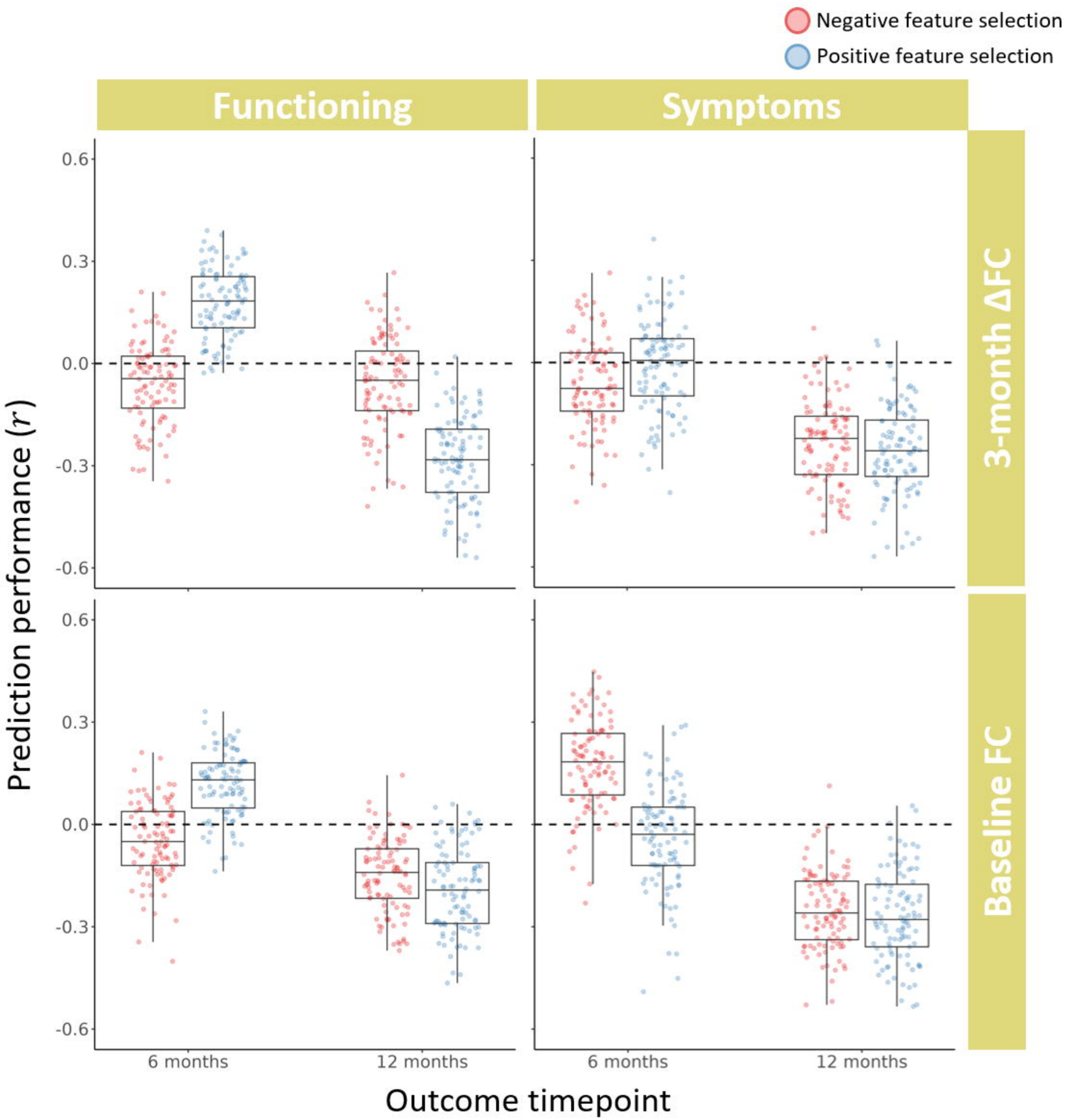
130. Ebdrup BH, Axelsen MC, Bak N, Fagerlund B, Oranje B, Raghava JM, et al. Accuracy of diagnostic classification algorithms using cognitive-, electrophysiological-, and neuroanatomical data in antipsychotic-naïve schizophrenia patients. Psychological Medicine. 2019 Dec;49(16):2754–63.

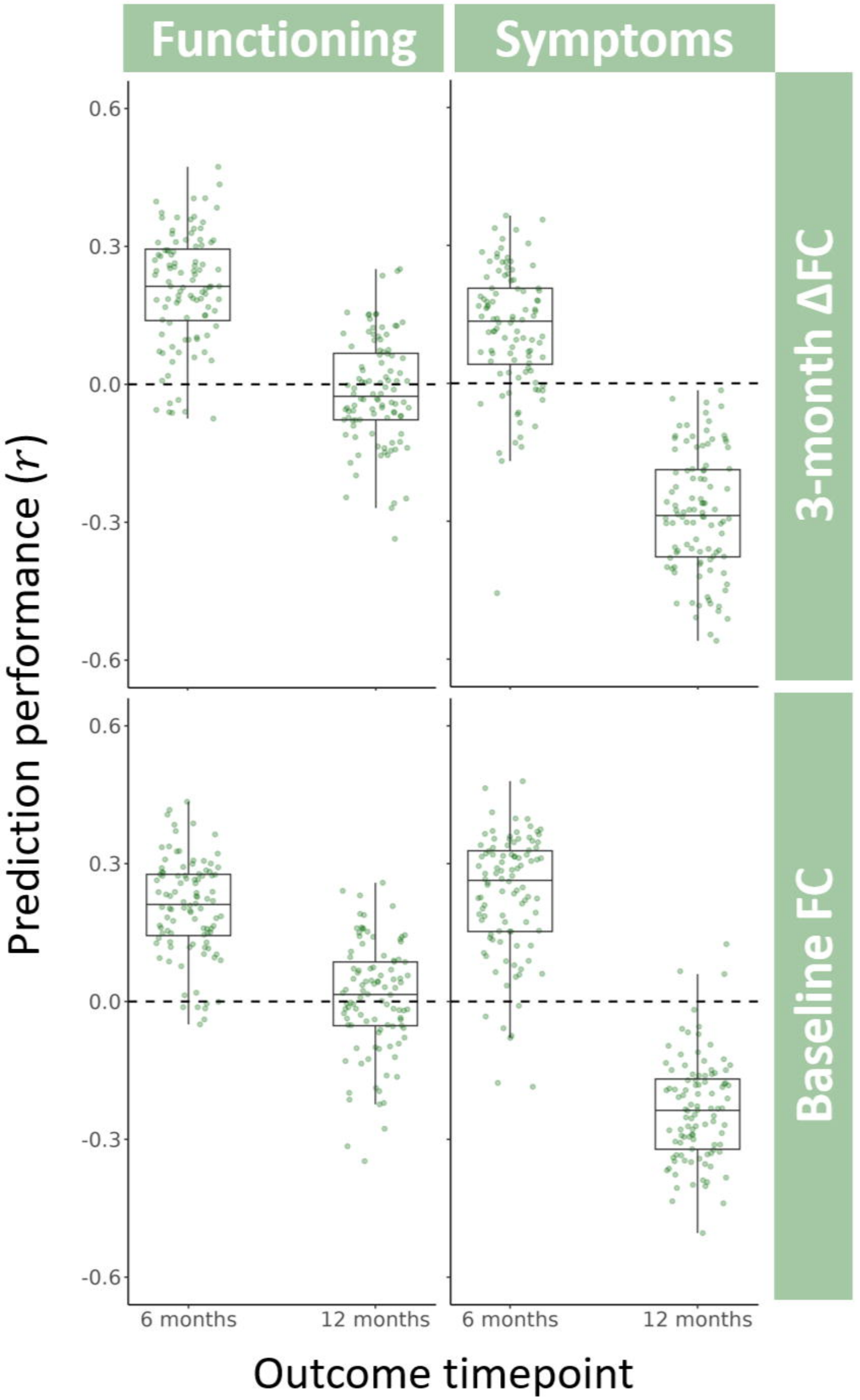
131. Etkin A, Mathalon DH. Bringing Imaging Biomarkers Into Clinical Reality in Psychiatry. JAMA Psychiatry. 2024 Nov 1;81(11):1142–7.

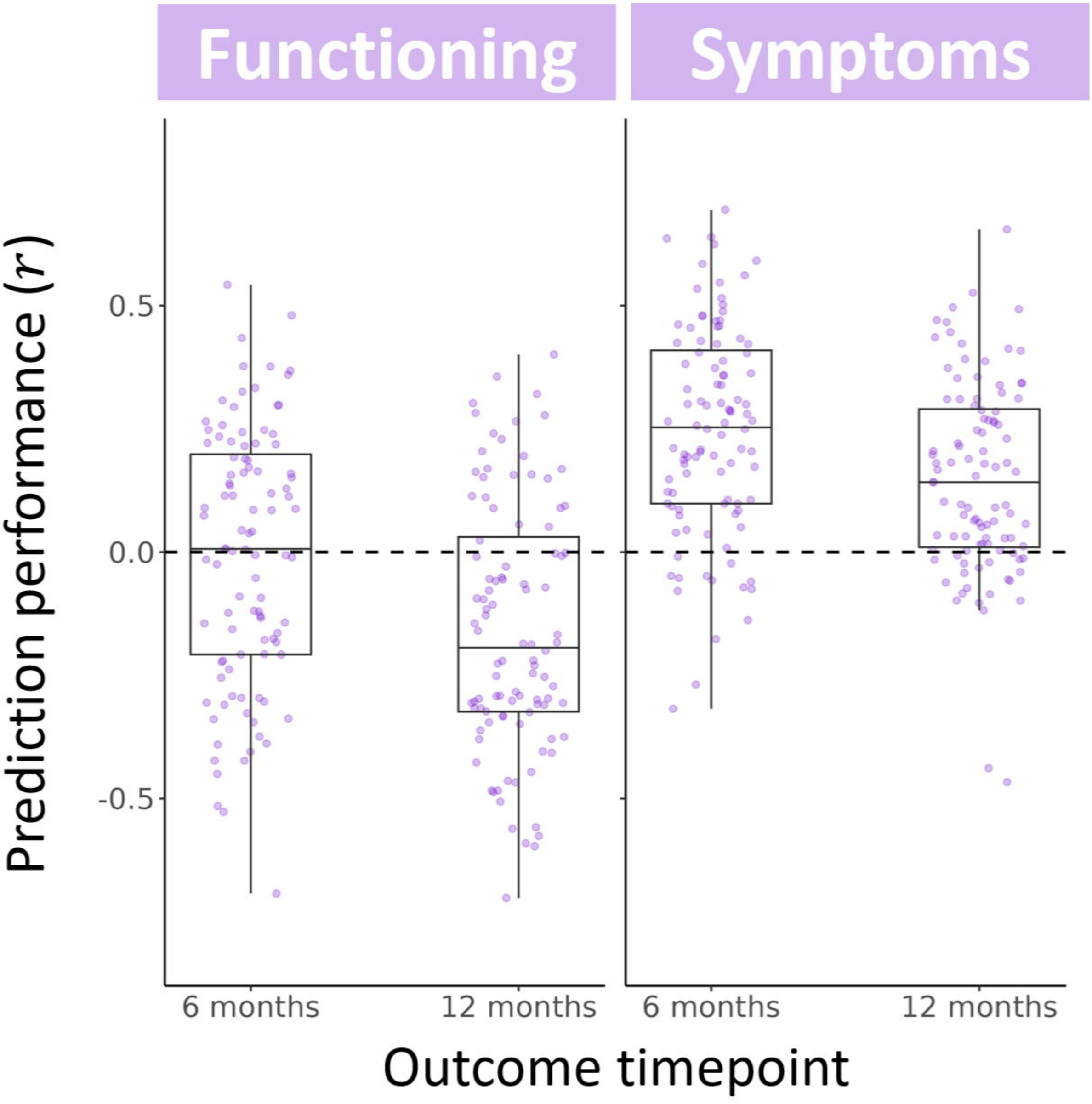
132. van Opstal DPJ, Kia SM, Jakob L, Somers M, Sommer IEC, Winter-van Rossum I, et al. Psychosis Prognosis Predictor: A continuous and uncertainty-aware prediction of treatment outcome in first-episode psychosis. Acta Psychiatrica Scandinavica. 2025;151(3):280–92.

133. Goscinski WJ, McIntosh P, Felzmann U, Maksimenko A, Hall CJ, Gureyev T, et al. The multi-modal Australian ScienceS Imaging and Visualization Environment (MASSIVE) high performance computing infrastructure: applications in neuroscience and neuroinformatics research. Front Neuroinform. 2014;8:30.


1279 Assessed for eligibility for clinical trial
1189 Excluded from trial
621 Previous antipsychotic medication
130 Unable to provide informed consent
128 Risk of suicide/hostility
90 Disallowed concomitant medication
73 Duration of untreated psychosis > 6 months
60 No consent
49 Unstable accommodation/support
25 Not approached
7 Pregnancy
3 Age < 16 years or > 25 years
3 Withdrew before randomization
90 Randomized to placebo or medication group
35 Excluded from analysis
32 Did not receive baseline fMRI
1 Localized MRI signal dropout
1 Poor fMRI signal-to-noise ratio
1 Excessive in-scanner head motion
55 Baseline fMRI and SOFAS/BPRS available
18 No 3-month fMRI
17 Discontinued trial or refused MRI
1 Localized MRI signal dropout
37 3-month fMRI available
49 6-month SOFAS available
37 3-month fMRI available
12 No 3-month fMRI
45 6-month BPRS available
36 3-month fMRI available
9 No 3-month fMRI
49 12-month SOFAS available
36 3-month fMRI available
13 No 3-month fMRI
45 12-month BPRS available
35 3-month fMRI available
10 No 3-month fMRI


Negative feature selection
Positive feature selection
Functioning
Symptoms
3-month ΔFC
Baseline FC
Prediction performance (r)
0.6
0.3
0.0
-0.3
-0.6
6 months
12 months
Outcome timepoint


Functioning
Symptoms
3-month ΔFC
Baseline FC
Prediction performance (r)
0.6
0.3
0.0
-0.3
-0.6
6 months
12 months
Outcome timepoint


Functioning
Symptoms
Prediction performance (r)
0.5
0.0
-0.5
6 months
12 months
6 months
12 months
Outcome timepoint