Structural and functional brain patterns predict formal thought disorder's severity and its persistence in recent-onset psychosis: Results from the PRONIA Study

Madalina-Octavia Buciuman[1,2,3‡]; Oemer Faruk Oeztuerk[1,2,3‡]; David Popovic[1,2,3]; Paolo Enrico[1,20]; Anne Ruef[1]; Nadia Bieler[1]; Elif Sarisik[1]; Johanna Weiske[1]; Mark Sen Dong[1]; Dominic B. Dwyer[1,18]; Lana Kambeitz-Ilankovic[4]; Shalaila S. Haas[5]; Alexandra Stainton[6]; Stephan Ruhrmann[4]; Katharine Chisholm[7]; Joseph Kambeitz[4]; Anita Riecher-Rössler[8]; Rachel Upthegrove[9]; Frauke Schultze-Lutter[10,11,12]; Raimo K. R. Salokangas[13]; Jarmo Hietala[13]; Christos Pantelis[14,15]; Rebekka Lencer[16]; Eva Meisenzahl[10]; Stephen J. Wood[6,17,18]; Paolo Brambilla[19,20]; Stefan Borgwardt[21]; Peter Falkai[1,3]; Linda A. Antonucci[1,22,23]; Alessandro Bertolino[22]; Peter Liddle[24]; Nikolaos Koutsouleris[1,3,25, CA] for the PRONIA Consortium[#]

‡ These authors contributed equally to the work

PRONIA consortium is described in the Acknowledgments

[1] Department of Psychiatry and Psychotherapy, Ludwig-Maximilian-University, Munich, Germany

[2] International Max Planck Research School for Translational Psychiatry, Munich, Germany

[3] Max Planck Institute for Psychiatry, Munich, Germany

[4] Department of Psychiatry and Psychotherapy, Faculty of Medicine and University Hospital of Cologne, University of Cologne, Germany

[5] Department of Psychiatry, Icahn School of Medicine at Mount Sinai, New York, NY, United States

[6] Centre for Youth Mental Health, University of Melbourne, Melbourne, Australia

[7] School of Psychology, Aston University, United Kingdom

[8] Faculty of Medicine, University of Basel, Basel, Switzerland

[9] Institute for Mental Health, University of Birmingham, Birmingham, United Kingdom

[10] Department of Psychiatry and Psychotherapy, Medical Faculty, Heinrich-Heine University, Düsseldorf, Germany

[11] Department of Psychology, Faculty of Psychology, Airlangga University, Surabaya, Indonesia

[12] University Hospital of Child and Adolescent Psychiatry and Psychotherapy, University of Bern, Bern, Switzerland

[13] Department of Psychiatry, University of Turku, Turku, Finland

[14] Melbourne Neuropsychiatry Centre, Department of Psychiatry, University of Melbourne, Carlton South, Vic, Australia

[15] NorthWestern Mental Health, Royal Melbourne Hospital, Parkville, VicMelbourne, Australia

[16] Department of Psychiatry and Psychotherapy, University of Münster, Germany

[17] School of Psychology, University of Birmingham, United Kingdom

[18] Orygen, the National Centre of Excellence for Youth Mental Health, Melbourne, Australia

[19] Department of Neurosciences and Mental Health, Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico, University of Milan, Milan, Italy

[20] Department of Pathophysiology and Transplantation, University of Milan, Milan, Italy

[21] Department of Psychiatry and Psychotherapy, University of Lübeck, Lüebeck, Germany

[22] Department of Basic Medical Sciences, Neuroscience and Sense Organs – University of Bari "Aldo Moro", Bari, Italy

[23] Department of Translational Biomedicine and Neuroscience (DiBraiN) – University of Bari "Aldo Moro", Bari, Italy

[24] Division of Mental Health and Clinical Neuroscience, Institute of Mental Health, School of Medicine, University of Nottingham, Nottingham, United Kingdom

[25] Department of Psychosis Studies, Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, UK

[CA] Corresponding author:
Nikolaos Koutsouleris, Department of Psychiatry and Psychotherapy, Ludwig-Maximilian-University, Nussbaumstr. 7, D-80336 Munich, Germany, Tel.: +49 (0) 89 4400, Fax: 0049 (0) 89 4400 55776, E-mail: Nikolaos.koutsouleris@med.uni-muenchen.de.

Abstract

Background and aims: Formal thought disorder (FThD) is a core feature of psychosis, and its severity and long-term persistence relates to poor clinical outcomes. However, advances in developing early recognition and management tools for FThD are hindered by a lack of insight into the brain-level predictors of FThD states and progression at the individual level.

Methods: 233 individuals with recent-onset psychosis were drawn from the multi-site European Prognostic Tools for Early Psychosis Management study. Support vector machine classifiers were trained within a cross-validation framework to separate two FThD symptom-based subgroups (high vs. low FThD severity), using cross-sectional whole-brain multi-band fractional amplitude of low frequency fluctuations (fALFF), gray-matter volume (GMV) and white-matter volume (WMV) data. Moreover, we trained machine learning models on these neuroimaging readouts to predict the persistence of high FThD subgroup membership from baseline to 1-year follow-up.

Results: Cross-sectionally, multivariate patterns of GMV within the salience, dorsal attention, visual and ventral attention networks separated the FThD severity subgroups (BAC=60.8%). Longitudinally, distributed activations/deactivations within all fALFF sub-bands (BAC_{slow-5}=73.2%, BAC_{slow-4}=72.9%, BAC_{slow-3}=68.0), GMV patterns overlapping with the cross-sectional ones (BAC=62.7%) and frontal WMV decrements (BAC=73.1%) predicted the persistence of high FThD severity from baseline to follow-up, with a combined multi-modal balanced accuracy of BAC=77%.

Conclusions: We report first evidence of brain structural and functional patterns predictive of FThD severity and persistence in early psychosis. These findings open the avenue for the development of neuroimaging-based diagnostic, prognostic and treatment options for the early recognition and management of FThD and associated poor outcomes.

1. Introduction

Early psychosis is characterized by marked symptom and outcome heterogeneity, which poses challenges for the development and implementation of tailored prevention and treatment options. In this context, modern frameworks such as the RDoC (Research Domain Criteria) (1) or HiTOP (2) highlight the relevance of symptom-based stratifications for the identification of more homogeneous neuroetiological pathways which could guide personalized early recognition, prevention and treatment.

As a multidimensional psychopathological construct, formal thought disorder (FThD) is a core and dynamically evolving feature of psychotic disorders (3,4). FThD encompasses symptoms of conceptual disorganization such as derailment, incoherence, tangentiality, neologisms, and thought blocking. Importantly, FThD has been associated with long-term adverse outcomes such as increased hospitalization, reduced quality of life, social, and occupational functioning, both in chronic schizophrenia patients (3,5–9), and in patients with recent-onset psychosis (ROP) (10–12). Employing a data-driven methodology for FThD stratification, a recent multi-site study from our group identified two ROP subgroups that differed in their FThD severity and provided support for the association between FThD intensity and functioning impairments, along with neurocognitive deficits (13). Longitudinally, FThD has been shown to exhibit a persistent symptom course in almost 40% of schizophrenia patients (7) and FThD persistence has been transdiagnostically associated with particularly poor clinical outcomes (14). Collectively, these results highlight the prognostic value of FThD symptoms and their persistence for relevant clinical outcomes. However, the currently limited insight into the brain-level alterations that underly FThD at the individual level hinders the development of FThD-tailored prevention and intervention tools.

Regarding brain anatomy, correlational evidence suggests structural gray matter volume (GMV) alterations within frontotemporal language networks associated with FThD in schizophrenia at the cross-sectional level (9,15). Furthermore, FThD sub-dimensions have been differentially related to volume reductions in specific brain regions. An extensive review (16) reported that volume reductions in six gray matter regions (superior temporal gyrus, orbitofrontal cortex, prefrontal lobe, amygdala-hippocampus, cerebellum vermis, nucleus accumbens) distinctly related to positive, negative, and global FThD dimensions. Additionally, white matter volume (WMV) reductions within language-related tracts have been associated with global scores of FThD in schizophrenia (17), and fractional anisotropy disruptions in the right posterior cingulum bundle, inferior longitudinal fasciculus and anterior thalamic radiation were differentially associated with FThD subdomains, such as disorganization and incoherence (18). Despite such FThD-associated structural brain abnormalities reported in

patients with chronic schizophrenia, the presence of similar changes underlying or predicting future FThD at first episode has rarely been studied. Vita et al. (1995) found that distractible speech and illogicality were inversely correlated with left prefrontal lobe volume, whereas incoherence and tangentiality were inversely correlated with left and right lobe volume in younger patients with schizophrenia (19). However, longitudinal approaches aiming to predict the course of FThD based on structural brain abnormalities have not been implemented so far, despite their potential for guiding preventive interventions.

Similarly, functional abnormalities within the language and executive networks have been related to FThD symptoms in schizophrenia using task-based and resting-state fMRI measures (9,20–23), with specificity for different FThD subdomains (24–27). Recent studies have increasingly transitioned from a univariate region-of-interest-based approach to a multivariate framework for analyzing functional MRI, to model the whole-brain functional correlates of FThD dimensions. For instance, Chen et al. (2021) used a whole-brain resting-state functional connectivity approach combined with machine learning algorithms to identify robust clusters associated with FThD dimensions (24). Another promising resting-state fMRI measure is the fractional amplitude of low frequency fluctuations (fALFF), a frequency-domain metric capturing the intensity of spontaneous fluctuations during rest at the voxel-level (28). fALFF is thought to reflect local functional integrity and is closely related to higher-order functional connectivity metrics (29,30). fALFF is commonly measured within multiple frequency sub-bands which seem to originate from distinct functional brain systems, encompassing both spatial and temporal dimensions of fMRI data (31–33). fALFF alterations have been consistently found in schizophrenia patients (33,34), and could therefore provide a promising measure for capturing complex patterns underlying FThD symptoms.

In summary, literature suggests that structure-function interdependencies underly FThD symptoms and calls for a better understanding of the predictive value of such changes for the development and persistence of FThD. Therefore, understanding potential neurobiological predictors of FThD severity and persistence at the individual level represents an important, yet uncharted research area, which could lead to a better characterization of the pathways underlying symptom heterogeneity in early psychosis. Our current study aimed to address this gap by investigating (I) whether the two ROP FThD subgroups previously identified by our group can be cross-sectionally differentiated based on whole-brain voxel-level gray matter volume, white matter volume, multi-band functional resting-state fALFF, and combined brain patterns, and (II) whether the persistence of high-FThD severity after 1-year of follow-up can be predicted by the same baseline neuroimaging modalities.

2. Methods

2.1. Study sample

Participants with ROP were recruited within a multisite, longitudinal study (PRONIA – Prognostic Tools for Early Psychosis Management, https://www.pronia.eu/), which included ten sites across five European countries (study protocol presented in Text S1 and Figure S1). General PRONIA inclusion and exclusion criteria are presented in Table S1. ROP patients fulfilled the DSM-IV-TR criteria for a lifetime affective and non-affective psychotic episode, having the psychotic episode within the past 3 months, with the onset of the psychotic episode occurring within the past 24 months. Being treated with antipsychotic medication for longer than 90 days at or above the minimum dosage of the first-episode psychosis range of the German Association for Psychiatry and Psychotherapy (DGPPN) S3 guideline (35) was a specific exclusion criterion. For the current analyses, we excluded 46 out of the 279 available ROP patients included in our previous clustering study (13) due to missing or low-quality structural or functional MRI scans, leading to the inclusion of 233 ROP patients (47,6% female, average age = 24.6 (SD = 5.6)) in the current study (exclusion steps are presented in Figure S1A). Additionally, for the persistence analyses, we additionally excluded 80 participants due to missing clinical follow-up data required for FThD subgroup assignment, leaving 153 ROP patients in the persistence analyses (43% female, average age = 24 (SD = 5.5)) (Figure S1A). The distribution of the ROP patients across diagnosis categories based on the *ICD-10 classification of mental and behavioural disorders* by *World Health Organization* (WHO) is provided in Figure S2.

Adult participants gave written informed consent, while patients younger than 18 years-of-age and their guardians provided written informed assent and consent, respectively. The study was registered at the German Clinical Trials Register (DRKS00005042) and approved by the local research ethics committees of all sites.

2.2. Formal thought disorder subgroup definition and assignment

The ROP patients were assigned to either a high or a low FThD subgroup based on a clustering analysis of FThD symptoms described in our previous work (13). We used 5 items from the Positive and Negative Symptom Scale (*PANSS)*(36) conceptual disorganization (P2); difficulty in abstract thinking (N5)) and the Scale for the Assessment of Negative Symptoms (*SANS)*(37); poverty of speech (9), poverty of content speech (10), increased latency of response (12)) in order to operationalize FThD as proposed elsewhere (3). An R-based clustering validation package (*ClValid*, (38)) was used to select a two-cluster k-means-based solution corresponding to high and low FThD subgroups.

In the current study, we used a subsample of the original ROP sample used by Oeztuerk et al., 2021 (13), according to MRI data availability (Text S1; Figure S1B). Cross-sectionally, we classified the high vs low FThD subgroups based on their structural and functional imaging data, directly using the subgroup assignment of the patients from our previous study (13). The same clustering model was used to assign individuals to the FThD subgroups at their 1-year follow-up study visit, using the squared Euclidean distance from each cluster center in the k-means based clustering model. This allowed us to identify and classify individuals with persistent high FThD severity based on the original clustering solution of the previous study (13), without building any new clustering models.

The two baseline FThD subgroups, as well as the persistent vs non-persistent high FThD symptom (at follow-up) groups were compared in terms of clinical and socio-demographic characteristics.

2.3. MRI data acquisition and processing

2.3.1. Structural MRI data

The MRI data was minimally harmonized between the different PRONIA sites to preserve the heterogeneity of MRI data acquisition present in the clinical context (39)(Text S2). The acquisition parameters used at each PRONIA site are presented in Table S2.

Images were pre-processed using a standardized pipeline implemented in the Statistical Parametric Mapping software (SPM12, http://www.fil.ion.ucl.ac.uk/spm)-based CAT12 toolbox (version r1155; http://dbm.neuro.uni-jena.de/cat12/). Briefly, the scans underwent segmentation (into white matter, gray matter, and cerebrospinal fluid) and registration to the *Montreal Neurological Institute* (MNI-152) space based on the DARTEL algorithm. More details on the CAT12 preprocessing procedures are provided in the Text S2 and in previous studies from our group (39,40). Following high-dimensional registration, the gray matter and white matter maps were modulated using the Jacobian determinants obtained during the registration to produce GMV and WMV maps, respectively. The volumetric images were resliced to a resolution of 3 mm^3 and corrected for total intracranial volume using global scaling. Images with a quality score lower or equal to C (satisfactory) based on the CAT12 Quality Assurance framework were excluded from further analyses (detailed in Text S2).

2.3.2. Resting-state fMRI data

The rs-fMRI data was collected using Echo Planar Imaging (EPI) sequences, a repetition time of 3 seconds and a total of 200 volumes across all sites (acquisition parameters per site are detailed in Table S3).

Data preprocessing was conducted using a pipeline previously developed in the PRONIA consortium (41), mainly based on SPM12 and the Resting State fMRI data analysis Toolkit (REST, version 1.848; http://www.restfmri.net/) (Text S2). In brief, the first 8 volumes were discarded, images were slice-time corrected and realigned to the first volume. The functional maps were co-registered to the T1 images, resliced and normalized to the common MNI space. Then, nuisance covariates were regressed out, including white matter, cerebrospinal fluid and Friston 24 motion parameters. Lastly, the images were smoothed, motion-corrected using time-series despiking and detrended (detailed in Text S2). Images with more than 38.5% of volumes with a mean framewise displacement over 0.50 mm were discarded (42).

The fALFF was computed as the ratio of the integrated power spectrum of the frequency range of interest to that of the entire frequency range, after time series were transformed to the frequency domain using a fast Fourier transform. Based on previous literature showing different properties and generators of the activity in different low-frequency sub-bands (31,43), we chose three different sub-bands: slow-5 (0.01 – 0.027 Hz), slow-4 (0.027 – 0.073 Hz) and slow-3 (0.073 – 0.198 Hz). The fALFF maps were z-score standardized voxel-wise within the gray matter mask for each participant.

2.4. Machine learning procedure

All machine learning models were built using the MATLAB-based open-source toolbox NeuroMiner (version 1.1; http://www.pronia.eu/neurominer), within a nested repeated cross-validation framework as detailed in Text S3. Based on the baseline structural and functional MRI data, supervised classifiers were trained to separate patients with high and low FThD cross-sectionally, as well as patients with persistent high-FThD (high-FThD group membership both at baseline and the 1-year follow-up) from those with any other FThD symptom courses.

Both the structural and functional MRI data were preprocessed within an outer-leave-site-out/ inner-pooled nested cross-validation structure for the cross-sectional models, and a pooled nested cross-validation structure for the persistence models, given the insufficient number of persistent cases available across different sites. Preprocessing steps included scaling to a range of 0 to 1, regression of covariates such as age, sex and symptom severity (*PANSS* total score), modality-specific site correction (based on generalizability theory-based voxel-level reliability maps (44) (Figure S3) and site correction of the voxel-level data based on the site means as detailed in Text S3), dimensionality reduction using principal component analysis, and standardization of the obtained components. Following preprocessing, we trained binary L2-regularized support vector machine classifiers with a Maximum Relevance Minimum Redundancy filter for feature selection to discriminate between the high and low FThD groups at baseline and between persistent vs. non-persistent high-FThD based on: 1) each of the

three fALFF sub-bands separately (slow-5, slow-4, slow-3), 2) the GMV maps, and 3) the WMV maps. Lastly, we trained stacked multimodal SVM models on the classifier scores derived from the models trained on the individual data modalities that performed above chance level, to assess the combined discriminative performance of structural and functional brain images. The models' significance was assessed using a label permutation approach and the obtained *P*-values were corrected using the false discovery rate to account for the multiple classifiers trained. The predictive features were visualized using a combination of feature importance (sign-based consistency) and feature stability (cross-validation ratio) metrics, as detailed in Text S3. Additionally, the performance differences between single-modality classifiers and the stacked multimodal models were tested for statistical significance by means of Quade tests followed by post-hoc pairwise classifier comparisons (45).

Post-hoc, we investigated the associations between the decision scores of significant classifiers using Pearson's correlation coefficients. Moreover, we explored the potential influence of medication on the models by computing Pearson's correlation coefficients between the significant classifiers' decision scores and antipsychotic medication converted into chlorpromazine equivalents as well as cumulative dosages of selective serotonin reuptake inhibitors (SSRI).

Supplementary analyses were conducted to evaluate possible methodological moderators of classification performance: 1) all main models were similarly trained on the non-standardized fALFF data; 2) all main models were trained without correcting for total symptomatology (*PANSS* total score); 3) FThD persistence prediction models were trained within an outer leave-site-out cross-validation framework despite the small sample size to further evaluate the impact of site effects; 4) potential differential relationships between positive and negative FThD subscores and the models' decision scores were evaluated using slope interaction analyses (Figure S4).

3. Results

3.1. Sociodemographic and clinical characteristics of the FThD subgroups

Sociodemographic and clinical comparisons of the two baseline FThD-based ROP subgroups, as well as between the patients with persistent and non-persistent high FThD are presented in Table 1. Briefly, high FThD patients at baseline were younger than those with low FThD (t(df)=2.46(231), p=.02), were more likely to have a stable partnership ($\chi 2$(df)=8.49(6), p=.02), but were similar in terms of other sociodemographic characteristics (Table 1). High FThD patients had higher levels of positive *PANSS* score (t(df)=-4.76(230), p<.001), negative *PANSS* score (t(df)=-10.93(230), p<.001), and *PANSS* general symptoms score (t(df)=-

4.09(226), p<.001). Lastly, high FThD patients had higher scores on all *SANS* subscales (Blunting: t(df)=-8.82(231), p<.001; Alogia: t(df)=-18.52(231), p<.001; Apathy: t(df)=-4.67(231), p<.001; Anhedonia: t(df)=-4.34(222), p<.001; Attention: t(df)=-6.85 205), p<.001)(Table 1).

Furthermore, patients with persistently high FThD were younger (t(df)=-2.05(151), p=.04), had fewer education years (t(df)=-3.87(151), p<.001), and were less likely to have worked in the past year ($\chi 2$(df)=-3.87(151), p=.003). Also, persistently high FThD patients had higher levels of general (t(df)=2.80(151), p=.006) and negative *PANSS* symptoms (t(df)=5.10(151), p<.001) and higher levels of affective blunting (t(df)=4.07(151), p<.001), alogia (t(df)=5.05(151), p<.001), apathy (t(df)=2.14(151), p=.03) and attention deficits (t(df)=4.45(134), p<.001) according to the *SANS* (Table 1).

3.2. Cross-sectional analyses

At baseline, high and low FThD subgroups could only be separated above chance level using the GMV data (Balanced Accuracy (BAC)=60.8%, Sensitivity=63.8%, Specificity=57.7%, p_{FDR}=.002, Table 2). The GMV pattern predicted high FThD group membership based on higher GMV in cingulate cortex regions pertaining to the prefrontal control network and dorsal attentional network as well as lower GMV in the visual network (Figures 1 and 2).

3.3. FThD persistence analyses

All imaging data modalities performed significantly above chance level in separating patients with persistently high FThD from those with other FThD symptom courses (slow-5 fALFF: BAC=73.2%, Sensitivity=83.3%, Specificity=63.1%, p_{FDR}<.001; slow-4 fALFF: BAC=72.3%, Sensitivity=83.3%, Specificity=62.4%, p_{FDR}<.001; slow-3 fALFF: BAC=68.0%, Sensitivity=75.0%, Specificity=61.0%, p_{FDR}<.001; GMV: BAC=62.7%, Sensitivity=75.0%, Specificity=50.4%, p_{FDR}=.048; WMV: BAC=73.1%, Sensitivity=91.7%, Specificity=54.6%, p_{FDR}<.001). Stacking all data modalities generated a performance of BAC=77.0% (Sensitivity=100%, Specificity=53.9%, p_{FDR}=.048), which was significantly higher than that of the GMV-based model, but not that of the fALFF- or WMV-based models, as evidenced by a Quade test followed by post-hoc comparisons (Table S4). Moreover, supplementary leave-site-out analyses showed that the slow-5-, slow-3- and WMV-based models remained statistically significant with a maximal multimodal BAC of the stacked model of 69% (Table S5).

fALFF predictive patterns of high FThD persistence consisted of distributed deactivations and activations within large-scale brain networks, such as the default-mode network, dorsal attention network, and salience network, as well as the cerebellum, with specificity for

frequency sub-bands (Figure 2, Figure 3A,B,C). Moreover, high FThD persistence was predicted by higher GMV within the dorsal attention and salience networks and lower GMV within regions of the ventral attention and visual networks (Figures 2 and 3D). Lastly, lower WMV within frontal tracts and higher WMV within subcortical tracts predicted the persistence of high FThD severity (Figure 3E).

3.4. Post-hoc analyses

Correlation analyses showed that the decision scores of the cross-sectional GMV-based model were positively correlated with those of the slow-4-based model predicting high FThD persistence ($r(151)$=.20, p_{FDR}<.05), as well as with those of the GMV- ($r(151)$=.60, p_{FDR}<.001) and WMV-based ($r(151)$=.30, p_{FDR}<.001) FThD persistence prediction models (Table S6). Moreover, we found significant positive correlations between the decisions scores of the persistence models trained on the different modalities (Table S6).

Furthermore, there were no significant correlations between antipsychotic medication and the decision scores of any of the significant cross-sectional models (GMV: $r(231)$=-0.09, p=.26) or any of the persistence classifiers (Slow-5: $r(151)$=-.001, p=.98; Slow-4: $r(151)$=-.19, p=.06; Slow-3: $r(151)$=-.11, p=.26; GMV: $r(151)$=-.01, p=.98; WMV: $r(151)$=-.15, p=.18). Similarly, no significant correlations were found between cumulative dosage of SSRI medication and the decision scores of the cross-sectional GMV model (r(231)=-0.12, p=.071). None of the patients included in the persistence analyses had been medicated with SSRIs.

Using the non-standardized fALFF did not have an influence on which models performed above chance-level but led to performance decreases of the FThD persistence prediction models (Table S7). Furthermore, removing the *PANSS* total score correction from the pipeline additionally led to the slow-5 fALFF sub-band being cross-sectionally predictive of high vs low FThD at baseline, without influencing the other significant results of the *PANSS*-corrected models (Table S8). Lastly, the slope interaction analyses showed similar relationships between the decision scores of all significant models and the positive and negative FThD subdimension scores (Figure S4).

4. Discussion

The current study provides preliminary evidence that FThD severity in recent-onset psychosis (ROP) can be cross-sectionally delineated and longitudinally predicted by multivariate structural and functional brain patterns spanning large-scale brain networks. Specifically, we show that GMV changes differentiate cross-sectionally between high vs. low FThD ROP subgroups with a relatively low BAC of 60.8%. Furthermore, the baseline activity in all fALFF

sub-bands, as well as GMV and WMV patterns predicted the persistence of high FThD severity over one year with BACs of up to 77%.

The cross-sectional results describing the separability of FThD severity subgroups based on GMV data are in line with univariate correlational studies showing associations between FThD symptoms and GMV alterations in schizophrenia (9). The GMV increases within fronto-cingulate regions and decreases localized within occipital regions, predictive of high FThD in our study, are consistent with previously findings in patients with schizophrenia (46,47). However, in contrast to other studies (9,15), we did not find prominent GMV decreases within the language networks to be particularly predictive of high-FThD in our sample. In addition to the multivariate methodology employed here, these inconsistencies could highlight the specificity of the predictive patterns reported here for delineating FThD heterogeneity in the early stages of psychosis, where the influence of general illness chronicity and medication effects are more limited. Our cross-sectional findings provide preliminary evidence for an FThD-discriminative pattern of limited clinical utility. Future studies showing cross-sectional classifications with higher specificity and sensitivity are needed to reach individual level clinical implications.

In contrast to the limited cross-sectional findings, we found that all structural and functional neuroimaging modalities investigated were predictive of 1-year high-FThD persistence. Volumetric brain patterns included GMV reductions/increases that partially overlapped with those found to be predictive of high-FThD at the cross-sectional level, as well as pronounced WMV decreases in frontal tracts. Furthermore, distributed functional alterations within the fALFF sub-bands across large-scale brain networks were predictive of high-FThD persistence. These functional patterns overlapped with those identified in previous studies, highlighting alterations of language, executive and default-mode networks as a possible substrate for positive (i.e., conceptual disorganization) and negative FThD symptoms (i.e., alogia, difficulty in abstract thinking)(20,24,25,27). Together with the WMV findings, our results suggest that connectivity patterns within and between long-range brain networks may represent relevant early biomarkers of FThD symptom progression, in line with the *disconnection hypothesis* of schizophrenia (48). Future studies may benefit from exploring the prognostic value of additional measures such as dynamic functional connectivity and diffusion tensor imaging, which provide more fine-grained information on the function and structure of these neural systems.

The current study addresses an important gap in the FThD stratification literature, providing a novel exploration of the predictive value of multi-modal neuroimaging data for disentangling the heterogeneity of FThD severity and persistence in ROP using machine learning methods

operating in a multi-site context. Collectively, our findings support the utility of symptom-based hierarchical approaches in psychiatric nosology (2) for identifying specific brain systems associated with psychopathological dimensions such as FThD, with noteworthy implications for personalized diagnosis and intervention. Firstly, the predictive patterns reported provide exploratory pathophysiological evidence linking specific brain networks to FThD severity and persistence. Following validation in future clinical studies, such imaging-based biomarkers could translate into targets for novel treatments (e.g. brain stimulation-enhanced neurocognitive training), tailored to patients suffering from severe FThD. Secondly, we provide preliminary evidence for the prognostic utility of neuroimaging measures for the stratification of FThD persistence, which has been associated with poor clinical outcomes (7). Although the small sample size available for this analysis calls for future validation studies, our results encourage the future development of neuroimaging-based prognostic tools for the early recognition of patients at-risk for progressing to chronic FThD syndromes, as a first step towards the indicated prevention of such debilitating disease trajectories.

In addition to the reduced sample size and lack of an external replication sample, which limit claims regarding the out-of-sample generalizability of present findings, several limitations of our study are noteworthy. Firstly, we did not use a construct-specific FThD scale, but instead used FThD subdomains corresponding to the *Thought and Language Disorder (TALD*)(3) scale (conceptual disorganization, poverty of content of speech, difficulty in abstract thinking, poverty of speech and increased latency of response), as measured by *PANSS* and *SANS* items. Although having an FThD-proxy within widely used clinical scales can facilitate clinical translation, further studies are needed to investigate FThD more broadly through validated scales (i.e. *TALD* or speech sample analyses(49)), to further evaluate the specificity of our findings for the FThD construct. Secondly, regarding FThD progression, we restricted our analyses to persistently high FThD symptom courses. The chronicity of FThD symptoms in comparison to single-time measurements has been associated with particularly poor clinical outcomes (7) and can be regarded as a more reliable prognostic marker of FThD severity for preventive clinical applications. This is further supported by our models, which performed with higher sensitivity in predicting high FThD persistence compared to single timepoint severity. Nonetheless, future studies may benefit from a manifold investigation of FThD longitudinal dynamics which may unravel symptom course patterns delineated by distinct structural and functional brain alterations.

6. Conclusion

Our findings align with modern symptom-based psychiatric frameworks by providing first evidence for multivariate structural and functional brain alterations within large-scale brain

networks predictive of FThD severity and its persistence in recent-onset psychosis. These preliminary results open the avenue for the development of refined early spectrum-wise diagnostics with preventive and long-term clinical outcome-oriented therapeutics, guided by and involving integrated neuroimaging data domains within larger multi-site samples.

Acknowledgements

Disclosures of Conflicts of Interest

Dr Bertolino reports speaker fees from Otsuka, Lundbeck, Angelini and Rovi outside of the submitted work. Dr Haas has received support by the NIH National Institute of Mental Health, grant T32MH122394, unrelated to the submitted work. Dr Hietala reports personal fees from Orion ltd, personal fees from Lundbeck, personal fees from Otsuka and other from Takeda during the conduct of the study. Dr Koutsouleris, Dr Ruhrmann, Dr Riecher-Rossler report grants from European Union over the duration of the study. Dr Meisenzahl and Dr Koutsouleris hold patent US20160192889A1 ('Adaptive pattern recognition for psychosis risk modelling'). Dr Koutsouleris reports speaker fees from Otsuka, Roche and Angelini outside of the submitted work. Dr Pantelis reports grants from Australian NHMRC during the study, and personal fees from Lundbeck, Australia Pty Ltd outside the submitted work. Dr Upthegrove reports speaker fees from Sunovion, Otsuka and Vitaris outside the submitted work as well as unpaid officership with the British Association for Pharmacology - Honorary General Secretary 2021-2024. She serves as Deputy Editor for The British Journal of Psychiatry. Dr Falkai reports he has received research support/honoraria for lectures or advisory activities from Boehringer-Ingelheim, Janssen, Lundbeck, Otsuka, Recordati and Richter outside the submitted work. Dr Pantelis was supported by an Australian National Health and Medical Research Council (NHMRC) L3 Investigator Grant (1196508) outside the submitted work. Dr Lana Kambeitz-Ilankovic reports receiving a NARSAD Young Investigator Award of the Brain & Behavior Research Foundation No◦ 28474 (PI: LK-I) outside the submitted work.

All other co-authors did not report any other financial disclosures or other conflicts of interests within or outside of the scope of the submitted work.

Funding Sources

PRONIA is a Collaboration Project funded by the European Union under the 7th Framework Programme under grant agreement n° 602152. Oemer Faruk Oeztuerk and David Popovic were supported by the Else-Kröner-Fresenius-Foundation through the Clinician Scientist Program 'EKFS-Translational Psychiatry'.

The PRONIA consortium

The following authors participated in the screening, recruitment, rating, examination, and follow-up of the study participants used in the current analyses. The authors are listed in alphabetical order according to the institution of affiliation.

Department of Psychiatry and Psychotherapy, Ludwig-Maximilian-University, Munich, Germany: Shalaila Haas, Alkomiet Hasan, Claudius Hoff, Ifrah Khanyaree, Camilla Krämer, Aylin Melo, Susanna Muckenhuber-Sternbauer, Yanis Köhler, Oemer Faruk Oeztuerk, Nora Penzel, David Popovic, Adrian Rangnick, Sebastian von Saldern, Rachele Sanfelici, Moritz Spangemacher, Ana Tupac, Maria Fernanda Urquijo, Johanna Weiske, Antonia Wosgien.

Department of Psychiatry and Psychotherapy, University of Cologne, Cologne, Germany: Karsten Blume, Dennis Hedderich, Dominika Julkowski, Nathalie Kaiser, Thorsten Lichtenstein, Ruth Milz, Alexandra Nikolaides, Tanja Pilgram, Mauro Seves, Martina Wassen.

Department of Psychiatry (Psychiatric University Hospital, UPK), University of Basel, Switzerland: Christina Andreou, Laura Egloff, Fabienne Harrisberger, Ulrike Heitz, Claudia Lenz, Letizia Leanza, Amatya Mackintosh, Renata Smieskova, Erich Studerus, Anna Walter, Sonja Widmayer.

Institute of Mental Health & School of Psychology, University of Birmingham, United Kingdom: Chris Day, Sian Lowri Griffiths, Mariam Iqbal, Mirabel Pelton, Pavan Mallikarjun, Alexandra Stainton, Ashleigh Lin.
Department of Psychiatry, University of Turku, Finland: Alexander Denissoff, Anu Ellilä, Tiina From, Markus Heinimaa, Tuula Ilonen, Päivi Jalo, Heikki Laurikainen, Antti Luutonen, Akseli Mäkela, Janina Paju, Henri Pesonen, Reetta-Liina Säilä, Anna Toivonen, Otto Turtonen.
Department of Psychiatry (Psychiatric University Hospital LVR/HHU Düsseldorf), University of Düsseldorf, Germany: Sonja Botterweck, Norman Kluthausen, Gerald Antoch, Julian Caspers, Hans-Jörg Wittsack.
Department of Basic Medical Science, Neuroscience and Sense Organs - University of Bari Aldo Moro: Giuseppe Blasi, Giulio Pergola, Grazia Caforio, Leonardo Fazio, Tiziana Quarto, Barbara Gelao, Raffaella Romano, Ileana Andriola, Andrea Falsetti, Marina Barone, Roberta Passiatore, Marina Sangiuliano.
Department of Psychiatry and Psychotherapy of the University of Münster, Germany: Marian Surmann, Olga Bienek, Udo Dannlowski.
General Electric Global Research Inc., USA.
Ana Beatriz Solana, Manuela Abraham, Timo Schirmer.
Workgroup of Paolo Brambilla, University of Milan, Italy:
Department of Neuroscience and Mental Health, Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico, University of Milan, Milan, Italy: Carlo Altamura, Marika Belleri, Francesca Bottinelli, Adele Ferro, Marta Re
Programma2000, Niguarda Hospital, Milan: Emiliano Monzani, Maurizio Sberna
San Paolo Hospital, Milan: Armando D'Agostino, Lorenzo Del Fabro
Villa San Benedetto Menni, Albese con Cassano (CO): Giampaolo Perna, Maria Nobile, Alessandra Alciati
Workgroup of Paolo Brambilla at the University of Udine, Italy.
Department of Medical Area, University of Udine, Udine, Italy: Matteo Balestrieri, Carolina Bonivento, Giuseppe Cabras, Franco Fabbro.
IRCCS Scientific Institute "E. Medea", Polo FVG, Udine: Marco Garzitto, Sara Piccin.

References

1. Insel T, Cuthbert B, Garvey M, Heinssen R, Pine DS, Quinn K, et al. Research domain criteria (RDoC): toward a new classification framework for research on mental disorders. Vol. 167, American Journal of psychiatry. Am Psychiatric Assoc; 2010. p. 748–51.
2. Kotov R, Krueger RF, Watson D, Achenbach TM, Althoff RR, Bagby RM, et al. The Hierarchical Taxonomy of Psychopathology (HiTOP): A dimensional alternative to traditional nosologies. J Abnorm Psychol. 2017;126(4):454.
3. Kircher T, Krug A, Stratmann M, Ghazi S, Schales C, Frauenheim M, et al. A rating scale for the assessment of objective and subjective formal thought and language disorder (TALD). Schizophr Res. 2014;160(1–3):216–21.
4. Roche E, Creed L, Macmahon D, Brennan D, Clarke M. The Epidemiology and Associated Phenomenology of Formal Thought Disorder: A Systematic Review. Schizophr Bull. 2015;41(4):951–62.
5. Andreasen NC, Grove WM. Thought, language, and communication in schizophrenia: diagnosis and prognosis. Schizophr Bull. 1986;12(3):348–59.
6. Hart M, Lewine RRJ. Rethinking thought disorder. Schizophr Bull. 2017;43(3):514–22.
7. Marengo JT, Harrow M. Schizophrenic Thought Disorder at Follow-up. Arch Gen Psychiatry. 1987;44.
8. Sigaudo M, Crivelli B, Castagna F, Giugiario M, Mingrone C, Montemagni C, et al. Quality of life in stable schizophrenia: The relative contributions of disorganization and cognitive dysfunction. Schizophr Res. 2014;153(1–3):196–203.
9. Cavelti M, Kircher T, Nagels A, Strik W, Homan P. Is formal thought disorder in schizophrenia related to structural and functional aberrations in the language network? A systematic review of neuroimaging findings. Schizophr Res. 2018;199:2–16.
10. Oeztuerk OF, Pigoni A, Antonucci LA, Koutsouleris N. Association between formal thought disorders, neurocognition and functioning in the early stages of psychosis: a systematic review of the last half-century studies. Vol. 272, European Archives of Psychiatry and Clinical Neuroscience. Springer Science and Business Media Deutschland GmbH; 2022. p. 381–93.
11. Remberk B, Namysłowska I, Rybakowski F. Clinical and cognitive correlates of formal thought disorder in early onset schizophrenia. Activitas Nervosa Superior Rediviva. 2012;54(2):68–76.
12. Roche E, Lyne J, O'Donoghue B, Segurado R, Behan C, Renwick L, et al. The prognostic value of formal thought disorder following first episode psychosis. Schizophr Res. 2016;178(1–3):29–34.
13. Oeztuerk OF, Pigoni A, Wenzel J, Haas SS, Popovic D, Ruef A, et al. The clinical relevance of formal thought disorder in the early stages of psychosis: results from the PRONIA study. Eur Arch Psychiatry Clin Neurosci. 2021 Apr 1;
14. Harrow M, Marengo JT. Schizophrenic Thought Disorder at Followup: Its Persistence and Prognostic Significance. Schizophr Bull. 1986 Jan 1;12(3):373–93.
15. Sans-Sansa B, McKenna PJ, Canales-Rodríguez EJ, Ortiz-Gil J, López-Araquistain L, Sarró S, et al. Association of formal thought disorder in schizophrenia with structural brain abnormalities in language-related cortical regions. Schizophr Res. 2013;146(1–3):308–13.
16. Sumner PJ, Bell IH, Rossell SL. A systematic review of the structural neuroimaging correlates of thought disorder. Neurosci Biobehav Rev. 2018;84:299–315.

17. Cavelti M, Winkelbeiner S, Federspiel A, Walther S, Stegmayer K, Giezendanner S, et al. Formal thought disorder is related to aberrations in language-related white matter tracts in patients with schizophrenia. Psychiatry Res Neuroimaging. 2018;279(November 2017):40–50.
18. Stein F, Buckenmayer E, Brosch K, Meller T, Schmitt S, Ringwald KG, et al. Dimensions of Formal Thought Disorder and Their Relation to Gray- and White Matter Brain Structure in Affective and Psychotic Disorders. Schizophr Bull. 2022;1–10.
19. Vita A, Dieci M, Giobbio GM, Caputo A, Ghiringhelli L, Comazzi M, et al. Language and thought disorder in schizophrenia: brain morphological correlates. Schizophr Res. 1995;15(3):243–51.
20. Horn H, Jann K, Federspiel A, Walther S, Wiest R, Müller T, et al. Semantic network disconnection in formal thought disorder. Neuropsychobiology. 2012 Jul;66(1):14–23.
21. Kircher TTJ, Bulimore ET, Brammer MJ, Williams SCR, Broome MR, Murray RM, et al. Differential activation of temporal cortex during sentence completion in schizophrenic patients with and without formal thought disorder [Internet]. 2001. Available from: www.elsevier.com/locate/schres
22. Sumner PJ, Bell IH, Rossell SL. A systematic review of task-based functional neuroimaging studies investigating language, semantic and executive processes in thought disorder. Neurosci Biobehav Rev. 2018;94:59–75.
23. Wensing T, Cieslik EC, Müller VI, Hoffstaedter F, Eickhoff SB, Nickl-Jockschat T. Neural correlates of formal thought disorder: An activation likelihood estimation meta-analysis. Hum Brain Mapp. 2017;38(10):4946–65.
24. Chen J, Wensing T, Hoffstaedter F, Cieslik EC, Müller VI, Patil KR, et al. Neurobiological substrates of the positive formal thought disorder in schizophrenia revealed by seed connectome-based predictive modeling. Neuroimage Clin. 2021 Jan 1;30.
25. Fuentes-Claramonte P, López-Araquistain L, Sarró S, Sans-Sansa B, Ortiz-Gil J, Maristany T, et al. Brain functional correlates of formal thought disorder in schizophrenia: Examining the frontal/dysexecutive hypothesis. Psychol Med. 2021;51(14):2446–53.
26. Kircher TTJ, Liddle ; P F, Brammer MJ, Williams SCR, Murray RM, Mcguire PK. Neural Correlates of Formal Thought Disorder in Schizophrenia Preliminary Findings From a Functional Magnetic Resonance Imaging Study. 2001.
27. Winkelbeiner S, Cavelti M, Federspiel A, Kunzelmann K, Dierks T, Strik W, et al. Decreased blood flow in the right insula and middle temporal gyrus predicts negative formal thought disorder in schizophrenia. Schizophr Res. 2018;9–11.
28. Zang YF, Yong H, Chao-Zhe Z, Qing-Jiu C, Man-Qiu S, Meng L, et al. Altered baseline brain activity in children with ADHD revealed by resting-state functional MRI. Brain Dev. 2007 Mar;29(2):83–91.
29. Fu Z, Tu Y, Di X, Du Y, Pearlson GD, Turner JA, et al. Characterizing dynamic amplitude of low-frequency fluctuation and its relationship with dynamic functional connectivity: An application to schizophrenia. Vol. 180, NeuroImage. Academic Press Inc.; 2018. p. 619–31.
30. Tomasi D, Shokri-Kojori E, Volkow ND. Temporal changes in local functional connectivity density reflect the temporal variability of the amplitude of low frequency fluctuations in gray matter. PLoS One. 2016 Apr 1;11(4).
31. Biswal B, Yetkin FZ, Haughton VM, Hyde JS. Functional Connectivity in the Motor Cortex of Resting Human Brain Using Echo-Planar MRI. 1995.

32. Liang X, Wang J, Yan C, Shu N, Xu K, Gong G, et al. Effects of different correlation metrics and preprocessing factors on small-world brain functional networks: A resting-state functional MRI study. PLoS One. 2012 Mar 6;7(3).
33. Yu R, Chien YL, Wang HLS, Liu CM, Liu CC, Hwang TJ, et al. Frequency-specific alternations in the amplitude of low-frequency fluctuations in schizophrenia. Hum Brain Mapp. 2014 Feb;35(2):627–37.
34. Hoptman MJ, Zuo XN, Butler PD, Javitt DC, D'Angelo D, Mauro CJ, et al. Amplitude of low-frequency oscillations in schizophrenia: A resting state fMRI study. Schizophr Res. 2010 Mar;117(1):13–20.
35. Falkai P, Gaebel W, Weinmann S, Wobrock T. DGPPN (Hrsg.) S3–Praxisleitlinien in Psychiatrie und Psychotherapie Band 1: Behandlungsleitlinie Schizophrenie. 2006.
36. Kay, S R Fiszbein A L a O. Positive and Negative Syndrome Scale (Panss) Rating Criteria. Schizophr Bull [Internet]. 1987;13(2):261–76. Available from: www.ncbi.nlm.nih.gov
37. Andreasen NC. The scale for the assessment of negative symptoms (SANS). British Journal of Psychiatry. 1989;155:49–52.
38. Burock G, Pihur V, Datta S, Datta S. clValid: An R Package for Cluster Validation. Journal of Statictical Software. 2008;25(4).
39. Koutsouleris N, Kambeitz-Ilankovic L, Ruhrmann S, Rosen M, Ruef A, Dwyer DB, et al. Prediction models of functional outcomes for individuals in the clinical high-risk state for psychosis or with recent-onset depression: a multimodal, multisite machine learning analysis. JAMA Psychiatry. 2018;75(11):1156–72.
40. Koutsouleris N, Dwyer DB, Degenhardt F, Maj C, Urquijo-Castro MF, Sanfelici R, et al. Multimodal Machine Learning Workflows for Prediction of Psychosis in Patients with Clinical High-Risk Syndromes and Recent-Onset Depression. JAMA Psychiatry. 2021 Feb 1;78(2):195–209.
41. Haas SS, Antonucci LA, Wenzel J, Ruef A, Biagianti B, Paolini M, et al. A multivariate neuromonitoring approach to neuroplasticity-based computerized cognitive training in recent onset psychosis. Neuropsychopharmacology. 2021 Mar 1;46(4):828–35.
42. Power JD, Mitra A, Laumann TO, Snyder AZ, Schlaggar BL, Petersen SE. Methods to detect, characterize, and remove motion artifact in resting state fMRI. Neuroimage. 2014;84:320–41.
43. Zuo XN, Di Martino A, Kelly C, Shehzad ZE, Gee DG, Klein DF, et al. The oscillating brain: complex and reliable. Neuroimage. 2010;49(2):1432–45.
44. Cronbach LJ, Rajaratnam N, Gleser GC. Theory of generalizability: A liberalization of reliability theory. British Journal of Statistical Psychology. 1963;16(2):137–63.
45. Quade D. Using weighted rankings in the analysis of complete blocks with additive block effects. J Am Stat Assoc. 1979;74(367):680–3.
46. Palaniyappan L, Mahmood J, Balain V, Mougin O, Gowland PA, Liddle PF. Structural correlates of formal thought disorder in schizophrenia: An ultra-high field multivariate morphometry study. Schizophr Res. 2015 Oct 1;168(1–2):305–12.
47. Stein F, Meller T, Brosch K, Schmitt S, Ringwald K, Pfarr JK, et al. Psychopathological Syndromes across Affective and Psychotic Disorders Correlate with Gray Matter Volumes. Schizophr Bull. 2021;47(6):1740–50.
48. Friston KJ. The disconnection hypothesis. Schizophr Res. 1998;30(2):115–25.
49. Çokal D, Sevilla G, Jones WS, Zimmerer V, Deamer F, Douglas M, et al. The language profile of formal thought disorder. NPJ Schizophr. 2018 Dec 1;4(1).
50. Beck AT, Steer RA, Brown G. Beck depression inventory–II. Psychol Assess. 1996;

51. Yeo BTT, Krienen FM, Sepulcre J, Sabuncu MR, Lashkari D, Hollinshead M, et al. The organization of the human cerebral cortex estimated by intrinsic functional connectivity. J Neurophysiol. 2011;

Table 1. Sociodemographic and clinical comparison of the high vs low FThD subgroups at baseline and of the persistence-based subgroups.

	Baseline subgroup membership				Baseline to follow-up high FThD persistence vs. other FThD courses			
	High FThD Subgroup	Low FThD Subgroup	t/χ2 value (df)	*P*-value	Persistent High FThD	Non-persistent High FThD	t/χ2 value (df)	*P*-value
Socio-demographic characteristics								
Participants, N	58	175			12	141		
Participants per site, N (%)								
Munich	17 (29.3)	66 (37.7)	1.34 (4)	.25	4 (33.3)	47 (33.3)	0.00 (4)	1.0
Turku	7 (12.1)	31 (17.7)	1.02 (4)	.31	2 (16.7)	23 (16.3)	0.00 (4)	.98
Milan	10 (17.2)	14 (8.0)	4.03 (4)	**.05**	4 (33.3)	13 (9.2)	6.51 (4)	**.01**
Birmingham	2 (3.4)	8 (4.6)	0.13 (4)	.71	0 (0.0)	8 (5.7)	0.72 (4)	.40
Basel	4 (6.9)	18 (10.3)	0.59 (4)	.44	0 (0.0)	17 (12.1)	1.63 (4)	.20
Udine	2 (3.4)	7 (4.0)	0.04 (4)	.85	0 (0.0)	6 (4.3)	0.53 (4)	.47
Cologne	13 (22.4)	25 (14.3)	2.11 (4)	.15	0 (0.0)	23 (16.3)	2.30 (4)	.13
Münster	3 (5.2)	6 (3.4)	0.36 (4)	.55	2 (16.7)	4 (2.8)	5.61 (4)	**.02**
Age, yrs., mean (SD)	23.9 (5.5)	26.0 (5.7)	2.46 (231)	**.02**	22.3 (5.2)	25.8 (5.7)	-2.05 (151)	**.04**
Male sex, N (%)	36 (62.1)	96 (54.9)	0.92 (4)	.34	6 (50.0)	81 (57.4)	0.25 (4)	.61
Education yrs., mean (SD)	12.9 (3.1)	14.9 (9.0)	1.65 (231,0)	.10	10.6 (2.3)	14.2 (3.2)	-3.87 (151)	**<.001**
Steady partnership/ married, N (%)	53 (91.4)	129 (73.7)	8.49 (6)	**.01**	11 (91.7)	109 (77.3)	1.39 (6)	.49
Work in the last year, N (%)	17 (31.5)	70 (41.4)	3.37 (6)	.19	3 (25.0)	58 (41.7)	11.51 (6)	**.003**
First-degree relatives with psychosis, N (%)	1.3 (1.8)	1.1 (1.6)	-1.08 (231)	.28	2 (16.7)	15 (10.6)	1.0 (16)	.99
Birth Complications, N (%)	12 (20.7)	29 (16.7)	0.48 (4)	.49	2 (16.7)	21 (15.0)	0.0 (4)	.88
Psychopharma-cological medication, N (%)	53 (91.4)	148 (85.1)	1.50 (4)	.22	11 (91.7)	120 (85.7)	0.33 (4)	.56
Chlorpromazine-equivalent antipsychotic dosages (mg), mean(SD)	650.8 (4082.5)	1225.5 (3932.4)	-0.94 (231,0)	.35	264.1 (207.6)	807.4 (4586.1)	-0.41 (151)	.68
Clinical characteristics								
Beck Depression Inventory, mean (SD)	18.9 (10.8)	20.4 (12.2)	0.75 (199)	.46	16.2 (13.4)	19.7 (11.5)	-0.92 (135)	.36
PANSS General, mean (SD)	42.4 (10.4)	32.2 (9.0)	-7.21 (230)	**<.001**	42.6 (8.1)	34.0 (10.3)	2.80 (151)	**.006**
PANSS Negative, mean (SD)	23.5 (8.0)	13.3 (5.5)	-10.93 (230)	**<.001**	24.4 (4.9)	14.4 (6.6)	5.10 (151)	**<.001**
PANSS Positive, mean (SD)	21.7 (6.1)	17.4 (5.8)	-4.76 (230)	**<.001**	19.6 (5.2)	18.6 (6.5)	0.49 (151)	.62
SANS Blunting, mean (SD)	17.4 (9.7)	6.3 (7.8)	-8.82 (231)	**<.001**	19.0 (7.5)	7.8 (9.3)	4.07(151)	**<.001**
SANS Alogia, mean (SD)	12.4 (5.7)	1.9 (2.8)	-18.52 (231)	**<.001**	10.8 (5.6)	3.4 (4.8)	5.05 (151)	**<.001**
SANS Apathy, mean (SD)	9.6 (4.8)	6.2 (4.8)	-4.67 (231)	**<.001**	9.4 (5.1)	6.2 (5.0)	2.14 (151)	**.03**

SANS Anhedonia, mean (SD)	13.6 (7.5)	8.6 (7.3)	-4.34 (222)	**<.001**	11.2 (6.9)	8.8 (7.6)	1.03 (143)	.31
SANS Attention, mean (SD)	6.5 (4.1)	2.5 (3.3)	-6.85 (205)	**<.001**	7.5 (2.8)	2.6 (3.5)	4.45 (134)	**<.001**

Note. BDI – Beck Depression Inventory (50); PANSS – Positive and Negative Schizophrenia Symptoms; SANS – Scale for the Assessment of Negative Symptoms.

*Statistically significant *P*-values (at an α = .05) are presented in bold.

Table 2. Performance metrics for all cross-sectional and persistence models predicting high vs low FThD subgroup membership at baseline and differentiating between persistent vs non-persistent high FThD from baseline to follow-up based on multiband fALFF data, gray matter and white matter volume.

Cross-sectional high vs. low FThD subgroup classification								
	Leave-site-out nested cross-validation							
	Sensitivity, %	**Specificity, %**	**BAC, %**	**AUC**	**PPV, %**	**NPV, %**	**PSI**	**P_{FDR} value**
Rs-fMRI slow-5	63.8	48.6	56.2	0.63	29.1	80.2	9.3	.053
Rs-fMRI slow-4	53.4	*55.4*	54.4	0.57	28.4	78.2	6.7	.056
Rs-fMRI slow-3	67.2	45.7	56.5	0.63	29.1	80.8	9.9	.053
GMV	63.8	57.7	60.8	0.66	33.3	82.8	16.1	**.01**
WMV	56.9	51.4	54.2	0.63	28.0	78.3	6.2	.19
Prediction of high FThD persistence (from baseline to 1-year follow-up)								
	Pooled nested cross-validation							
	Sensitivity, %	**Specificity, %**	**BAC, %**	**AUC**	**PPV, %**	**NPV, %**	**PSI**	**P_{FDR} value**
Rs-fMRI slow-5	83.3	63.1	73.2	0.78	16.1	97.8	13.9	**<.001**
Rs-fMRI slow-4	83.3	62.4	72.9	0.83	15.9	97.8	13.7	**<.001**
Rs-fMRI slow-3	75.0	61.0	68.0	0.76	14.1	96.6	10.7	**<.001**
GMV	75.0	50.4	62.7	0.76	11.4	95.9	7.3	**.048**
WMV	91.7	54.6	73.1	0.75	14.7	98.7	13.4	**<.001**
Stacked model – all modalities	100.0	53.9	77.0	0.87	15.6	100.0	15.6	**.048**

Note. P-values were adjusted using the false-discovery rate (FDR) correction for multiple comparisons and the values exceeding the threshold for statistical significance based on the permutation testing procedure employed (α = .05) are presented in bold.

AUC – area under the curve, BAC – Balanced Accuracy, PPV – positive predictive value, NPV – negative predictive value, PSI – prognostic summary index

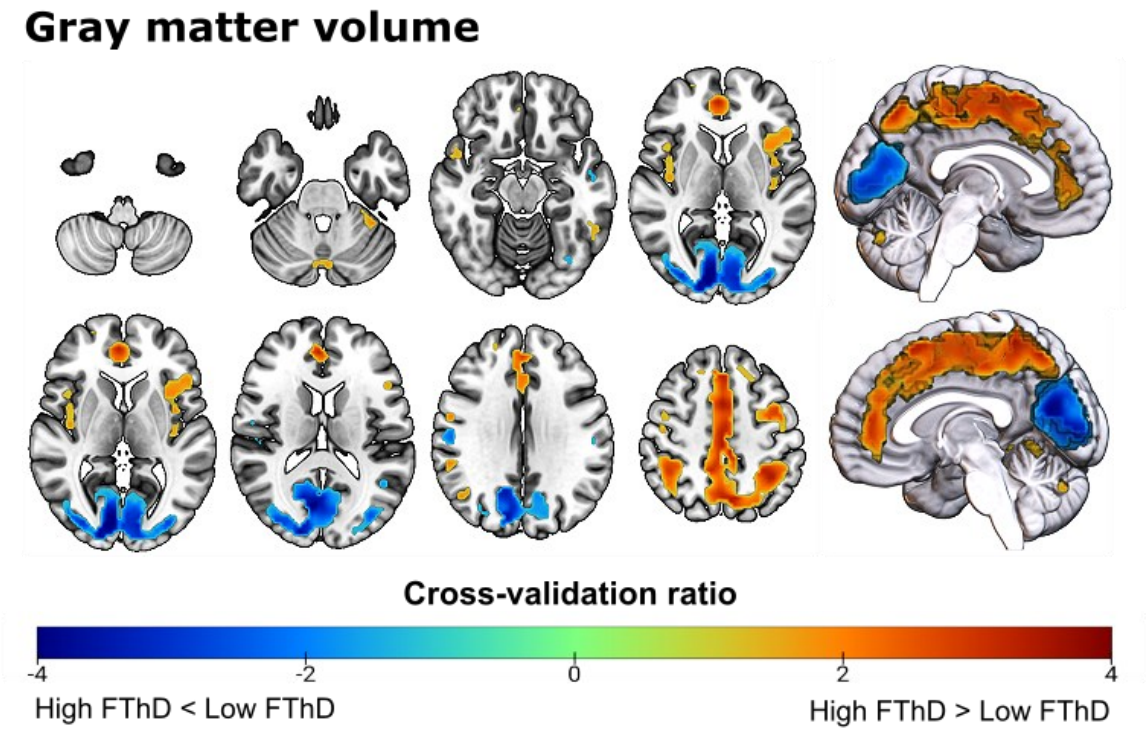

Figure 1. Reliable features for predicting membership to the high vs low formal thought disorder clusters at baseline based on GMV data aquired at the same study visit. The patterns of the other data modalities explored (fALFF sub-bands and WMV) are not presented here, as the models performed at chance level. The reliability of the features is displayed using a grand mean cross-validation ratio, thresholded based on FDR-corrected sign-based consistency maps at α=.05 (detailed in the Supplement). Warm colors represent voxels with larger gray-matter volume for individuals in the high formal thought disorder subgroup, while cold colors indicate smaller volume for this subgroup.

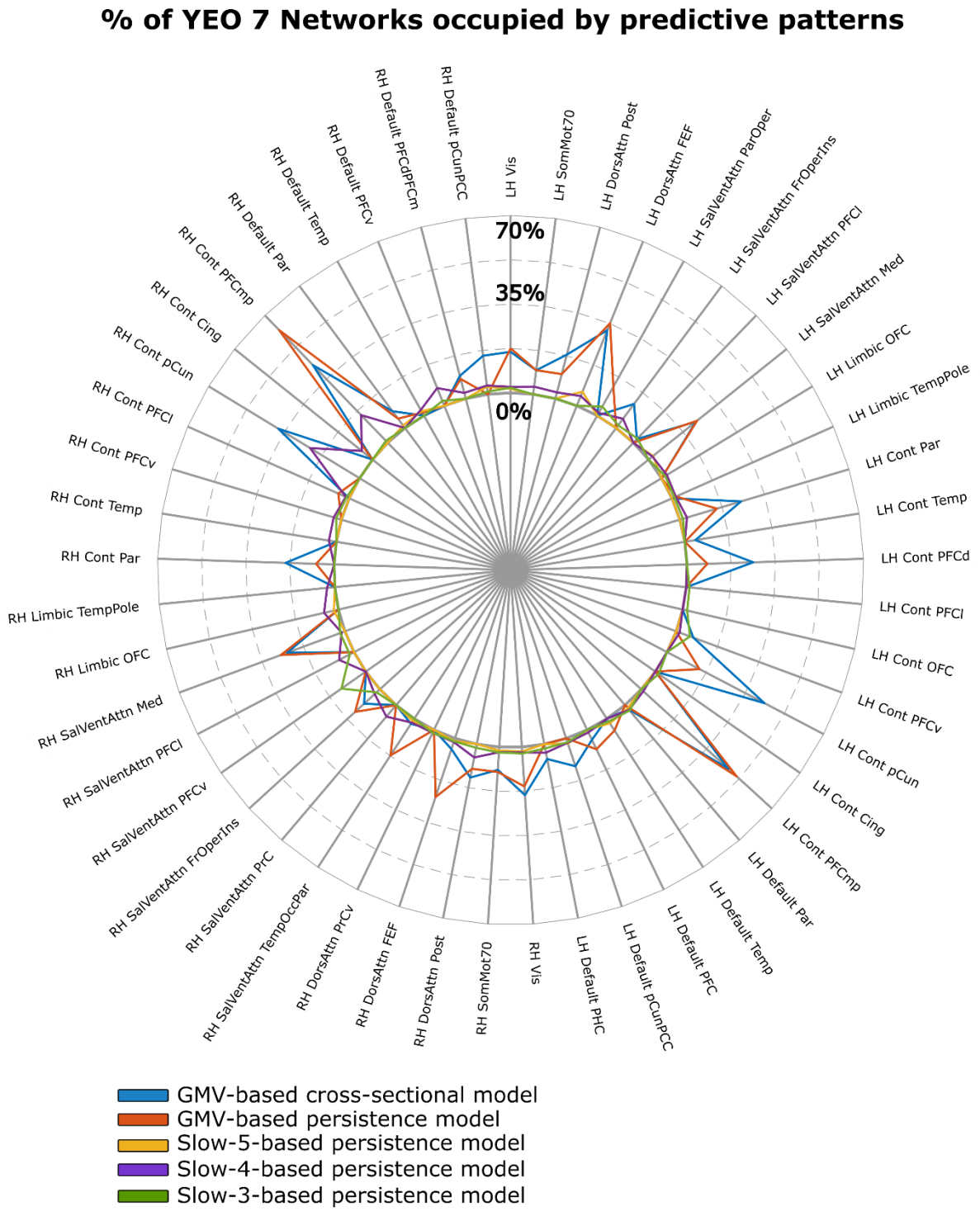

Figure 2. Percentage of the YEO 7 resting-state networks' parcellations (51) **occupied by the predictive patterns of the significant cross-sectional and longitudinal fALFF- and GMV-based models.** Voxels were considered to be predictive based on the derived sign-based consistency metric of feature importance (detailed in the Supplementary methods) thresholded at 1.3, corresponding to α = .05. Only YEO resting-state network parcellations with at least 1% occupation by any of the models' predictive patterns were included in the figure.

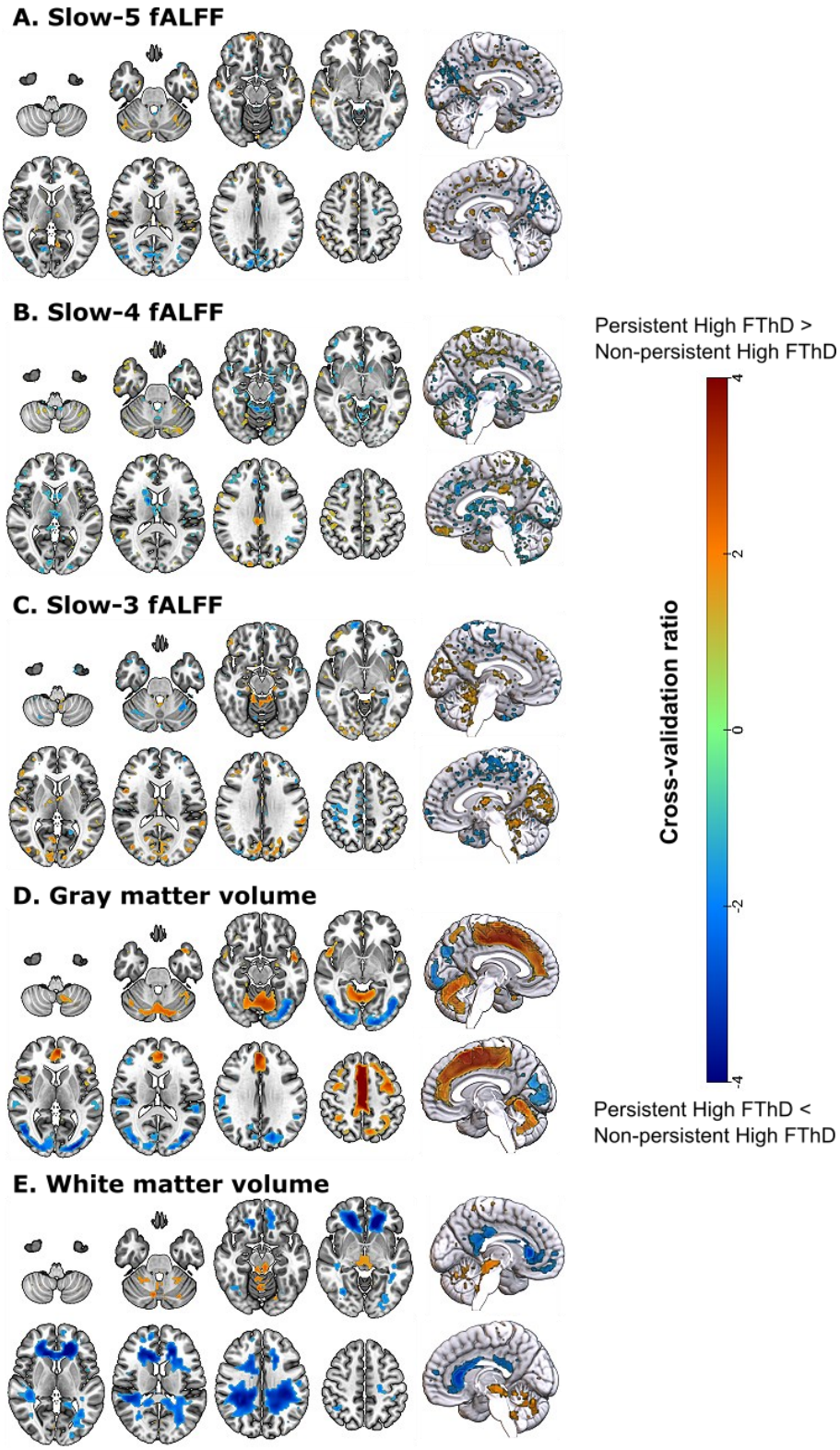

Figure 3. Reliable features for predicting the persistence of high FThD symptomatology from baseline to follow-up relative to other symptom courses based on **A.** slow-5 fALFF data, **B.** slow-4 fALFF data, **C.** slow-3 fALFF data, **D.** gray matter volume data, and **E.** white matter volume data. The reliability of the features is displayed using a grand mean cross-validation ratio, thresholded based on *FDR*-corrected sign-based consistency maps at α=.05 (detailed in Text S3). Warm colors represent voxels with increased activity/volume for individuals with persistently high FThD from baseline to follow-up, while cold colors indicate decreases for this subgroup.